ISSN: 1646-9895

risti

Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Janeiro 19 • January 19

©AISTI 2019 http://www.aisti.eu

Nº E17

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Edição / Edition
Nº. E17, 01/2019

ISSN: 1646-9895

Indexação / Indexing
Academic Journals Database, CiteFactor, Dialnet, DOAJ, DOI, EBSCO, GALE, Index-Copernicus, Index of Information Systems Journals, Latindex, ProQuest, QUALIS, SCImago, SCOPUS, SIS, Ulrich's.

Propriedade e Publicação / Ownership and Publication
AISTI – Associação Ibérica de Sistemas e Tecnologias de Informação
Rua Quinta do Roseiral 76, 4435-209 Rio Tinto, Portugal
E-mail: aistic@gmail.com
Web: http://www.aisti.eu

Director
Álvaro Rocha, Universidade de Coimbra, PT

Coordenadores da Edição / Issue Coordinators
Manolo Paredes, Universidad de las Fuerzas Armadas ESPE, EC
Álvaro Rocha, Universidade de Coimbra, PT

Conselho Editorial / Editorial Board
Carlos Ferrás Sexto, Universidad de Santiago de Compostela, ES
Gonçalo Paiva Dias, Universidade de Aveiro, PT
Jose Antonio Calvo-Manzano Villalón, Universidad Politécnica de Madrid, ES
Manuel Pérez Cota, Universidad de Vigo, ES
Ramiro Gonçalves, Universidade de Trás-os-Montes e Alto Douro, PT

Conselho Científico / Scientific Board
Adolfo Lozano-Tello, Universidad de Extremadura, ES
Adrián Hiebra Pardo, Universidad de Santiago de Compostela, ES
Alberto Fernández, Universidad Rey Juan Carlos, ES
Alberto Bugarín, Universidad de Santiago de Compostela, ES
Alejandro Medina, Universidad Politécnica de Chiapas, MX
Alejandro Rodríguez González, Universidad Politécnica de Madrid, ES
Alejandro Peña, Escuela de Ingeniería de Antioquia, CO
Alexandre L'Erario, Universidade Tecnológica Federal do Paraná, BR
Ana Azevedo, Instituto Politécnico do Porto, PT
Ana Cristina Ramada Paiva, FEUP, Universidade do Porto, PT
Ana Isabel Veloso, Universidade de Aveiro, PT
Ana Maria Correia, ISEGI, Universidade Nova de Lisboa, PT
Ana Paula Afonso, Instituto Politécnico do Porto, PT
Anabela Mesquita, Instituto Politécnico do Porto, PT
Angelica Caro, Universidad del Bío-Bío, CL
António Coelho, FEUP, Universidade do Porto, PT
António Godinho, ISLA-Gaia, PT
Antonio Jesus Garcia Loureiro, Universidad de Santiago de Compostela, ES

António Pereira, Instituto Politécnico de Leiria, PT
Armando Mendes, Universidade dos Açores, PT
Arnaldo Martins, Universidade de Aveiro, PT
Arturo J. Méndez, Universidad de Vigo, ES
Baltasar García Pérez-Schofield, Universidad de Vigo, ES
Benjamim Fonseca, Universidade de Trás-os-Montes e Alto Douro, PT
Bráulio Alturas, ISCTE - Instituto Universitário de Lisboa, PT
Brenda L. Flores-Rios, Universidad Autónoma de Baja California, MX
Brígida Mónica Faria, ESTSP, Instituto Politécnico do Porto, PT
Carlos Costa, Universidade de Aveiro, PT
Carlos Carreto, Instituto Politécnico da Guarda, PT
Carlos Morais, Instituto Politécnico de Bragança, PT
Carlos Vaz de Carvalho, Instituto Politécnico do Porto, PT
Carmen Galvez, Universidad de Granada, ES
Ciro Martins, Universidade de Aveiro, PT
Daniel Castro Silva, Universidade de Coimbra, PT
Daniel Polónia, Universidade de Aveiro, PT
Daniel Riesco, Universidad Nacional de San Luis, AR
David Fonseca, Universitat Ramon Llull, ES
David Ramos Valcarcel, Universidad de Vigo, ES
Dora Simões, Universidade de Aveiro, PT
Eduardo Sánchez Vila, Universidad de Santiago de Compostela, ES
Emiliano Reynares, CIDISI - UTN FRSF - CONICET, AR
Enric Mor, Universitat Oberta de Catalunya, ES
Eusébio Ferreira da Costa, Escola Superior de Tecnologias de Fafe, PT
Fábio Magalhães, Universidade Nove de Julho, BR
Feliz Gouveia, Universidade Fernando Pessoa, PT
Fernando Bandeira, Universidade Fernando Pessoa, PT
Fernando Moreira, Universidade Portucalense, PT
Francisco Restivo, Universidade Católica Portuguesa, PT
Frederico Branco, Universidade de Trás-os-Montes e Alto Douro, PT
Gerardo Gonzalez Filgueira, Universidad da Coruña, ES
Germano Montejano, Universidad Nacional de San Luis, AR
Guilhermina Lobato Miranda, Universidade de Lisboa, PT

Hélder Zagalo, Universidade de Aveiro, PT
Hélia Guerra, Universidade dos Açores, PT
Henrique Gil, Instituto Politécnico de Castelo Branco, PT
Henrique Santos, Universidade do Minho, PT
Higino Ramos, Universidad de Salamanca, ES
Hugo Paredes, Universidade de Trás-os-Montes e Alto Douro, PT
Isabel Pedrosa, Instituto Politécnico de Coimbra, PT
Isaura Ribeiro, Universidade dos Açores, PT
Isidro Calvo, Universidad del País Vasco (UPV/EHU), ES
Ismael Etxeberria-Agiriano, Universidad del País Vasco (UPV/EHU), ES
Ivan Garcia, Universidad Tecnologica de la Mixteca, MX
Javier Garcia Tobio, CESGA-Centro de Supercomputacion de Galicia, ES
Jezreel Mejia, Centro de Investigación en Matemática (CIMAT), MX
João Pascual Faria, FEUP, Universidade do Porto, PT
João Paulo Costa, Universidade de Coimbra, PT
João Tavares, FEUP, Universidade do Porto, PT
Joaquim José Gonçalves, Instituto Politécnico do Cávado e do Ave, PT
Joaquim Madeira, Universidade de Aveiro, PT
Joaquim Reis, ISCTE - Instituto Universitário de Lisboa, PT
Jörg Thomaschewski, University of Applied Sciences OOW - Emden, DE
Jorge Dias Villegas, Universidad de la Frontera, CL
Jose Alfonso Aguilar, Universidad Autonoma de Sinaloa, MX
José Augusto Fabri, Universidade Tecnológica Federal do Paraná, BR
José Braga de Vasconcelos, Universidade Atlântica, PT
José Cascalho, Universidade dos Açores, PT
José Felipe Cocón Juárez, Universidad Autónoma del Carmen, MX
Jose J. Pazos-Arias, Universidad de Vigo, ES
José Luís Silva, Instituto Universitário de Lisboa (ISCTE-IUL), PT
José Paulo Lousado, Instituto Politécnico de Viseu, PT
José Luis Pestrana Brincones, Universidad de Málaga
José Luís Reis, ISMAI - Instituto Superior da Maia, PT
Jose M Molina, Universidad Carlos III de Madrid, ES
Jose Maria Zavala Perez, Eticas Research & Consulting, ES
José Martins, Universidade de Trás-os-Montes e Alto Douro, PT

Jose Maria de Fuentes, Universidad Carlos III de Madrid, ES
Jose R. R. Viqueira, Universidade de Santiago de Compostela, ES
José Silvestre Silva, Academia Militar, PT
Josep M. Marco-Simó, Universitat Oberta de Catalunya, ES
Juan D'Amato, PLADEMA-UNCPBA-CONICET, AR
Juan M. Santos Gago, Universidad de Vigo, ES
Juan Manuel Fernández-Luna, Universidad de Granada, ES
Juan-Manuel Lopez-Zafra, Universidad Complutense de Madrid, ES
Leonardo Bermon, Universidad Nacional de Colombia, CO
Leila Weitzel, Universidade Federal Fluminense, BR
Lilia Muñoz, Universidad Tecnológica de Panamá, PA
Luis Alvarez Sabucedo, Universidad de Vigo, ES
Luís Bruno, Instituto Politécnico de Beja, PT
Luis de Campos, Universidad de Granada, ES
Luis Fernandez-Sanz, Universidad de Alcalá, ES
Luís Ferreira, Instituto Politécnico do Cávado e do Ave, PT
Luis Vilán-Crespo, Universidad de Vigo, ES
Luisa María Romero-Moreno, Universidad de Sevilla, ES
Luisa Miranda, Instituto Politécnico de Bragança, PT
Lus Sussy Bayona Ore, Universidad Nacional Mayor de San Marcos, PE
Magdalena Arcilla Cobián, Universidade Nacional de Educación a Distancia, ES
Manuel Jose Fernandez Iglesias, Universidad de Vigo, ES
Marco Painho, ISEGI, Universidade Nova de Lisboa, PT
Maria Hallo, Escuela Politécnica Nacional, EC
María J. Lado, Universidad de Vigo, ES
Maria João Castro, Instituto Politécnico do Porto, PT
Maria João Ferreira, Universidade Portucalense, PT
Maria João Gomes, Universidade do Minho, PT
Maria José Angélico, Instituto Politécnico do Porto, PT
Maria José Marcelino, Universidade de Coimbra, PT
Maria José Sousa, Universidade Europeia, PT
Marisol B. Correia, Universidade do Algarve, PT
Maristela Holanda, Universidade de Brasília, BR
Martín Llamas Nistal, Universidad de Vigo, ES

Matías García Rivera, Universidad de Vigo, ES
Mercedes Ruiz, Universidad de Cádiz, ES
Miguel A. Brito, Universidade do Minho, PT
Miguel Bugalho, Universidade Europeia, PT
Miguel Casquilho, IST, Universidade de Lisboa, PT
Mirna Ariadna Muñoz Mata, Centro de Investigación en Matemáticas (CIMAT), MX
Nelson Rocha, Universidade de Aveiro, PT
Nuno Lau, Universidade de Aveiro, PT
Nuno Ribeiro, Universidade Fernando Pessoa, PT
Oscar Mealha, Universidade de Aveiro, PT
Paula Peres, Instituto Politécnico do Porto
Paula Prata, Universidade da Beira Interior, PT
Paulo Martins, Universidade de Trás-os-Montes e Alto Douro, PT
Paulo Pinto, FCT, Universidade Nova de Lisboa, PT
Pedro Abreu, Universidade de Coimbra, PT
Pedro Miguel Moreira, Instituto Politécnico de Viana do Castelo, PT
Pedro Nogueira Ramos, ISCTE - Instituto Universitário de Lisboa, PT
Pedro Sánchez Palma, Universidad Politécnica de Cartagena, ES
Pedro Sanz Angulo, Universidad de Valladolid, ES
Pilar Mareca Lopez, Universidad Politécnica de Madrid, ES
Raul Laureano, ISCTE - Instituto Universitário de Lisboa, PT
Reinaldo Bianchi, Centro Universitário da FEI, BR
Rita Santos, Universidade de Aveiro, PT
Rodolfo Miranda Barros, Universidade Estadual de Londrina, BR
Rogério Eduardo Garcia, Universidade Estadual Paulista, BR
Rubén González Crespo, Universidad Internacional de La Rioja, ES
Rui Cruz, IST, Universidade de Lisboa, PT
Rui José, Universidade do Minho, PT
Rui Pedro Marques, Universidade de Aveiro, PT
Santiago Gonzales Sánchez, Universidad Inca Garcilaso de la Vega, PE
Sergio Gálvez Rojas, Universidad de Málaga, ES
Sérgio Guerreiro, Universidade Lusófona de Humanidades e Tecnologias, PT
Silvia Fernandes, Universidade do Algarve, PT
Solange N Alves de Souza, Universidade de São Paulo, BR

Tomás San Feliu Gilabert, Universidad Politécnica de Madrid, ES
Valéria Farinazzo Martins, Universidade Presbiteriana Mackenzie, BR
Victor Hugo Medina Garcia, Universidad Distrital Francisco José de Caldas, CO
Vitor Carvalho, Instituto Politécnico do Cávado e do Ave, PT
Vitor Santos, ISEGI, Universidade Nova de Lisboa, PT
Wagner Tanaka Botelho, Universidade Federal do ABC, BR
Xose A. Vila, Universidad de Vigo, ES

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 07/12/2018
Aceitação/Acceptance: 11/12/2018

Editorial

Tecnología de la Información y Sistemas

Information Technology and Systems

Manolo Paredes[1], Álvaro Rocha[2]

dmparedes@espe.edu.ec, amrocha@dei.uc.pt

[1] Universidad de Las Fuerzas Armadas - ESPE, Av. General Rumiñahui s/n y Ambato, angolqui 171103, Equador

[2] Departamento de Engenharia Informática, Universidade de Coimbra, Pólo II – Pinhal de Marrocos, 3030-290 Coimbra

Pages: ix–x

El vertiginoso desarrollo de la sociedad mundial obedece en gran parte a los avances científicos y tecnológicos, avances que son generados por investigadores involucrados con instituciones públicas y privadas, los cuales entregan sus mayores esfuerzos por generar nuevos conocimientos o aplicar de mejor forma la literatura científica disponible, para mejorar las condiciones de vida de la sociedad.

Sin lugar a dudas que uno de los puntales del desarrollo tecnológico es el campo de las tecnologías de la información y comunicación, puesto que estas, están relacionadas directamente con la productividad y la calidad de vida del ser humano ya que están presentes en todas y cada una de sus actividades. Pero este avance no sería tal, si no se hubiese observado prospectivamente la posibilidad de conformar redes de investigación y desarrollo que dieron cabida al trabajo multidisciplinario, integrando así conocimientos de diversa índole a través de modernos sistemas de comunicación y manejo de la información.

Los procesos de desarrollo tienen su génesis en la investigación, y esta a su vez, no puede ser fructífera si no es acompañada de una adecuada difusión de información, puesto que es a través de esta globalización del conocimiento que se puede alcanzar un resultado con la robustez requerida, para así trascender desde la investigación básica hasta la aplicada y su posterior industrialización.

En ese contexto, esta edición de RISTI presenta los resultados de varias investigaciones realizadas a través de equipos multidisciplinarios que generan una importante contribución para disminuir los niveles de incertidumbre, fortalecer la metodología de investigación, transferir conocimientos, y sobre todo, poner a consideración de la comunidad científica y en general, sapiencias que sin duda son un aporte para solventar pequeños y grandes problemas que actualmente asedian a la sociedad mundial.

Índice / Index

EDITORIAL

Medios, Tecnologías Aplicadas y Comunicación ix
Palmira Chavero Ramírez, Paulo Carlos López-López, Iván Puentes-Rivera, Álvaro Rocha

ARTIGOS / ARTICLES

Sistema IOT para el auto-diagnóstico de enfermedades del corazón usando la evaluación matemática de la dinámica cardiaca basada en la teoría de la probabilidad 1
Carlos Gutierrez-Ardila[1] , Juan Cubillos-Calvachi[1] , Juan Piedrahita-Gonzalez[1], Carlos Montenegro-Marín. Paulo Gaona-García

Herramienta automatizada para el análisis de la satisfacción en pruebas de pensamiento en voz alta 11
Gabriel E. Chanchí G., Luis Freddy Muñoz S., Wilmar Y. Campo M.

Sistemas Biométricos aplicados en smartphones 25
David Suarez, Teresa Guarda

Seguridad Informática en el Uso de los Nuevos Equipos Tecnológicos 32
Jhonny Flores, Teresa Guarda, Lilian Molina

Educação a Distância: as características do líder aplicada ao papel do tutor no processo de ensino aprendizagem 39
Arcângelo dos Santos Safanelli, Dalton Francisco de Andrade, Juliana Brito, Luiz Salgado Klaes, Léia Mayer Eyng, Vânia Ribas Ulbricht

Minería de datos educativa para identificar la relación entre cociente intelectual, estilos de aprendizaje, inteligencia emocional e inteligencias múltiples de estudiantes de ingeniería 48
Pedro Arevalo-Marin, Yonder Cabrera-Piedra, Miguel Cabrera-Sarango, Jonathan Caicedo-Rodriguez, Andres Camacho-Veliz, Sergio Cartuche-Morocho, Cristian Castillo-Aguilar, Daniel Castillo-Betancourt, Jason Chamba-Tigre, Alex Condoy-Carrion, Monica Coronel-Cardenas, Jorlan Elizalde-Cando, Elvis Freire-Feijoo, Adriana Gomez-Jara, Johnny Gonzalez-Guaman, Manuel Mora-Medina, Andres Morocho-Cumbicus, Maria-Jose Ojeda-Bazaran, Nelson Puchaicela-Godoy, Guido Quezada-Alvarez, Nayo Salinas-Minga, Javier Sarango-Espinoza, Luis Chamba-Eras, Milton Labanda-Jaramillo, Edison Coronel-Romero, Jose-Luis Granda y Maria Roman-Sanchez

Modelo Learning Vectors na Avaliação Formativa no AVA Moodle: de Emoticons a GIFs Animados 64
Gilvandenys Sales, Monck C. de Albuquerque, Ricky P. Paz, Eliana A. M. Leite

Alert-LV: Um Sistema de Monitoramento e Apoio à tutoria para suporte à avaliação Formativa em AVA 77
Francisco Adaias Gomes da Silva, Gilvandenys Leite Sales

Avaliação Heurística e Teste de Usabilidade para Software de *Design* de Interiores 90

Noemi Pereira Scherer Cilumbriello, Valéria Farinazzo Martins, Maria Amélia Eliseo, André Luiz Satoshi Kawamoto

OpenChatBotUNL: Propuesta de plataforma de ejecución de Agentes Conversacionales 102

Jose-Luis Granda, Luis Chamba-Eras, Milton Labanda-Jaramillo, Edison Coronel-Romero, Rene Guaman-Quinche, Carlos Maldonado-Ortega

Análisis comparativo de predicción dentro de bases de datos de cáncer: una aplicación de aprendizaje automático..... 113

Gabriel MauricioMartínez-Toro, Dewar Rico-Bautista, Efrén Romero-Riaño

Madurez de la Gestión del Conocimiento en la Ingeniería de Requisitos..... 123

Edgar Serna M., Alexei Serna A.

Estrategias de gamificación con fines de mejora de procesos software en la gestión de proyectos 142

Liliana Machuca-Villegas, Gloria Piedad Gasca-Hurtado

Caracterización de factores que influyen en la baja producción científica de las universidades usando análisis de redes sociales..... 156

Jéssica Ponce, Ariosto Vicuña, Orlando Erazo, Eduardo Samaniego

Propuesta de un test heurístico de accesibilidad para sitios web basados en la norma NTC 5854 170

Daniela Pérez Oliveros, Gabierl E. Chanchi, Maria Isabel Vidal

Granja inteligente: Definición de infraestructura basada en internet de las cosas, IpV6 y redes definidas por software..... 183

Edwin Barrientos-Avendaño, Dewar Rico-Bautista, Luis Anderson Coronel-Rojas, Fabian Ranulfo Cuesta-Quintero

Modelo de arquitectura Cloud Federada utilizando recursos oportunistas de universidades 198

Diego Rincon Cesar Diaz, Olmer Garcia, Holman Bolivar

Principios de la Web Semántica y Computación Afectiva en un Ecoturismo Sustentable mediante el Desarrollo de Aplicación Web Educativa..... 212

Pablo Alejandro Quezada-Sarmiento, Santiago Mengual- Andres

Impactos da implantação da Telemedicina no Tratamento e Prevenção do Câncer 222

Antônio Augusto Gonçalves, Carlos Henrique Fernandes Martins, Sandro Luís Freire de Castro Silva, Cezar Cheng, Roberto Luiz Silva dos Santos, Saulo Barbará de Oliveira

A Governança de Tecnologia da Informação e Comunicação em Organizações de Saúde Pública: Estudo de Caso do Instituto Nacional de Câncer..... 231

Antônio Augusto Gonçalves, Carlos Henrique Fernandes Martins, Sandro Luís Freire de Castro Silva, Cezar Cheng, Roberto Luiz Silva dos Santos, Saulo Barbará de Oliveira

Modelo Ontológico de los Ciberdelitos: Caso de estudio Colombia 244
Jhon Marin, Yuri Nieto, Freddy Huertas, Carlos Montenegro

Modelo para predecir el rendimiento académico basado en redes neuronales y analítica de aprendizaje 258
Nelson Salgado Reyes, Jéfferson Beltrán Morales, Javier Guaña Moya, Charles Escobar Teran, Damian Nicolalde Rodriguez, Gustavo Chafla Altamirano

Jardín botánico: Prototipo desoftware para la gestión y divulgación de plantas nativas basado en código QR y realidad aumentada 267
Edwin Barrientos-Avendaño, Dewar Rico-Bautista, Luis Anderson Coronel-Rojas, Fabian Ranulfo Cuesta-Quintero

Relación entre Elementos y Ventajas de la Cartera Estratégica de Proyectos de TI. Caso de Estudio: Universidades Públicas Españolas 283
Francisco Valverde-Alulema, Faraón Llorens-Largo

Técnicas de mitigación para principales vulnerabilidades de seguridad en aplicaciones web 299
Alex Zambrano, Teresa Guarda, Edward Vladimir Haro Valenzuela, Geovanni Ninahualpa Quiña

Desarrollo backend para aplicaciones web, Servicios Web Restful: Node.js vs Spring Boot 309
Edward Haro[1], Teresa Guarda, Alex Omar Zambrano Peñaherrera, Geovanni Ninahualpa Quiña

Estudio de la Identidad Federada y su Aplicación en el Ecuador 322
Gabriel Patricio Luna Arias, Teresa Guarda, Lucía Liliana Valencia Quezada, Geovanni Ninahualpa Quiña

Un modelo basado en datos abiertos para la creación de un Espacio Latinoamericano y del Caribe de educación superior 336
Nelson Piedra, Lucía Puertas-Bravo, Santiago Mengual-Andrés, Roser Grau

A tecnologia ao serviço da música – Um estudo no sentido de se evitar lesões e maximizando-se a *performance* 349
Eliseu Silva, Manuel Au-Yong-Oliveira, Pedro Fonseca Rui Garganta, Christopher Bochmann

Análisis de la estructura intelectual y la evolución de la investigación en la interacción humano -computador: un análisis bibliométrico..... 363
Gabriel Mauricio Martínez-Toro, Dewar Rico-Bautista[2], Efrén Romero-Riaño, Claudia Jazmin Galeano-Barrera, Cesar D. Guerrero, Jorge A. Parra-Valencia

App Móvil Desarrollada con Metodología Ágil para IoT Controlada desde una Red LAN/WAN con Placa de Desarrollo de Hardware Libre (Arduino) 379
Xavier Villamil, Teresa Guarda

Seguridad de la Información en WSN aplicada a Redes de Medición Inteligentes basado en técnicas de criptografía 393
Liliana Valencia, Teresa Guarda, Gabriel Patricio Luna Arias, Geovanni Ninahualpa Quiña

Sistema de Recuperación de Información para Apoyar la Revitalización del Nasa Yuwe..... 407
Luz Marina Sierra Martínez, Carlos Alberto Cobos Lozada, Juan Carlos Corrales Muñoz, Tulio Rojas Curieux, Luis Carlos Gómez Flórez

Sobre el Éxito de Proyectos y el Framework PSA: Un estudio Fenomenológico en la industria chilena 423
Manuel Cáceres, Aldo Quelopana

Esquema para la actualización de Ontologías de Competencias en base al Procesamiento del Lenguaje Natural y la Minería Semántica 433
Alexandra González-Eras, Jose Aguilar

Algoritmo Memético Autoadaptativo para Solución de Problemas Combinatorios 448
Yuri Cristian Bernal Peña, Alvaro E. Ortiz Dávila, Astrid Z. Peñaranda Capera

Estabelecimento de Requisitos para Aplicações de Realidade Aumentada..... 463
Marcelo de Paiva Guimarães, Valéria Farinazzo Martins, Claudio Kirner, Maria Amélia Eliseo, André Luiz Satoshi Kawamoto

Plataforma y aplicación móvil para proporcionar información del transporte público utilizando un dispositivo embebido de bajo costo.....476
Jorge Buele, Franklin Salazar L., Santiago Altamirano, Abigail Aldás R., Pilar Urrutia-Urrutia

Recuperación de Datos en Dispositivos de Almacenamiento SSD Utilizando File Carving.....490
Geovanni Ninahualpa Quiña, Sang Guun Yoo, Teresa Guarda

Monitorização cognitiva longitudinal numa comunidade: estudo de implementação em contexto real 499
Ana Isabel Martins, Isabel Araujo, Andreia Sousa, Catarina Ruano, Joana Pais, Sara Félix, Eduardo Conde, Luís Ruano, Alexandra Queirós, Vítor Tedim Cruz, Nelson Pacheco Rocha

Herramienta virtual usando dispositivo háptico para terapias de concentración en niños con déficit atencional.....513
Ivón Escobar, Paulina Guanoluisa, Javier Montaluisa, Edwin Pruna y Cristian Duque

Incorporando pruebas basadas en modelos para servicios web en un proceso de desarrollo ágil: Un caso de estudio en la industria 526
Brenda Aymerich, Andrés Solano, Christian Quesada-López, Alexandra Martínez

Máquinas de soporte vectorial para clasificación supervisada de imágenes en bases de datos espaciales 538
Raúl Alejandro Murillo Castañeda, Álvaro Enrique Ortiz Davila

Evaluación Subjetiva de la Calidad de Vídeo para el Servicio YouTube Live en un Dispositivo Móvil 552
Alexander Sócola, Marta Solera, Luís Roberto Jiménez, Matías Toril

Motivaciones hedónicas para el uso de sitios de redes sociales por adultos mayores: un estudio exploratorio en Chile 563
Patricio Ramírez-Correa, Giselle Painén-Aravena, Jorge Alfaro-Peréz, Ari Melo Mariano, Felipe Machorro-Ramos

O Papel da Aceitação da Tecnologia nas Cidades Inteligentes: Um estudo das percepções dos usuários do Uber Brasil 571
Ari Melo Mariano, Patricio Ramírez-Correa, Jorge Alfaro-Pérez, Giselle Painén-Aravena, Felipe Machorro-Ramos

Regresión con polinomios ortogonales completos de alto grado para mejorar la predicción de la distancia mediante RSSI 584
Diego Ordóñez-Camacho, Héctor Ávalos, Estevan Gómez

Un estudio comparativo de técnicas de minería de datos y aprendizaje máquina para la estimación del esfuerzo utilizando puntos de función 595
Christian Quesada López, Juan Murillo-Morera, Marcelo Jenkins

Pedagogia Transformadora e o Reaprender nas Humanidades Digitais 610
João Carlos Wiziack, Vitor Duarte dos Santos

Explicando el uso de Instagram en Chile: Una aplicación de la técnica de árboles de decisión 626
Giselle Painén-Aravena, Patricio Ramírez-Correa, Jorge Alfaro-Peréz, Ari Melo Mariano, Felipe Machorro-Ramos

Metodología para hallar patrones de movimiento en trayectorias sobre DBMS 636
Diego Fernando Rodriguez Lamus, Álvaro Enrique Ortiz Dávila

Pruebas de concepto automatizado sobre aplicaciones web utilizando Raspberry Pi 648
Geraldín Vergara Fajardo, Diana Marcela Montaño, Siler Amador Donado, Katerine Márceles Villalba

Explorando el poder predictivo del modelo TAM: una comparación empírica de la estimación basada en CBSEM, PLS y ANN 660
Felipe Machorro-Ramos, Patricio Ramírez-Correa, Giselle Painen-Aravena, Jorge Alfaro-Peréz, Ari Melo Mariano

Pharming y Phishing: Delitos Informáticos Penalizados por la Legislación Ecuatoriana 671
Hamilton Villón, Marlon Sojos, Carlos Mendoza, Teresa Guarda, Arturo Clery

Herramienta para el seguimiento del estrés mental en pruebas de usabilidad 678
Gabriel Elías Chanchí G., Martha Giraldo G., Wilmar Yesid Campo Muñoz

Transferencia de conocimiento tácito entre administrador y trabajador a través de la socialización: estudio en el sector minero chileno 691

Raúl Caussin, Felipe Machorro-Ramos, Jorge Alfaro-Pérez, Patricio Ramírez-Correa, Giselle Painen-Aravena, Ari Melo Mariano

Enfoques Inteligentes para Identificar Estilos de Aprendizaje de los estudiantes mediante las Emociones en un salón de clases 703

Jorge Cordero, Jose Aguilar, Kristell Aguilar

Sistema para la Detección y Seguimiento de Afecciones Cardíacas Soportado en SBC 717

Darly Mildred Delgado Agudelo, Diego Fernando Girón Timaná, Gabriel Elías Chanchí Golondrino, Katerine Márceles Villalba, Stivens Antonio Dionizio Solarte

Desenvolvimento de Aplicações Acessíveis no Contexto de Sala de Aula da Disciplina de Interação Humano-Computador 729

Valéria Farinazzo Martins, Cibelle Albuquerque de La Higuera Amato, Glaucia Roxo Ribeiro, Maria Amélia Eliseo

Diseño de un Framework de Arquitectura Empresarial para Instituciones Públicas de Educación Superior 742

Inés Meriño Fuentes, Wilson Nieto Bernal, Silvia Moreno Trillos, Yamid Hernández Julio

Análisis del valor potencial de la tecnología: Caso Universidad Francisco de Paula Santander Ocaña 756

Dewar Rico-Bautista, Cesar D. Guerrero, Yurley Medina-Cárdenas, Antón García-Barreto

Introducción de metodologías participativas en la modernización de la administración de tierras en Colombia acercando la tecnología a la comunidad 775

Luz Angela Rocha Salamanca, Alvaro Enrique Ortiz Dávila, Edilberto Suárez

Pensamiento Computacional para la mejora de las capacidades en Geometría Espacial – una experiencia constructivista con Sphero 787

Jason Cahuana, Elizabeth Vidal

Educação a Distância: as características do líder aplicada ao papel do tutor no processo de ensino aprendizagem 795

Arcângelo dos Santos Safanelli, Dalton Francisco de Andrade, Juliana Brito, Luiz Salgado Klaes, Léia Mayer Eyng, Vânia Ribas Ulbricht

Ingeniería de Requerimientos: Análisis ontológico de su desarrollo en la literatura. 804

Jonathan Vásquez, Ariel La Paz

Uso de Arquitecturas Empresariales en e-Government Municipal: Un Mapeo Sistemático de la Literatura 816

Daniela Gallegos-Baeza[1], Ignacio Velásquez[1], Alfonso Rodríguez[1], Angélica Caro[1] 816

Tecnologias educacionais para cegos e tecnologias emergentes: um estudo bibliográfico 830

Graziela de Souza Sombrio, Vania Ribas Ulbricht

Arquitectura IoT para la identificación de personas en entornos educativos.....841
Darly Mildred Delgado Agudelo, Diego Fernando Girón Timaná, Gabriel Elías Chanchí Golondrino, Katerine Márceles Villalba

A evolução cultural graças à adoção da tecnologia e ao fenómeno "Mix-Tech" – Um estudo exploratório baseado na observação 854
Manuel Au-Yong-Oliveira, Frederico Branco, Carlos Costa

Sistema de Gestión de las emociones en un salón de clases inteligente basado en Modelos de confianza y reputación..... 870
Jorge Cordero, Jose Aguilar, Kristell Aguilar, Miguel Martinez

Criptoanálisis del algoritmo RSA de al menos 330 bits con dispositivos de bajo costo SBC 886
Nelsón Darío Pantoja, Anderson Felipe Jiménez, Siler Amador Donado, Katerine Márceles Villalba

Modelo de tópicos para repositorios Git: Un enfoque basado en la estructura de directorios de código fuente.....900
Paul Mendoza del Carpio

Prototipo de monitoreo basado en inferencia difusa para la prevención de incendios en habitaciones urbanas mediante tecnologías IoT 913
Kevin Riaño-Vargas, Jorge Alonso-Echeverri, Paulo Gaona-García, Carlos Montenegro-Marín

Implementación de un Laboratorio Remoto (LR), como recurso de apoyo en un sistema de Educación a Distancia 923
Jairo Antonio Contreras-Mendieta, Celia Paola Sarango-Lapo, Dunia Inés Jara-Roa, Martha Vanessa Agila-Palacios

Modelo de serious game para mejorar la atención en niños con trastorno por déficit de atención e hiperactividad (TDAH). 936
Mayra Albán Taipe

Explorando la relación entre Gestión del Conocimiento y el Rendimiento Organizativo en Instituciones de Educación Superior Universitaria.....947
Sergio Araya-Guzmán, Carolina Henríquez Bravo, Patricio Ramírez-Correa, Ana Barra Salazar

Interpretación de lenguaje de señas ecuatoriano empleando visión por computador 960
Javier P. Pichucho, Patricia N. Constante, Andrés M. Gordón, Darío J. Mendoza

Realidad Aumentada como recurso de apoyo en el proceso enseñanza-aprendizaje en Medicina Legal 972
Jennifer Betzabe Samaniego-Franco, Martha Vanessa Agila-Palacios, Dunia Inés Jara-Roa, Celia Paola Sarango-Lapo

Métricas que podrían usarse en el Desarrollo de Aplicaciones Móviles para Personas con deficiencias Visuales: Una Revisión de Literatura Preliminar..... 985
Erick Zambrano, Rolando P. Reyes Ch., John W. Castro, Efraín R. Fonseca C.

MOMMIV: Modelo para descomposición de una arquitectura monolítica hacia una arquitectura de microservicios bajo el Principio de Ocultación de Información 1000
Víctor Velepucha, Pamela Flores, Jenny Torres

Aplicación del modelo COBIT para la administración de datos (DS11): el caso de un centro de análisis clínico 1010
Jorge Buele, John Espinoza, Daniel Fiallos, Mónica Ruiz R., Ana-María Pilco-Salazar, John Reyes, Franklin Salazar L.

Comparación de comportamiento de redes neuronales recurrentes con otros métodos en pronósticos con series de tiempo – caso series hidrológicas. 1022
Carlos Hernán Fajardo-Toro, Luis Cobo, Paloma María Teresa Martínez-Sánchez

Propuesta de aplicación de minería de procesos para evaluar las rutas de aprendizaje de estudiantes con discapacidad visual en cursos en línea..... 1035
Diego Buenaño-Fernández, Sergio Luján-Mora

Modelo de educación virtual a distancia en el contexto universitario latinoamericano..... 1048
Freire-Avilés Roger, Díaz-Nava Judith, Rodríguez-Suárez William, Vera-Lucio Néstor

Metodología SCRUM y desarrollo de Repositorio Digital 1062
Margarita Ramírez Ramírez, María del Consuelo Salgado Soto, Hilda Beatriz Ramírez Moreno, Esperanza Manrique Rojas, Nora del Carmen Osuna Millán, Ricardo Fernando Rosales Cisneros

Complejidad y gestión del conocimiento e innovación 1073
María del Consuelo Salgado Soto, Ricardo Fernando Rosales Cisneros, Nora del Carmen Osuna Millán, Margarita Ramírez Ramírez, Hilda Beatriz Ramírez Moreno, Maricela Sevilla Caro

Sistema de Gestión Académica a través del desarrollo de Modelo-Vista-Controlador 1083
Esperanza Manrique Rojas, Margarita Ramírez Ramírez, Hilda Beatriz Ramírez Moreno, María del Consuelo Salgado Soto, Nora del Carmen Osuna Millán, Luis Manuel Cerda Suarez

Procesamiento de las opiniones de un personaje público en el Ecuador..... 1094
Boris Enrique Herrera Flores

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 02/09/2018
Aceitação/Acceptance: 20/11/2018

Sistema IOT para el auto-diagnóstico de enfermedades del corazón usando la evaluación matemática de la dinámica cardiaca basada en la teoría de la probabilidad

Carlos Gutierrez-Ardila[1] , Juan Cubillos-Calvachi[1] , Juan Piedrahita-Gonzalez[1], Carlos Montenegro-Marín[2]. Paulo Gaona-García[2]

[1]{chgutierreza, jdcubillosc, jspiedrahitag} @correo.udistrital.edu.co [2]{cemontenegrom, pagaonag}@udistrital.edu.co

Facultad de ingeniería, Universidad Distrital Francisco José de Caldas, Bogotá D.C., Colombia

Pages: 1–10

Resumen: Las enfermedades cardíacas son una de las causas de muerte más comunes a nivel mundial, aun cuando estas pueden ser tratadas si son detectadas de forma temprana. A pesar de saber esto, bien sea por falta de tiempo o de recursos las personas no asisten a chequeos de salud cardiovascular periódicamente. Por tal motivo este artículo presenta el desarrollo de un sistema IOT con arquitectura orientada a servicios, que permite al paciente realizar chequeos preventivos en cualquier momento usando un dispositivo móvil Android y una banda de ritmo cardíaco. Esta solución usa una metodología de análisis basada en la teoría de la probabilidad, lo que permite generar un diagnóstico preventivo que informa al paciente cuando es necesario consultar a un servicio médico especializado, logrando de esta manera reducir los factores de costo y tiempo que hacen que los pacientes omitan el seguimiento a su salud cardiovascular.

Palabras-clave: Análisis de datos, Auto-diagnóstico, IOT, Salud Cardiovascular, Teoría de la probabilidad.

IOT system for self-diagnosis of heart diseases using mathematical evaluation of cardiac dynamics based on probability theory

Abstract: Heart diseases are one of the most common causes of death worldwide, although they can be treated if they are detected early. Despite knowing this, either due to lack of time or resources, people do not attend cardiovascular health check-ups periodically. For this reason, this article presents the development of an IOT system with a service-oriented architecture, which allows the patient to perform preventive check-ups at any time using an Android mobile device and a heart rate band. This solution uses a methodology of analysis based on the theory of probability, which allows generate a diagnosis that informs the patient when he needs to consult a specialized medical service, accomplishing the objective of

reducing the cost and time factors that makes people lose interest in tracking their cardiovascular health.

***Keywords**:* Data analysis, Cardiovascular health, Internet of Things, Probability Theory, Self-Heart Diagnosis.

1. Introducción

Las enfermedades cardiovasculares (ECV) se han convertido en una de las principales causas de muerte alrededor del mundo; de acuerdo con cifras de la Asociación Americana del Corazón, en 2013 se reportó que alrededor de 17.3 de las 54 millones de muertes totales en el mundo estaban relacionadas con ECV, esto corresponde a un 31.5% del total de muertes en el mundo (Benjamín et al., 2017). Con el objetivo de contribuir a reducir este porcentaje, la medicina ha optado por implementar acciones preventivas y monitoreo en tiempo real de la salud del paciente, sin embargo, asistir a chequeos médicos puede costarle al paciente mucho tiempo y dinero, lo que hace que prefiera posponerlos o cancelarlos.

El principal objetivo de esta investigación es brindarle herramientas al paciente para llevar a cabo chequeos preventivos en cualquier momento, para lo cual hemos desarrollado una herramienta que le permite al paciente conocer un estado aproximado de su salud cardiovascular actual por medio de un análisis de sus datos cardiovasculares, sin la necesidad de salir de su hogar o de incurrir en costosos equipos médicos. Esta herramienta consiste en una aplicación móvil para dispositivos Android, que tiene la capacidad de recolectar datos desde un monitor genérico de ritmo cardiaco Bluetooth Low Energy (BLE), y enviarlos a un servidor donde son almacenados, analizados y procesados con el fin de realizar un diagnóstico del estado del paciente cuando este lo desee

2. Estado del arte

Existen diversos proyectos de investigación que tienen el objetivo de crear sistemas que permitan monitorear los signos vitales de un paciente. Dentro de estos sistemas se pueden diferenciar 2 tipos de aplicaciones de monitoreo, primero están las aplicaciones que usan los sensores del celular y segundo las que se conectan a sensores externos de mayor precisión, tales como bandas de ritmo cardiaco (Mohammed et al., 2014), relojes inteligentes o pulseras ópticas. Estos sistemas son capaces de monitorear signos vitales tales como: Presión arterial (Antonovici, Chiuchisan, Geman, & Tomegea, 2014), nivel de glucosa en la sangre (Triventi et al., 2009), ritmo cardiaco(Bathilde, Then, Chameera, Tay, & Zaidel, 2018) y la actividad eléctrica del corazón (ECG) (Gia et al., 2017), entre otros. Los primeros proyectos e investigaciones como (Gao, Greenspan, Welsh, Juang, & Alm, 2005) están orientados a encontrar cambios críticos en los signos vitales de un paciente y a enviar una notificación o alerta si se encuentra un comportamiento anormal.

Se pueden encontrar diferentes ejemplos de sistemas de notificación en (Muppalla, Suraj, Suman, & Mukherjee, 2017; Watanabe, Kawarasaki, Sato, & Yoshida, 2012; Azariadi, Tsoutsouras, Xydis, & Soudris, 2016), estos sistemas envían alertas vía mensaje

de texto cuando el paciente presenta una arritmia cardiaca. Los datos utilizados para detectar las arritmias son tomados usando sensores ECG y son procesados con distintos métodos y algoritmos como la detección de onda R, Transformada Discreta de Pequeñas Ondas (DWT) o las Máquinas de Vector de Soporte (SVM) (Hernández-Ruiz, Villagrán Martínez, Ortiz Guzmán, & Gaona Garcia, 2018). Aunque este enfoque parece bastante atractivo, adquirir los datos de ECG puede ser complejo ya que los electrodos del sensor deben ser ubicados correctamente con el fin de evitar medidas erróneas y muchas veces los usuarios no cuentan con los conocimientos suficientes para realizar la ubicación correcta de los mismos, a esto se debe añadir que tener electrodos cableados en el cuerpo del paciente por largos periodos de tiempo requiere de quietud casi absoluta del mismo lo cual limita sus actividades.

Así, por medio de estos trabajos se puede ver el potencial de los dispositivos móviles en la telemedicina ya que son una de las formas más rápidas de enviar las notificaciones incluyendo datos relevantes como identificaciones, ubicaciones y otros detalles. Sin embargo, con la tecnología actual y los avances de la ciencia médica y de los dispositivos IOT es posible implementar técnicas que permitan dar alertas anticipadas cuando el sistema encuentra signos de una enfermedad incluso antes de que se produzcan los cambios críticos relacionados a estas, este enfoque de predicción tiene una mayor probabilidad de brindar un tratamiento adecuado para la enfermedad del paciente.

Trabajos como el presentado en (Narváez, Villacís, Chalen, & Velásquez, 2017) muestran que para la tarea de diagnóstico aún se necesita la intervención de un especialista, mientras que para algunas aplicaciones con el enfoque predictivo no es de esta forma, un ejemplo de esto se puede ver en (Ani, Krishna, Anju, Aslam, & Deepa, 2017) donde se desarrolla un sistema que con datos del nivel de glucosa en la sangre y el ritmo cardiaco permite predecir posibles casos de derrame cerebral en personas de la tercera edad, el problema de este sistema es que el usuario debe accionar el dispositivo cada vez que quiere realizar una medición de sus signos, por lo tanto, el sistema no cuenta con un monitoreo continuo. Lo anterior nos motiva a juntar las ventajas de estos trabajos, y lograr una herramienta que mediante un monitorio continuo del ritmo cardiaco genere un diagnóstico de la salud que le permita saber al paciente si es necesario que consulte a un médico en algún momento.

3. Descripción del sistema

El sistema se compone de dos partes, la primera es la arquitectura, que define los componentes y sus canales de comunicación. La segunda parte corresponde al componente de análisis de datos, cada una de estas partes será explicada en las secciones presentadas a continuación.

3.1. Arquitectura

El sistema posee 5 componentes, estos componentes interactúan entre sí usando diferentes medios de conexión, que dependen de cada par de componentes. Estos componentes permiten al sistema recolectar y mostrar información al usuario, como se puede ver en la Figura 1. Las mediciones del ritmo cardiaco son tomadas a través del monitor que lleva puesto el usuario, luego de esto, la información es leída en el dispositivo

Android y enviada finalmente al servidor, donde los datos se almacenan en una base de datos y pueden ser leídos para ser procesados. La comunicación entre el servidor y el dispositivo Android se realiza a través de internet y le da al usuario la ventaja de usar el dispositivo para enviar y almacenar los datos en cualquier lugar donde disponga de una conexión a la red.

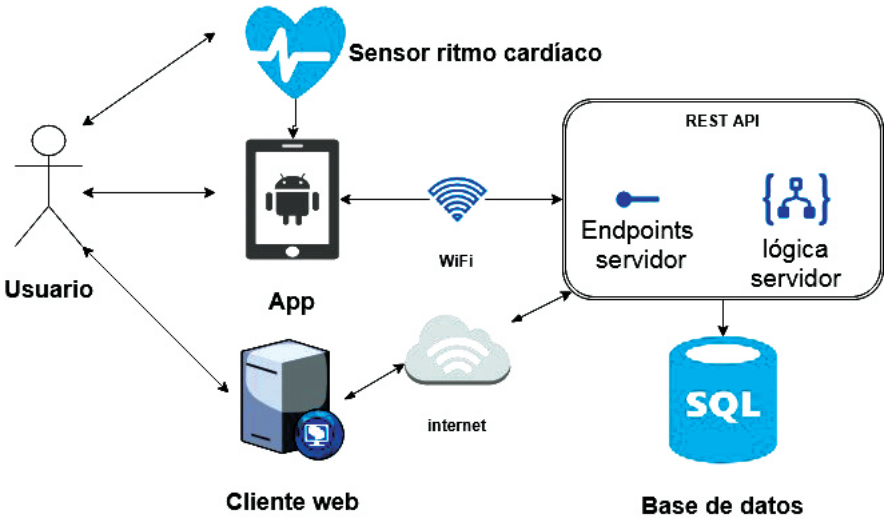


Figura 1 – Componentes de la arquitectura del sistema y flujos de información.

Cada componente tiene funciones específicas y características que permiten dar soporte a la comunicación entre componentes, estas especificaciones se listan a continuación:

- Monitor de ritmo cardiaco. El sensor de ritmo cardiaco funciona bajo el protocolo BLE, el sensor expone servicios que son consumidos por otros dispositivos que operen bajo este protocolo de comunicación. Los servicios expuestos se denominan servicios de Atributos de Perfil de Servicio Genérico (GATT), los cuales establecen en detalle cómo se realiza el intercambio de información a través de la conexión BLE ya que cada servicio posee un identificador único el cual lo diferencia de otros y describe la función del mismo, estos identificadores únicos de servicio están estandarizados en la especificación de servicios GATT de (Bluetooth SIG, 2018).
- Dispositivo Android con aplicación de autodiagnóstico. Se ha desarrollado una aplicación para plataformas Android que establece un puente de comunicación entre el Monitor de ritmo cardiaco y el *Representational State Transfer Application Programing Interface* (REST API), el dispositivo Android debe ser compatible con la tecnología BLE para conectarse a los servicios expuestos por el sensor. Cuando se establece la conexión entre el monitor de ritmo cardiaco y se empiezan a enviar las mediciones que son leídas por la app y mostradas al usuario, (ver Figura 2), se envían luego al servidor donde son almacenadas para continuar con el proceso.

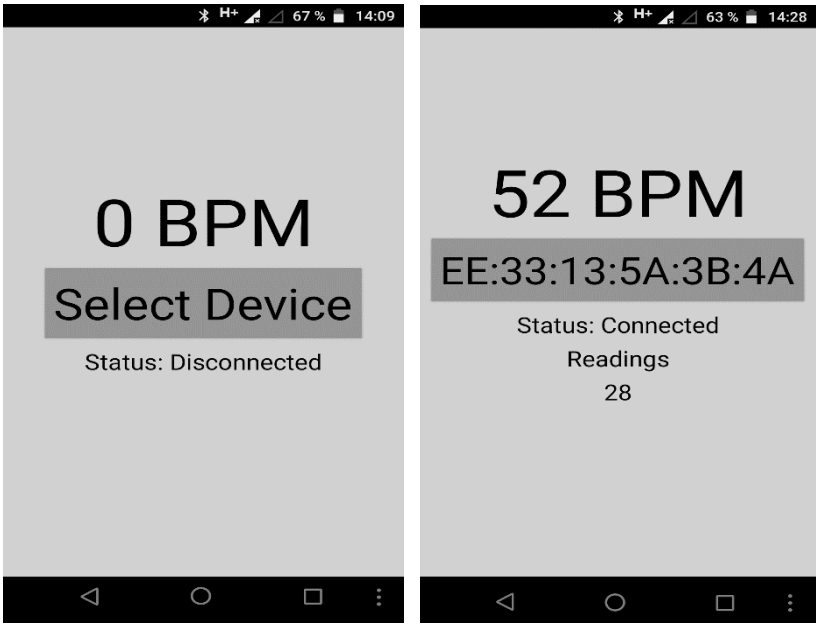


Figura 2 – Pantalla de medición de ritmo cardiaco de la aplicación. En la figura de la izquierda la conexión no ha sido establecida, una vez se activa el Bluetooth y se selecciona un dispositivo, en la imagen de la derecha se muestra al usuario el valor de la medición actual y el número de mediciones tomadas, así como la dirección MAC del dispositivo seleccionado.

- REST API. Este componente realiza todas las operaciones de un CRUD es decir: permite crear, leer, actualizar y borrar información en la base de datos. Además, contiene los procedimientos lógico-matemáticos que se realizan para hacer el análisis de datos y obtener los resultados con el diagnóstico del paciente, los endpoints definidos por el REST API son:
  - login: Hace el inicio de sesión en la app con los datos del usuario.
  - getMeasurements: Obtiene todas las medidas del paciente, si se usa como parámetro el ID del paciente el resultado obtenido solo corresponderá a las mediciones de un paciente en específico.
  - postMeasurement: Guarda en la base de datos las mediciones tomadas a un paciente.
  - getPatients: Obtiene los datos de los pacientes, incluyendo los de contacto. Se puede utilizar el ID de un paciente como parámetro para consultar sus datos.
  - postPatient: Registra un paciente y guarda su información básica.
  - getDiagnosis: Lee los diagnósticos asociados con el paciente que realiza la consulta.
  - postDiagnosis: Guarda los diagnósticos cuando el paciente selecciona la acción en la aplicación.
  - makeDiagnosis: Realiza el análisis de datos con las medidas recolectadas en las últimas 21 horas, y como resultado genera un diagnóstico.

- Servidor BD. La base de datos se encarga de guardar toda la información básica del usuario, las medidas recolectadas y los diagnósticos que el usuario escoge almacenar. La información solo puede ser accedida desde el REST API y puede ser vista en la aplicación o en un cliente web.

3.2. Análisis de datos

Uno de los componentes más importantes del sistema es análisis de datos, el cual permite al usuario acceder a un diagnóstico preventivo de su salud cardiovascular, mediante el REST API, las medidas de la frecuencia cardiaca almacenadas en la base de datos son analizadas con una metodología basada en la teoría de la probabilidad, dicha metodología propuesta en (Rodríguez et al., 2009) se usa para analizar los datos de las mediciones de ritmo cardiaco del paciente por un mínimo de 21 horas, para así establecer si su estado es característico de afección cardiaca, se encuentra sano o en evolución hacia una enfermedad. Esta metodología de análisis fue probada posteriormente en (Rodríguez et al., 2011), presentando buenos resultados.

Con la metodología se determinó que, así como en los sistemas dinámicos, se pueden establecer espacios característicos con el fin de determinar comportamientos asociados tanto a pacientes sanos como enfermos, y ser capaz de reconocer en casos particulares estos patrones, arrojando resultados concluyentes. Usando rangos de 5 latidos por minuto se calcula la probabilidad que tiene el paciente de estar en un rango durante la prueba, de la misma forma se calcula la probabilidad de presentar un numero n de latidos en una hora, estableciendo rangos de 250 latidos por minuto, la metodología se puede ver resumida en la tabla 1.

Resumen criterios para análisis de datos		
Criterio	**Descripción**	**Conclusión**
1	*Rangos RC>=17.	Normalidad
	Rangos RC<=14.	Característico de enfermedad
	14>Rangos RC>17.	No concluyente
2	Se presenta parámetro A.	Característico de enfermedad
	Se presentan parámetros A y B.	
	Solo parámetro B y numero de latidos en una hora >=3000 o >=6250.	
	Solo se presenta parámetro B.	Progreso hacia enfermedad
3	Suma de los dos rangos RC más frecuentes >=0.39, y además se presenta alguna de las condiciones del criterio 2.	Característico de enfermedad

*Rango RC: valores de ritmo cardiaco entre los que se distribuyen las probabilidades del total de los datos de la medición de un paciente

Tabla 1 – Resumen criterios para el análisis de datos.

Viendo de una manera más detallada lo expuesto en la tabla 1, el siguiente criterio determina el estado de salud del paciente:

1. Si los datos de ritmo cardiaco de un paciente se encuentran en más de 17 rangos, su condición es normal, por otro lado, si se encuentra en 14 o menos será característica de enfermedad, de estar entre 14 y 17 rangos este criterio no es concluyente.
2. Los parámetros a y b son aplicados simultáneamente.
    a. Diferencia entre los dos rangos de frecuencia cardiaca que presentan la mayor probabilidad, mayor o igual a 15 es característico de enfermedad.
    b. La mayor probabilidad del número de latidos con valor menor o igual a 0.217 o mayor o igual a 0.304 es característico de enfermedad.

    De acuerdo con los parámetros a y b se tiene:
    - Si se presenta solo el parámetro a, entonces existe una enfermedad.
    - Si los parámetros a y b se presentan, entonces existe una enfermedad.
    - Si solo el parámetro b se presenta, hay avance hacia una enfermedad.
    - Si se presenta el parámetro b y a su vez el número de latidos en una hora es menor que 3000 o mayor que 6250, entonces hay una enfermedad.
3. La suma de las 2 probabilidades más frecuentes en Holters enfermos es característica de enfermedad cuando la misma presenta valores por encima de 0.319 en los siguientes casos:
    - Cuando los parámetros a y b son también característicos de enfermedad.
    - Cuando el parámetro b es característico de enfermedad.
    - Cuando el parámetro b es característico de enfermedad y se presenta un número de latidos menor a 3000 o mayor a 6250 en una hora.

En ese orden de ideas, las posibles respuestas del REST API cuando se ejecuta un análisis son:

- Para el primer análisis:
    - Característico de normalidad.
    - Característico de enfermedad.
    - No concluyente.
- Para el segundo análisis:
    - Característico de enfermedad cardiaca por (primer o segundo) criterio.
    - Característico de enfermedad en evolución por tercer criterio.
    - Sano.
- Para el tercer análisis:
    - Característico de enfermedad cardiaca por (primer, segundo o tercer) criterio.
    - Sano.
    - No concluyente.

4. Caso de estudio

El caso de estudio para esta investigación consistió en la implementación del sistema con un paciente sano. La implementación comenzó cuando el usuario descargaba la aplicación y la instala, luego debe ponerse el monitor de ritmo cardiaco y lanzar la aplicación, El

paciente realiza el proceso de registro e ingresa, selecciona el sensor y la app comienza a recolectar datos. El paciente tiene que usar el dispositivo y recolectar datos por 21 horas como mínimo, como es recomendado en (Rodríguez et al., 2009), mientras realiza sus actividades diarias normales. En este caso el ritmo cardiaco fue monitoreado durante ese tiempo, cuando la prueba finalizó se obtuvieron 75603 registros de mediciones almacenados en la base de datos, cada uno con su momento exacto de captura y su valor.

Cuando el usuario posee suficientes datos para realizar el diagnóstico, se dirige al menú en la app y selecciona la opción donde puede ejecutar el proceso de diagnóstico (ver Figura 3) y ver los resultados del proceso, en este caso, como se esperaba desde un principio de la prueba, el resultado del diagnóstico concluyo que el paciente estaba sano.

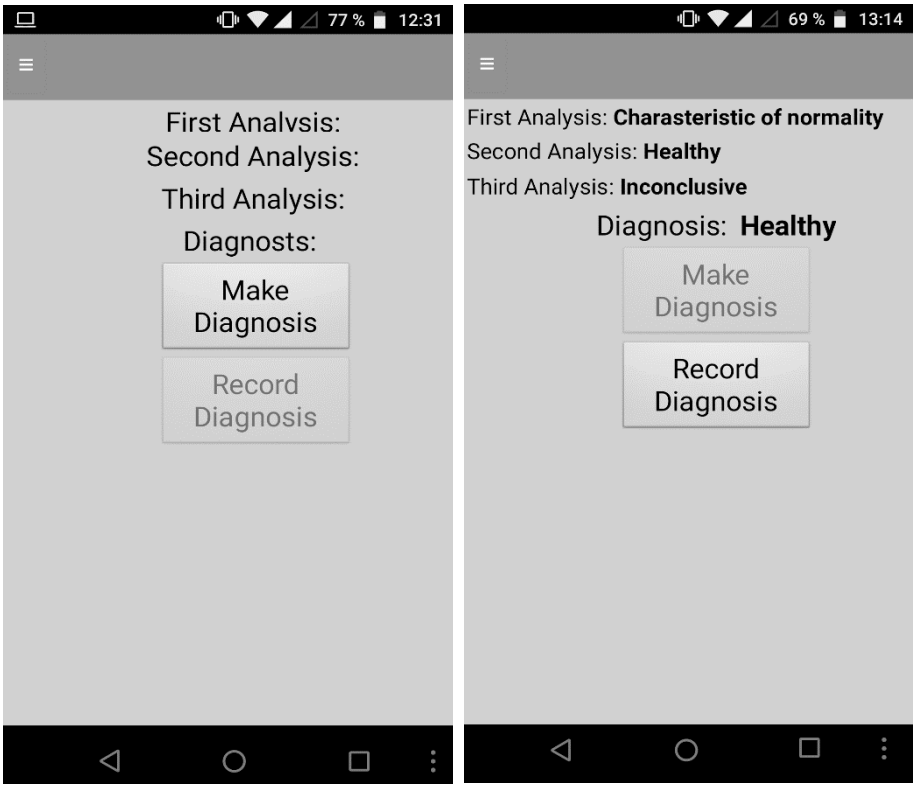


Figura 3 – Pantalla de diagnóstico, en la imagen a la izquierda, el usuario no ha ejecutado el análisis, mientras que la derecha muestra la pantalla después de que el usuario solicita el análisis y obtiene el diagnóstico, se pueden ver los resultados y grabarlos si se desea.

5. Conclusiones

El diseño de este sistema, que incluyó un componente de análisis sobre datos recolectados usando tecnologías móviles e IOT, así como dispositivos portables que se comunican con el sistema mediante protocolos confiables como BLE, surge como una alternativa ante

los diagnósticos tradicionales, que se deben realizar con equipos médicos profesionales. Usando este sistema, el paciente tiene la posibilidad de llevar a cabo chequeos se su salud cardiovascular cuando lo desee, y tan frecuentemente como lo necesite, sin tener siquiera la necesidad de salir de casa ya que el único equipo necesario para tener un diagnóstico está compuesto por un dispositivo móvil Android y un monitor de ritmo cardiaco bluetooth.

Como resultado del proceso el paciente obtiene un diagnóstico del estado de salud cardiovascular del paciente y recomendaciones sobre si es necesario consultar un experto o no, todo esto basado en una variedad de criterios que usan la teoría de la probabilidad. Nuestro trabajo a futuro estará orientado a hacer pruebas estructuradas del sistema con pacientes de cardiología en hospitales.

Agradecimientos

Los autores agradecemos sinceramente el apoyo brindado a este proyecto por parte del grupo de investigación GIIRA de la Facultad de Ingeniería de la Universidad Distrital Francisco José de Caldas.

Referencias

Ani, R., Krishna, S., Anju, N., Aslam, M. S., & Deepa, O. S. (2017). Iot based patient monitoring and diagnostic prediction tool using ensemble classifier. 2017 International Conference On Advances In Computing, Communications And Informatics (ICACCI) (p. 1588-1593). doi: 10.1109/ICACCI.2017.8126068

Antonovici, D., Chiuchisan, I., Geman, O., & Tomegea, A. (2014). Acquisition and management of biomedical data using internet of things concepts. 2014 International Symposium On Fundamentals Of Electrical Engineering (ISFEE) (pp. 1–4). doi: 10.1109/ISFEE.2014.7050625

Azariadi, D., Tsoutsouras, V., Xydis, S., & Soudris, D. (2016). Ecg signal analysis and arrhythmia detection on iot wearable medical devices. 2016 5th International Conference On Modern Circuits And Systems Technologies (MOCAST) (p. 1-4). doi: 10.1109/MOCAST.2016.7495143

Bathilde, J. B., Then, Y. L., Chameera, R., Tay, F. S., & Zaidel, D. N. A. (2018). Continuous heart rate monitoring system as an iot edge de-vice. 2018 IEEE Sensors Applications Symposium (SAS) (p. 1-6). doi: 10.1109/SAS.2018.8336777

Benjamin, E.J., Blaha, M.J., Chiuve, S.E., Cushman, M., Das, S.R., Deo, R., de Ferranti, S.D., Floyd, J., Fornage, M., Gillespie, C., Isasi, C.R., Jimenez, M.C., Jordan, L.C., Judd, S.E., Lackland, D., Lichtman, J.H., Lisabeth, L., Liu, S., Longenecker, C.T., Mackey, R.H., Matsushita, K., Mozaffarian, D., Mussolino, M.E., Nasir, K., Neumar, R.W., Palaniappan, L., Pandey, D.K., Thiagarajan, R.R., Reeves, M.J., Ritchey, M., Rodriguez, C.J., Roth, G.A., Rosamond, W.D., Sasson, C., Towfighi, A., Tsao, C.W., Turner, M.B., Virani, S.S., Voeks, J.H., Willey, J.Z., Wilkins, J.T., Wu, J.H., Alger, H.M., Wong, S.S., Muntner, P.(2017). Heart disease and stroke statistics—2017 update: A report from the American Heart Association. Circulation, 135 (10), 146–603. doi: 10.1161/cir.0000000000000485

Bluetooth. Gatt services. Retrieved 2018-05-19, de https://www.bluetooth.com / specifications/gatt/services

Gao, T., Greenspan, D., Welsh, M., Juang, R. R., & Alm, A. (2005). Vital signs monitoring and patient tracking over a wireless network. 2005 IEEE Engineering In Medicine And Biology 27th Annual Conference (p. 102-105). doi: 10.1109/IEMBS.2005.1616352

Gia, T. N., Jiang, M., Sarker, V. K., Rahmani, A. M., Westerlund, T., Liljeberg, P., & Tenhunen, H. (2017). Low-cost fog-assisted health-care iot system with energy-efficient sensor nodes. 2017 13th International Wireless Communications And Mobile Computing Conference (IWCMC) (p. 1765-1770). doi: 10.1109/IWCMC.2017.7986551

Hernández-Ruiz, C. M., Villagrán Martínez, S. A., Ortiz Guzmán, J. E., & Gaona Garcia, P. A. (2018). Model based on support vector machine for the estimation of the heart rate variability. In V. Kurková, Y. Manolopoulos, B. Hammer, L. Iliadis, & I. Maglogiannis (Eds.), Artificial neural networks and machine learning – icann 2018 (pp. 186–194). Cham: Springer International Publishing.

Mohammed, J., Lung, C., Ocneanu, A., Thakral, A., Jones, C., & Adler, A. (2014). Internet of things: Remote patient monitoring using web services and cloud computing. 2014 IEEE International Conference On Internet Of Things (iThings), and IEEE Green Computing And Communications (GreenCom) and IEEE Cyber, Physical And Social Computing (Cpscom) (p. 256-263). doi: 10.1109/iThings.2014.45

Muppalla, V., Suraj, N. S. S. K., Suman, D., & Mukherjee, B. (2017). Implementation of an arrhythmia detection scheme with cellular based alert framework. 2017 IEEE Calcutta Conference (CALCON) (p. 438-442). doi: 10.1109/CALCON.2017.8280771

Narváez, R. B., Villacís, D. M., Chalen, T. M., & Velásquez, W. (2017). Heart rhythm monitoring system and iot device for people with heart problems. 2017 International Symposium On Networks, Computers and Communications (ISNCC) (p. 1-5). doi: 10.1109/ISNCC.2017.8072004

Rodríguez, J., Correa, C., Ortiz, L., Prieto, S., Bernal, P., & Ayala, J. (2009). Evaluación matemática de la dinámica cardiaca con la teoría de la probabilidad. Revista Med , 20 (4), 183–189.

Rodríguez, J., Correa, C., Prieto, S., Bernal, P., Forero, G., y Sarith Vitery y Luisa Álvarez, G. S., & Puerta, G. (2011). Confirmación del método de ayuda diagnóstica de la dinámica cardiaca de aplicación clínica desarrollado con base en la teoría de la probabilidad. Revista Med , 19 (2), 167–178. doi: 10.18359/rmed.1278

Triventi, M., Mattei, E., Censi, F., Calcagnini, G., Strano, S., & Bartolini, P. (2009). A SMS-based platform for cardiovascular tele-monitoring. IFMBE Proceedings (pp. 295–298). Springer Berlin Heidelberg. doi: 10.1007/9783642039041_82

Watanabe, H., Kawarasaki, M., Sato, A., & Yoshida, K. (2012). Development of wearable heart disease monitoring and alerting system associated with smartphone. 2012 IEEE 14th International Conference On E-health Networking, Applications and Services (HealthCom) (p. 292-297). doi: 10.1109/HealthCom.2012.6379423

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 05/10/2018
Aceitação/Acceptance: 02/12/2018

Herramienta automatizada para el análisis de la satisfacción en pruebas de pensamiento en voz alta

Gabriel E. Chanchí G.[1], Luis Freddy Muñoz S.[2], Wilmar Y. Campo M.[3]

gchanchi@unimayor.edu.co, lfreddyms@gmail.com, wycampo@uniquindio.edu.co.

1 Institución Universitaria Colegio Mayor del Cauca, Cra. 7 # 2-34, 190003, Popayán, Colombia.

2 Fundación Universitaria de Popayán, Cll 5 # 8-58, 190003, Popayán, Colombia.

3 Universidad del Quindío, Cra. 15 Cll 12 norte, 630004, Armenia, Colombia

Pages: 11–24

Resumen: Uno de los modos de evaluar la usabilidad de una aplicación es mediante las llamadas pruebas de usuario, las cuales se desarrollan en un laboratorio de usabilidad. Estas pruebas buscan obtener indicadores de los tres atributos que definen la usabilidad según la ISO 9241-11: eficiencia, eficacia y satisfacción. De los tres atributos el más subjetivo es la satisfacción, dado que se obtiene mediante la percepción del coordinador de la prueba sobre el comportamiento del usuario. Este artículo propone como aporte una herramienta automatizada para el análisis de la satisfacción en pruebas de usuario bajo el enfoque de pensamiento en voz alta y co-descubrimiento a partir del análisis emocional de la voz. De este modo, la herramienta automatizada extrae a partir de los segmentos de audio de la voz, las propiedades acústicas de arousal y valence, mediante los cuales es posible obtener emociones en diferentes momentos de la prueba.

Palabras-clave: Emociones; prueba de usuario; satisfacción; usabilidad.

Automated tool for the analysis of satisfaction in thinking aloud tests

Abstract: One of the ways to evaluate the usability of an application is through the so-called user tests, which are developed in a usability lab. These tests seek to obtain indicators of the three attributes that define usability according to ISO 9241-11: efficiency, effectiveness and satisfaction. Of these three attributes, the most subjective is satisfaction, considering that it is obtained through the perception of the test coordinator on the user's behaviour. This article proposes as a contribution an automated tool for the analysis of satisfaction in user tests under the approach of thinking aloud and co-discovery, from the emotional analysis of the voice. In this way, the automated tool extracts from the audio segments of the voice, the acoustic properties of arousal and valence, through which it is possible to obtain emotions at different moments of the test.

Keywords: Emotions; satisfaction; usability; user tests.

1. Introducción

Cada vez es mayor la cantidad de usuarios que consumen aplicaciones desplegadas en internet o en entornos móviles. Por lo anterior es importante considerar al usuario como parte fundamental en el proceso de desarrollo, puesto que generalmente las empresas del área de la informática dan mayor relevancia al aspecto funcional del software por sobre las características de la interacción (Retnani, Prasetyo, Prayogi, Nizar, & Abdul, 2017). Así, la usabilidad ha venido cobrando importancia, al considerarse un atributo que define la calidad de un producto software (González Sánchez, 2011). De acuerdo a la ISO 9241-11, la usabilidad es entendida como el grado en que un producto software puede ser usado por usuarios específicos, para lograr los objetivos con eficacia, eficiencia y satisfacción, en un contexto de uso específico (ISO, 2018). Dentro de las principales ventajas de la usabilidad se encuentran: es un atributo fundamental en el desarrollo de aplicaciones interactivas, ya que permite un desarrollo competitivo dentro de la industria del software al considerar al usuario como parte esencial, incrementa la productividad del usuario al tener productos más adecuados a sus necesidad, reduce los costos de desarrollo, mantenimiento y capacitación (González Soto Juan Domingo Farnós Miró, 2009).

Para evaluar las aplicaciones interactivas a nivel de usabilidad, se hace uso de los llamados test de usuarios, los cuales son desarrollados en un laboratorio de usabilidad. En estas pruebas los coordinadores de la prueba observan y analizan los diferentes aspectos de la interacción de un usuario con la herramienta automatizada, con el propósito de obtener indicadores de eficacia, eficiencia y satisfacción (Calvo-Fernández Rodríguez, Santamaría, & Valls Saez, n.d.). Dentro de un laboratorio de usabilidad la eficacia puede ser obtenida a partir del porcentaje de tareas cumplidas, la eficiencia puede determinarse con el tiempo que tarda un usuario en ejecutar dichas tareas, mientras que la satisfacción es el atributo más subjetivo, pues depende de la observación del comportamiento emocional del usuario (gestos, posturas, expresiones faciales) por parte del coordinador de la prueba a través del vidrio unidireccional del laboratorio de usabilidad; así mismo la valoración de la satisfacción se suele hacer también mediante la aplicación de cuestionarios de percepción al final de la prueba (Enriquez gabriel, 2013). En este sentido la estimación de la satisfacción, se convierte en un desafío para aprovechar los diferentes datos obtenidos en un laboratorio de usabilidad, con el fin de obtener indicadores menos subjetivos.

De acuerdo a la ISO 9241-11, la satisfacción es un atributo que se define como: “ausencia de incomodidad y existencia de actitudes positivas hacia la utilización del producto” (ISO, 2018). Es así como el estudio de los datos capturados en un laboratorio, haciendo uso de técnicas de computación afectiva posibilitan la obtención de indicadores de la satisfacción en pruebas de usabilidad (González E, 2017). Una de las variables asociadas al usuario y que puede considerarse adecuada para el estudio de la satisfacción es la voz, teniendo en cuenta que es el medio natural para la comunicación y expresión de ideas, del mismo modo en la actualidad uno de los enfoques más difundidos para la evaluación de aplicaciones interactivas es la técnica de pensamiento en voz alta. Este enfoque busca capturar las impresiones o pensamientos de un usuario mientras interactúa con un aplicativo software en el laboratorio de usabilidad, de tal modo que ante cada interacción, el usuario expresa sus opiniones en voz alta (Granollers Toni, 2014). Una variante del anterior enfoque es la técnica de co-descubrimiento, en la cual varios usuarios dentro

del laboratorio conversan de manera grupal, mientras interactúan con un prototipo software (César & González, 2013).

Otra de las ventajas para considerar la voz como variable de estudio de la satisfacción, es el hecho de que a partir de ella pueden obtenerse propiedades acústicas, tales como la activación (arousal) y la valencia (valence) (Solarte Moncayo, Sánchez Barragán, Chanchí Golondrino, Durán Dorado, & Arciniegas Herrera, 2016), las cuales pueden ser usadas para clasificar emocionalmente una muestra de audio en el plano emocional propuesto por Russell (Russell, 1980). Del mismo modo, una ventaja adicional es que la obtención de la voz en un laboratorio de usabilidad es uno de los métodos menos intrusivos de obtener indicadores de la satisfacción, comparado con técnicas asociadas al análisis de variables como: el ritmo cardiaco, las ondas cerebrales, la conductividad de la piel, entre otras (Rodriguez Carrillo, Jimenez Builes, Paterno, & Paterno, 2015).

Este artículo propone una herramienta automatizada para el análisis de la satisfacción en pruebas de usabilidad bajo el enfoque de pensamiento en voz alta y co-descubrimento, a partir del análisis emocional de la voz. Esta herramienta pretende aprovechar los comentarios en voz alta capturados durante una determinada prueba de usabilidad, con el fin de realizar un procesamiento emocional y en el tiempo de las variables acústicas de arousal y valence asociadas a los segmentos de audio de dichas grabaciones. En este sentido dentro de la presente investigación, se consideró el modelo de Rusell para la clasificación emocional de las muestras de audio, el cual incluye en el plano de emociones la propiedad de arousal (eje y) y la propiedad de valence (eje x). La herramienta propuesta ha sido denominado WDEV (Wolf Detector Emotions Voice), fue escrita en el lenguaje java y hace uso en segundo plano de la librería de código abierto openEAR (Eyben, Wollmer, & Schuller, 2009), la cual permite la extracción de las propiedades acústicas arousal y valence a partir de una muestra de audio. Así mismo, la herramienta permite el análisis emocional durante cada una de las tareas definidas en una prueba de usabilidad. A modo de caso de estudio, en el presente artículo se realizó el análisis de la satisfacción en una prueba de usabilidad realizada bajo los enfoques de pensamiento en voz alta y co-descubrimiento sobre el portal de la Institución Universitaria Colegio Mayor del Cauca. Para la construcción de la herramienta se hizo uso de las fases metodológicas tradicionales del desarrollo de software: análisis, diseño, implementación y pruebas, las cuales son abordadas a lo largo del artículo.

El resto del artículo está organizado de la siguiente forma: en la sección 2 se definen los conceptos y herramientas que se tuvieron en cuenta para el desarrollo de este artículo. En la sección 3 se presenta el modelo de emociones considerado por la herramienta propuesta. En la sección 4 se presenta el diseño y construcción de la herramienta para el seguimiento emocional en pruebas de pensamiento en voz alta. En la sección 5 se describe el caso de estudio realizado a través de la herramienta desarrollada. Finalmente en la sección 6 se muestran las conclusiones y trabajos futuros derivados de este trabajo.

2. Marco conceptual

A continuación se presentan algunos de los conceptos que se tuvieron en cuenta para el desarrollo del presente artículo. Dentro de estos se encuentran: satisfacción, modelo emocional de Russell, la voz, librería openEAR, laboratorio de usabilidad, enfoques en prueba de usabilidad.

2.1. Satisfacción

Diferentes definiciones se encuentran en la literatura en cuanto al atributo satisfacción, de este modo según la ISO 9241-11, la satisfacción es entendida como la ausencia de incomodidad y existencia de actitudes positivas hacia la utilización de un producto software (ISO, 2018). Como puede observarse en la anterior definición, la satisfacción está relacionada directamente con las actitudes positivas o emocionales de un usuario con respecto al producto software, por lo cual realizar un seguimiento emocional de un usuario mientras interactúa con una aplicación, puede aportar en cuanto a la estimación de este atributo. Por otro lado según la ISO 9126-1, la satisfacción es definida como la capacidad del software para cumplir con las expectativas del usuario en un contexto de uso determinado (Botta Nicolás, 2014). Finalmente en (Hassan-Montero & Ortega-Santamaría, 2014) el atributo satisfacción es entendido como la medida en la que las tareas realizadas por un usuario en un producto software le resultan agradables y sencillas. En un test de usuario, la satisfacción es obtenida a partir de la percepción del coordinador de la prueba con respecto al comportamiento del usuario (gestos, expresiones faciales, posturas, etc) mientras interactúa con un software determinado.

2.2. Modelo de Russell

El modelo de Russell o modelo circunflejo, es uno de los modelos más utilizados e investigados para el análisis de emociones. Este modelo tiene una estructura circular de dos dimensiones (valencia/activación), que parte el espacio en cuatro cuadrantes, en el cual las emociones son trazadas basándose en su nivel de actividad (activo/pasivo) y su valencia (positiva/negativa) (Schall Andrew, 2015).

Así mismo el modelo circunflejo del afecto, muestra que los estados afectivos surgen de interpretaciones cognitivas de sensaciones nerviosas centrales, que son el producto de dos sistemas neurofisiológicos independientes, uno está relacionado con la valencia (placer/desagrado) y el otro al arousal (estado de alerta). Cada emoción puede ser vista como una combinación lineal de estas dos dimensiones o diversos grados de valencia (Solarte Moncayo, Sánchez Barragán, Chanchí Golondrino, Duran Dorado, & Arciniegas Herrera, 2016).De acuerdo con lo anterior existen dos variables acústicas que hace parte del modelo, las cuales son valencia y arousal.

La valencia es una propiedad auditiva que describe la positividad musical trasmitida por una pista de audio. Las pistas con alta valencia están asociadas con emociones positivas tales como: estar feliz, alegre eufórico, entre otros. Por otra parte, las pistas con baja valencia están asociadas a emociones negativas o estados de ánimo tales como: tristeza, depresión, enojo (Solarte Moncayo, Sánchez Barragán, Chanchí Golondrino, Duran Dorado, et al., 2016). El arousal por su parte representa una medida de percepción de la intensidad y la actividad a lo largo de la pista de audio. Típicamente las pistas de audio rápidas que cuentan con sonidos fuertes y ruido, tendrían una alta energía, mientras que una pista de audio que contenga sonidos leves y poco ruido, estaría bajo en la escala de energía. Al obtener los dos valores arousal y valencia definidos anteriormente se establece un punto en el espacio cartesiano de Russell, que indica la emoción actual de una persona.

2.3. La Voz

Es el principal medio de comunicación entre seres humanos, la voz experimenta cambios constantemente según la persona, en este caso según las emociones esta puede tomar diferentes tonalidades (pitch). Pitch es la frecuencia fundamental (F0) en la que las cuerdas vocales vibran (Mariscal Juan Carmona, 2012), esta variable es considerada como una de las principales portadoras de información acerca de las emociones, el rango pitch refleja el grado de exaltación del hablador, si se tiene una media elevada de F0 esta indica un nivel alto de excitación, existen fluctuaciones en el pitch las cuales son analizadas como el cambio abrupto o suave entre valores altos y bajos, normalmente se genera una curva discontinua en el tono de la voz, esta se considera que refleja las emociones negativas tales como miedo, enfado, etc. Para las emociones positivas como la alegría el cambio del tono de voz es suave.

2.4. Test de usuarios

Un test de usuarios se basa en la observación de un grupo de usuarios que ejecutan tareas específicas y reales, es así como los usuarios son observados por un grupo de evaluadores que analizan los problemas de usabilidad encontrados (Yusef & J., n.d.). Las pruebas de usabilidad se enfocan en medir la capacidad de un producto software para cumplir el propósito para el cual fue diseñado. Estas pruebas son ejecutadas por lo general en un laboratorio de usabilidad, el cual es definido como una plataforma de observación de la interacción entre un usuario y un aplicativo software que está formado por dos módulos aislados, un módulo de observación en el cual se sitúan los evaluadores o coordinadores de la prueba, y el otro de prueba en el cual estará el usuario de manera aislada, donde normalmente hay un cristal unidireccional que permite ver al usuario desde la habitación del evaluador pero no al contrario (Chanchí Gabriel, 2015).

En este trabajo se tuvo en cuenta los enfoques de pensamiento en voz alta y co-descubrimiento (César & González, 2013). El enfoque de pensamiento en voz alta consiste en que el usuario que esté realizando la prueba de usabilidad comenta en voz alta sus percepciones y/o apreciaciones de los diferentes elementos funcionales del software que está evaluando. El enfoque de co-descubrimiento es una variante del enfoque de pensamiento en voz alta, el cual hace referencia a que mientras un grupo de usuarios realiza las tareas de una prueba, dialogan de manera grupal con respecto a los diferentes elementos funcionales del software evaluado.

3. Modelo de emociones considerado

Para el desarrollo de la herramienta se partió del modelo de emociones presentado en (Solarte Moncayo, Sánchez Barragán, Chanchí Golondrino, Duran Dorado, et al., 2016), el cual es una adaptación del modelo de Rusell (Russell, 1980). En dicho modelo se manejan cinco estados de ánimo formando una circunferencia en el plano bidimensional la cual es dividida en cinco sectores iguales de los cuales con base en los resultados de las variables (arousal y valence) se posiciona un ángulo apuntando a una determinada emoción (ver figura 1). Los estados de ánimo considerados en el modelo de cinco emociones son: excitado, feliz, relajado, triste y enojado.

Figura 1 – Modelo de emociones considerado.

A continuación en la tabla 1 se puede observar que cada emoción del modelo y descrita en la figura 2 tiene una amplitud de 72°, así mismo se muestra el rango para el que está determinado cada estado de ánimo (Solarte Moncayo, Sánchez Barragán, Chanchí Golondrino, Duran Dorado, et al., 2016).

Rango de ángulos	**Emoción**
<54° y ≥342°	Feliz
≥54° y <126°	Excitado
≥126° y 198°	Enojado
≥198° y <270°	Triste
≥270° y <342°	Relajado

Tabla 1 – Emociones por rango

4. Diseño y construcción de la herramienta

En esta sección se presenta el procedimiento empleado por la herramienta para el análisis de la satisfacción en pruebas de usuario bajo el enfoque de pensamiento en voz alta y co-descubrimiento. Así mismo se muestra el esquema funcional de la herramienta propuesta, las interfaces de alto nivel y las interfaces finales obtenidas.

4.1. Procedimiento para el análisis de la satisfacción

En la figura 2 se presentan las diferentes etapas consideradas por la herramienta para el análisis de la satisfacción en pruebas de usuario, considerando la voz como variable de entrada. En este caso la herramienta propuesta está enmarcada en el contexto de las pruebas de usabilidad bajo el enfoque de pensamiento en voz alta y/o co-descubrimiento, en la cual se graban los comentarios en voz alta realizados por un usuario o grupos de usuarios mientras interactúan con un software determinado.

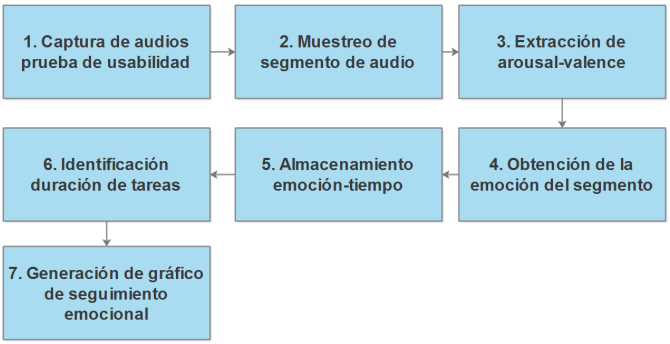


Figura 2 – Procedimiento para el análisis de la satisfacción.

En la etapa 1, se realiza la prueba de usabilidad bajo el enfoque de pensamiento en voz alta y/o co-descubrimiento sobre un determinado producto software, capturando el audio de los comentarios en voz alta de los usuarios durante toda la prueba. En la etapa 2, una vez termina la prueba de usuario y se cuenta con los audios asociados a esta, la herramienta propuesta se encarga de segmentar los audios en archivos de corta duración para analizar la emocionalidad del usuario en esos segmentos. En la etapa 3 una vez segmentados los audios, la herramienta procede a obtener las propiedades acústicas de arousal y valence de cada segmento, las cuales permiten en la etapa 4 la clasificación del segmento en una emoción de acuerdo al modelo de Rusell. Una vez obtenida la emoción asociada a cada segmento, en la etapa 5 la herramienta procede a almacenar la emoción y el tiempo en el cual dicha emoción es obtenida. Para que lo anterior pueda ser realizado las etapas 2, 3 y 4 deben ser realizadas de manera secuencial y hasta que todos los segmentos se hayan procesados. En la etapa 6 es necesario identificar la duración de las diferentes tareas de la prueba para que luego en la etapa 7 la herramienta propuesta genere un gráfico por tarea con el comportamiento emocional del usuario.

4.2. Esquema funcional de la herramienta

En esta sección se presenta el diagrama de bloques de la herramienta WDEV para el seguimiento emocional del usuario en una prueba de usabilidad, teniendo en cuenta como variable fisiológica de entrada la voz (ver figura 3). El diagrama de bloques de la herramienta propuesta está conformado por los siguientes módulos: cargar audio, procesar audio, configurar tareas y generar reporte.

La herramienta WDEV está pensado para ser usado por el coordinador de la prueba, una vez ha terminado la interacción del usuario con la aplicación dentro del laboratorio de usabilidad. De este modo, la herramienta propuesta toma como parámetro de entrada el archivo de audio previamente capturado en las pruebas de usabilidad basadas en los enfoques de pensamiento de voz alta y/o el co-descubrimiento.

Para comenzar, el evaluador haciendo uso de la herramienta WDEV procede en el módulo de cargar audio, a subir el archivo de la prueba el cual debe estar en formato mp3 o wav. Posteriormente en el módulo de procesar audio, la herramienta haciendo uso de las librerías FFmpeg y openEAR segmenta el audio y analiza la emoción asociada a cada segmento de

audio, de tal manera que en la interfaz se puede visualizar como la emoción va cambiando en el tiempo. Para realizar el proceso de clasificación de la emoción, se usan los parámetros musicales de arousal y valence arrojados por la librería openEAR, los cuales una vez obtenidos permiten definir una emoción dentro del modelo de emociones considerado.

Finalmente en el módulo generar reporte, la herramienta propuesta permite la visualización de manera gráfica del comportamiento emocional del usuario durante las diferentes tareas de una prueba de usabilidad. Para realizar el proceso anterior, el coordinador de la prueba debe previamente indicar el número de tareas y los tiempos asociados a cada tarea, separando los datos mediante punto y coma. Como aporte complementario, los datos del seguimiento emocional del usuario y las gráficas de los mismos, pueden ser exportados a los formatos csv y pdf.

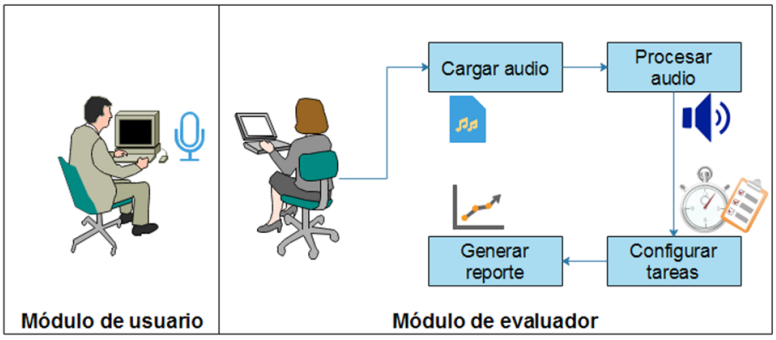


Figura 3 – Esquema funcional de la herramienta.

Con base en el diagrama presentado en la figura 3, a continuación se describen de manera detallada los 4 principales módulos de la herramienta propuesta (cargar audio, procesar audio, configurar tareas y generar reporte) con su respectivo mockup.

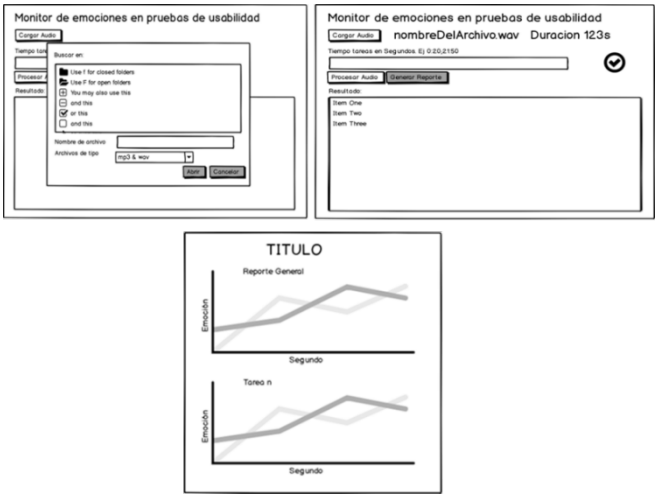


Figura 4 – Mockups de la herramienta software propuesta.

En el módulo cargar audio, el evaluador hace clic sobre el botón cargar audio, la herramienta muestra una interfaz para seleccionar el audio en formato WAV o MP3. Dentro del módulo procesar audio, el evaluador hace clic en el botón procesar, la herramienta comenzará a segmentar el archivo de audio haciendo uso de FFmpeg y a analizar cada segmento empleando la librería openEAR en segundo plano, obteniendo la emoción de cada segmento (ver figura 4).

En el módulo configurar tareas el evaluador se encarga de diligenciar el número de tareas de la prueba y su duración. En cuanto al módulo generar reporte, el evaluador hace clic sobre el botón generar reporte y la herramienta le muestra una interfaz en donde deberá elegir la ubicación del reporte. El reporte generado contiene el comportamiento emocional de un usuario al realizar las diferentes tareas de la prueba de usabilidad. De igual modo, la herramienta propuesta genera un reporte en pdf y csv con los datos del comportamiento emocional del usuario.

4.3. Prototipo final de la herramienta propuesta.

En la figura 5 se presenta la interfaz principal de la herramienta WDEV, donde se aprecian las diferentes funcionalidades del diagrama de bloques. Esta herramienta fue

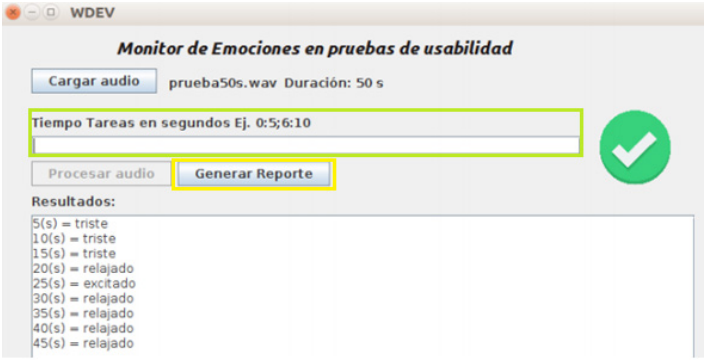


Figura 5 – Interfaz final de la herramienta.

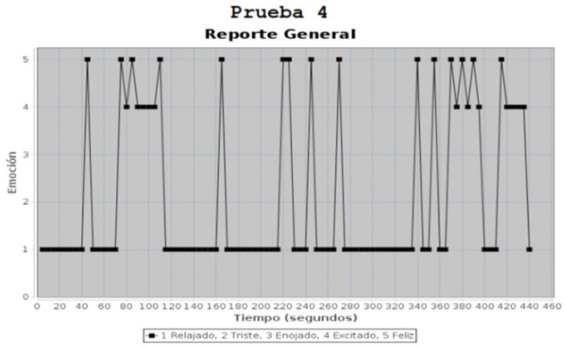


Figura 6 – Reporte general obtenido.

escrita en el lenguaje java y admite formatos de audio mp3 y wav. En caso de que el formato sea MP3, la herramienta hace uso de la librería FFmpeg para convertir el archivo al formato WAV, además WDEV hace uso de la librería openEAR en segundo plano, para obtener las propiedades acústicas de arousal y valence.

De la misma manera en la figura 6 se presenta el reporte gráfico generado por la herramienta a lo largo de la prueba. Así mismo, la herramienta presenta una gráfica por cada una de las tareas de la prueba. Finalmente en el eje Y de la gráfica se pueden apreciar con un número las 5 emociones del modelo de la figura 1.

5. Caso de estudio

A modo de caso de estudio se desarrolló una prueba de usabilidad bajo los enfoques de pensamiento en voz alta y co-descubrimiento sobre el portal web de la Institución Universitaria Colegio Mayor del Cauca. Con los audios obtenidos a partir de la prueba se hizo uso de la herramienta WDEV para realizar el análisis emocional.

5.1. Diseño de la prueba.

Dentro de las tareas consideradas en la prueba se encuentran: consultar calendario académico (tarea 1), buscar la sección de quejas y reclamos (tarea 2), buscar horarios de actividades deportivas (tarea 3).

La prueba se desarrolló en el laboratorio de usabilidad de la Institución Universitaria Colegio Mayor del Cauca (ver figura 7) de acuerdo con las recomendaciones de Nielsen, en las que se menciona que para obtener al menos el 70% de los problemas de un software se debe contar con al menos 5 usuarios (Rodriguez Carrillo et al., 2015). Para la presente prueba se seleccionaron 7 usuarios (6 estudiantes y un administrativo) con edades comprendidas entre 21 y 29 años.

Figura 7 – Laboratorio de usabilidad.

5.2. Resultados de la prueba de pensamiento en voz alta.

Al momento de generar las gráficas de seguimiento emocional, la herramienta propuesta asigna por facilidad convenciones numéricas a cada una de las emociones. Como resultados generales del análisis de la prueba de pensamiento en voz alta, se puede concluir que de las pruebas realizadas en el laboratorio, 4 de los 7 usuarios

experimentaron un cambio de emoción al momento de interactuar con la sección de calendario. Así mismo, 6 de los usuarios que realizaron la prueba cambiaron de una emoción positiva a una negativa al momento de buscar la sección quejas y reclamos. Por otra parte la mayoría de los usuarios presentaron un estado relajado en gran parte del transcurso de la prueba. Del mismo modo, es importante mencionar que se presentaron emociones positivas por parte de 6 usuarios, al momento de buscar información en la sección deportes.

A modo de ejemplo, a continuación se presenta en específico la descripción del seguimiento emocional para 1 de los 7 usuarios que realizaron la prueba usando el enfoque de pensamiento en voz alta. Para realizar esta descripción se tuvo en cuenta las gráficas generadas la herramienta propuesta (ver figura 8). A modo de convención el usuario seleccionado para el ejemplo será llamado usuario A.

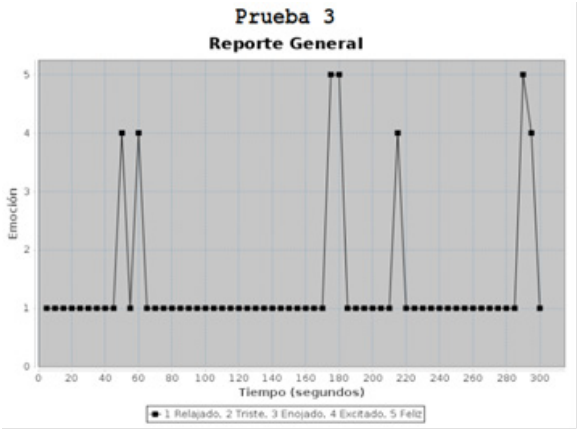


Figura 8 – Reporte general usuario A.

En la figura 8, se presenta el reporte general de la prueba de usabilidad para el usuario A. Así, el usuario A inicia la prueba en un estado de relajación, hasta los segundos 50 y 65, en los cuales se aprecia una intermitencia entre los estados relajado y excitado, lo anterior como consecuencia de la búsqueda infructuosa de la opción "calendario académico". Luego de lo anterior, el usuario A continúa en el estado relajado hasta el segundo 175, en el cual se observa un cambio positivo (estado feliz) mientras se busca la opción "quejas y reclamos". Posteriormente el usuario en el segundo 215 regresa al estado excitado, al tiempo que se percibe un problema con el captcha en el formulario de quejas y reclamos. Después de lo anterior, en el segundo 290 se observa un cambio al estado feliz por parte del usuario A al acceder al "área deportes" del portal. Luego de lo anterior, el usuario A experimenta un cambio hacia el estado enojado al buscar su deporte favorito en el scroll de dicha sección. Finalmente, el usuario termina la prueba sobre el segundo 300 en un estado relajado.

Tal como puede observarse en el reporte general de la prueba, la mayor parte del tiempo el usuario A permanece en el estado relajado. Confrontando lo anterior con la información de video capturado en el laboratorio, es posible observar que los aspectos de la interfaz

que incluyen emociones negativas, están asociados a los elementos de navegación de las secciones “calendario académico” y “quejas y reclamos”.

5. Conclusiones y trabajos futuros

La herramienta propuesta en este artículo representa un aporte importante para las pruebas de usabilidad, en el sentido que permite al evaluador o coordinador de la prueba obtener indicadores que posibilitan la estimación de la satisfacción en las diferentes tareas asociadas a un test de usuario. Lo anterior gracias al análisis emocional que realiza la herramienta WDEV, a partir del audio capturado en una prueba de usabilidad bajo los enfoques de pensamiento en voz alta y/o co-descubrimiento. Así mismo, la herramienta WDEV, permite la obtención de los datos asociados al comportamiento emocional del usuario en formato csv y pdf.

La librería openEAR es un gran aporte para el desarrollo de la herramienta WDEV, ya que esta posibilita la obtención de las variables acústicas asociadas a los segmentos de audio del archivo capturado en la prueba. Estas variables son la base para estimar la emoción percibida por un usuario durante diferentes momentos de la prueba, tomando en consideración el modelo de la figura 2.

Por medio de características musicales como arousal y valence es posible clasificar un contenido de audio en el espacio cartesiano de las emociones. Con estas dos variables la herramienta WDEV pueda calcular las emociones que experimenta un usuario en el desarrollo de una prueba de usabilidad. En este sentido, es adecuado el uso de la herramienta propuesta en conjunto con los enfoques de pensamiento en voz alta y co-descubrimiento, puesto que estos posibilitan la obtención de un archivo de audio con las impresiones en voz alta de un usuario sobre el software a evaluar.

Los enfoques de pensamiento en voz y/o co-descubrimiento requieren de personas con facilidad de expresarse verbalmente, dado que en este tipo de pruebas de usabilidad lo más importantes es obtener la opinión de los usuarios acerca del prototipo a evaluar mediante la voz. En este sentido, los resultados que puede entregar la herramienta dependen de una participación activa de los usuarios en la prueba.

Como trabajo futuro se pretende integrar dentro de la herramienta otro tipo de variables fisiológicas del usuario para monitorear el estado emocional, tales como: el ritmo cardiaco, la conductividad de la piel, los movimientos del usuario, la expresión facial, entre otras.

Referencias

Botta Nicolás. (2014). Usabilidad en sitios Web accesibles. Universidad Abierta Interamericana. Retrieved from http://imgbiblio.vaneduc.edu.ar/fulltext/files/TC116500.pdf

Calvo-Fernández Rodríguez, A., Santamaría, S. O., & Valls Saez, A. (n.d.). Métodos de evaluación con usuarios. Retrieved from http://openaccess.uoc.edu/webapps/02/bitstream/10609/9861/4/PID_00176614.pdf

César, J., & González, S. (2013). Métodos de Evaluación de Usabilidad para Sistemas de Información Web: Una revisión. Bogotá. Retrieved from http://bdigital.unal.edu.co/54171/1/m_usabilidad.pdf

Chanchí Gabriel, M. C. y C. M. (2015). Experimentos de Usabilidad en Unimayor. Almenara, 14, 4–7.

Enriquez gabriel, C. S. (2013). Usabilidad en aplicaciones móviles. Informe Científico Técnico UNPA, ISSN-e 1852-4516, Vol. 5, No. 2, 2013, págs. 25–47 (Vol. 5). Universidad Nacional de la Patagonia Austral. Retrieved from https://dialnet.unirioja.es/servlet/articulo?codigo=5123524

Eyben, F., Wollmer, M., & Schuller, B. (2009). OpenEAR — Introducing the munich open-source emotion and affect recognition toolkit. In 2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops (pp. 1–6). IEEE. https://doi.org/10.1109/ACII.2009.5349350

González E, A. J. (2017). Arquitectura de reconocimiento multimodal del estado emocional de un niño en un contexto educativo. Cali.

González Sánchez, J. L. (2011). Jugabilidad y Videojuegos Análisis y Diseño de la Experiencia del Jugador en Sistemas Interactivos de Ocio Electrónico. Editorial Académica Española.

González Soto Juan Domingo Farnós Miró, Á.-P. (2009). Usabilidad y accesibilidad para un e-learning inclusivo, 2, 49–60. Retrieved from http://www.ujaen.es/revista/rei/linked/documentos/documentos/2-3.pdf

Granollers Toni. (2014). Pensando en voz alta (Thinking Aloud) – MPIu+a. Retrieved August 22, 2018, from http://mpiua.invid.udl.cat/pensando-en-voz-alta-thinking-aloud/

Hassan-Montero, Y., & Ortega-Santamaría, S. (2014). Informe APEI sobre usabilidad. Retrieved from http://eprints.rclis.org/13253/

ISO. (2018). ISO 9241-11:2018(en), Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts. Retrieved August 22, 2018, from https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en

Mariscal Juan Carmona. (2012). Desarrollo de un segmentador fonético automático para habla expresiva basado en modelos ocultos de Markov. Universidad politéctica de Madrid Escuela Técnica superior de Ingenieros de Telecomunicaciones. Retrieved from http://lorien.die.upm.es/juancho/pfcs/JCM/PFCJuanCarmonaMariscal.pdf

Retnani, W. E. Y., Prasetyo, B., Prayogi, Y. P., Nizar, M. A., & Abdul, R. M. (2017). Usability testing to evaluate the library's academic web site. In 2017 4th International Conference on Computer Applications and Information Processing Technology (CAIPT) (pp. 1–4). IEEE. https://doi.org/10.1109/CAIPT.2017.8320714

Rodriguez Carrillo, P. J., Jimenez Builes, J. A., Paterno, F., & Paterno, F. (2015). Monitoreo de la actividad cerebral para la evaluación de la satisfacción. Revista de Ingeniería, 0(42), 16. https://doi.org/10.16924/riua.v0i42.750

Russell, J. A. (1980). A circumplex model of affect. Journal of Personality and Social Psychology, 39(6), 1161–1178. https://doi.org/10.1037/h0077714

Schall Andrew. (2015). Develando las emociones reales de nuestros usuarios: el futuro de la investigación de experiencia del usuario User Experience Magazine. Retrieved August 24, 2018, from http://uxpamagazine.org/the-future-of-ux-research/?lang=es

Solarte Moncayo, L. A., Sánchez Barragán, M., Chanchí Golondrino, G. E., Duran Dorado, D. F., & Arciniegas Herrera, J. L. (2016). Dataset de contenidos musicales de video, basado en emociones. Ingenierías USBmed, 7(1), 37. https://doi.org/10.21500/20275846.2460

Solarte Moncayo, L. A., Sánchez Barragán, M., Chanchí Golondrino, G. E., Durán Dorado, D. F., & Arciniegas Herrera, J. L. (2016). Video on demand service based on inference emotions user. Sistemas y Telemática, 14(38), 29–45. https://doi.org/10.18046/syt.v14i38.2286

Yusef, H. M., & J., M. F. F. (n.d.). No sólo usabilidad : revista multidisciplinar sobre diseño, personas y tecnología. Yusuf Hassan Montero. Retrieved from http://www.nosolousabilidad.com/articulos/interfaces_afectivas.htm

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 11/09/2018
Aceitação/Acceptance: 09/11/2018

Sistemas Biométricos aplicados en smartphones

David Suarez[1], Teresa Guarda[1,2]

david_villon@hotmail.es, tguarda@gmail.com

[1] Universidad Estatal Península de Santa Elena – UPSE, La Libertad, Ecuador

[2] Algoritmi Centre, Minho University, Guimarães, Portugal.

Pages: 25–31

Resumen: Debido a los avances tecnológicos surgidos en los últimos años, el uso smartphones se ha convertido en un recurso indispensable para el ser humano, esto se debe a que permite realizar actividades cotidianas como las llamadas telefónicas, envío de mensajes, tomar fotos, uso redes sociales, correo electrónico, administración de cuentas bancarias, compras y pagos en línea. Por ello, los usuarios deben de tener privacidad de la información personal utilizando sistemas de seguridad que no permitan el fácil acceso en los dispositivos. En la actualidad la mayoría de empresas que se dedican a vender smartphones de gama media y alta ya incluyen sistemas biométricos, en este trabajo hablaremos específicamente de tres sistemas los cuales son el reconocimiento facial, escáner de iris y huellas digitales. Elegimos estos tres mecanismos de seguridad ya que son las técnicas de identificación más seguras que existen en el mercado. Estas técnicas básicamente realizan un reconocimiento de patrones los cuales han sido almacenados previamente en una base de datos.

Palabras clave: reconocimiento de facial, escáner de iris, escáner de huellas digitales, smartphone, sistemas biométricos.

Biometric systems applied in smartphones

Abstract: Due to the technological advances that have arisen in recent years, the use of smartphones has become an indispensable resource for the human being, this is due to the fact that it allows carrying out daily activities such as telephone calls, sending messages, taking photos, using social networks, email, bank account management, purchases and online payments. Therefore, users must have privacy of personal information using security systems that do not allow easy access on devices. Currently most companies that are dedicated to sell mid-range and high-end smartphones already include biometric systems, in this work we will talk specifically about three systems which are facial recognition, iris scanner and fingerprints. We chose these three security mechanisms since they are the safest identification techniques that exist in the market. These techniques basically perform a pattern recognition which has been previously stored in a database.

Keywords: facial recognition, iris scanner, fingerprint scanner, smartphones, biometric systems.

1. Introducción

El origen de la telefonía móvil y su rápida expansión ha supuesto un cambio drástico en nuestra forma de vida, especialmente a lo largo de los últimos 30 años. Hoy en día es una de las tecnologías más comúnmente utilizadas a diario, llegando a considerarse casi indispensable en el día a día. El avance que se ha producido en los medios de comunicación, ya no únicamente por voz sino también a través de mensajes de texto y contenidos multimedia, nos ha permitido un contacto permanente y de forma casi instantánea independientemente de la distancia, y ha dado forma en gran parte a lo que es hoy en día nuestra sociedad y la forma en la que interactuamos. Actualmente el uso de dispositivos móviles ha cobrado gran auge debido, en mayor medida, al peso y tamaño reducido de estos aparatos, lo cual ofrece la posibilidad de llevarlos prácticamente a cualquier lugar; a su capacidad de conectarse a internet y a otros dispositivos; además sus costos son accesibles (Jaime-Ortega, Uribe-Hernández, & Hernández-Domínguez, 2015).

Actualmente los smartphones manejan diferentes tipos de información como son: personal, salud, bancaria, laboral, entre otros; esto implica que su nivel de seguridad debe ser el suficiente para mantener la privacidad del contenido del dispositivo. Por esta razón, la mayoría de fabricantes está implementando sistemas biométricos para mantener la mejor seguridad posible al momento de acceder a los teléfonos.

La biometría es una tecnología que se encuentra hoy en día en pleno desarrollo, tanto en el ámbito de la investigación como en el de la inclusión en la vida cotidiana. Esta rápida evolución se debe a la creciente preocupación actual por el tema de la seguridad y a la vinculación que tiene esta técnica para garantizar la misma (Etchart, Luna, Leal, Benedetto, & Alvez, 2012).

En el presente artículo hablaremos específicamente sobre la seguridad aplicada a los smartphones, en la actualidad hacemos uso de información importante en nuestros dispositivos, es de vital importancia que tengamos la debida seguridad para que cuando otra persona quiera acceder a nuestra información se le haga muy difícil obtenerla, por eso los nuevos smarphones implementan sistemas biométricos los cuales permiten o denegan el acceso dependiendo si los patrones guardados coinciden con la persona que esta intentando hacer uso del dispositivo.

Los sistemas que brindan mayor seguridad son el reconocimiento de facial, escáner de iris y escáner de huellas digitales.

2. Seguridad

La Real Academia Española dice que el termino seguro significa estar libre, sin ningún tipo de peligro o riesgo. Este mismo concepto tiene el mismo sentido si lo aplicamos los sistemas informáticos que existen en la actualidad.

La seguridad es la disciplina que se ocupa de crear las normas, procedimientos, métodos y técnicas destinados a conseguir que los sistemas sean confiables y seguros ante cualquier tipo de amenaza (López, 2010).

Pese a que se aplican las medidas de seguridad, no deja de existir siempre un margen de error. Por ello, debemos a fondo cuales son los elementos que componen los sistemas informáticos, los peligros que afectan al sistema ya sean accidentales o provocados y conocer los mecanismos de seguridad que reducirán los riegos al máximo. Si hablamos de seguridad en smartphones, es importante saber cuales son los sistemas de seguridad con los que cuentan, por ejemplos los sistemas biométricos.

3. Sistemas Biométricos

La biometría hace uso de las características únicas de una persona para permitir su reconocimiento automático. Se han desarrollado sistemas biométricos basados en las huellas dactilares, la voz, la fisiología de las manos o la cara, la escritura, la retina o el iris. En estos últimos años la biometría ha crecido exponencialmente. De usar simplemente la huella dactilar, cuyos orígenes datan del siglo XIV en China, se ha pasado a emplear muchos métodos distintos teniendo en cuenta varias medidas físicas y de comportamiento.

A pesar de la diversidad de estos sistemas, el procedimiento es común a todos. Primero se captura una muestra del rasgo estudiado, por ejemplo, una foto un color de la cara o una muestra de voz. De la muestra se extrae un conjunto de datos de datos de interés, que luego son transformados, mediante algoritmos matemáticos, en una plantilla biométrica. Esta plantilla proporciona una representación normalizada del rasgo, que permite su posterior comparación con otras plantillas pertenecientes al mismo o a otro individuo (Gonzalez, Reconocimiento de iris,in Tesis de licenciatura, 2010).

Estos sistemas tienen un software especializado que procesará esa información en un conjunto de los datos que se pueden comparar con los modelos de los usuarios que se han introducido previamente al sistema. Si se encuentra un “matching” con la base de datos, se confirma la identidad de la persona y se concede el acceso (Tolosa & Giz).

Gracias a los avances tecnológicos, es fácil encontrar varios dispositivos que permiten generar un reconocimiento biométrico del ser humano de una forma fácil y segura. A continuación, describiremos los sistemas biométricos más utilizados actualmente por la seguridad que nos brindan.

3.1. Reconocimiento Facial

Los sistemas basados en reconocimiento facial clasifican la apariencia de la persona e intenta medir algunos puntos nodales del rostro como la distancia que existe entre los ojos, el ancho de la nariz, la distancia del ojo a la boca, o la longitud de la línea de la mandíbula. El análisis tridimensional de la cara elimina algunos inconvenientes que se pueden tener en un reconocimiento bidimensional, como son: la iluminación y las sombras, la orientación o pose de la cara, y la variación de expresiones faciales (Fig. 1).

El reconocimiento facial es muy bueno, pero menos seguro que el análisis de iris y de huellas digitales teniendo la gran ventaja de no ser un método invasivo. Se puede decir que incluso es menos seguro que un patrón, un PIN o una contraseña. En la actualidad

existen muchos códigos fuentes ya desarrollados que permiten un análisis facial de forma simple como los implementados en la red social Facebook (Osorio Cortés, Aguirre Medina, & Escobar, 2010), en algunas ocasiones estos sistemas de reconocimiento facial tienden a fallar si son sometidos a pruebas rigurosas.

Por ejemplo, si el sistema analiza a una persona o un objeto similar al dueño, hay grandes posibilidades de que se desbloquee el smartphone.

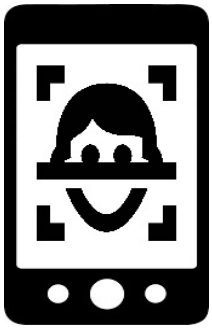


Figura 1 – Reconocimiento Facial en un Smartphone.

3.2. Escáner de Iris

El reconocimiento de iris es una técnica biométrica considerada como uno de los modelos más efectivos para la identificación de una persona.

El propósito de incluir esta función es obtener en tiempo real, acceso a la información de los smartphones sin posibilidades de ser vulnerado por otra persona. Representación del sistema de escáner de iris básico (Fig. 2).

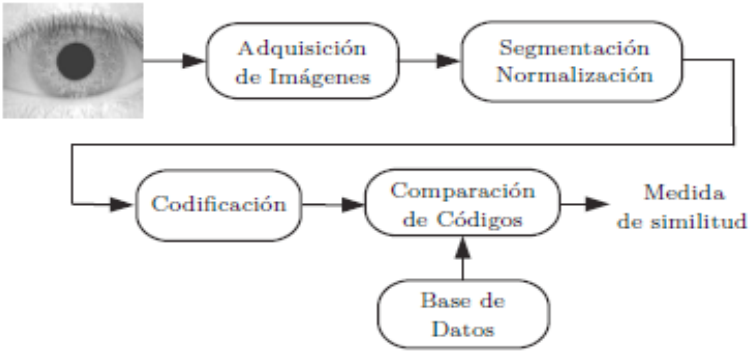


Figura 2 – Diagrama en bloques del sistema de reconocimiento de patrón de iris.

A continuación, mostraremos los pasos resumidos que se realizan en el escáner de iris.

En la primera etapa, se adquiere la imagen del iris de la persona que va a ser reconocida. Posteriormente, la imagen digital es procesada para localizar el iris en la misma y así poder normalizar su tamaño. Luego, la información contenida en el patrón de iris es extraída y un código asociado con el iris es generado. Finalmente, en la etapa de comparación, se decide, en base al porcentaje de similitud obtenido, si los códigos comparados fueron generados por el mismo iris, o sea, por la misma persona, o no (Terissi, Cipollone, & Baldino, 2006).

Todo este proceso se lo lleva a cabo ya que el dispositivo trae un emisor de infrarrojos y una cámara en la parte frontal de la cámara. Al activar el escáner, se encienden ambos componentes, permitiendo ver una fuerte luz roja (el emisor de infrarrojos).

Los infrarrojos impactarán contra los ojos permitiendo que la cámara vea como rebotan incluso de noche, aquí interviene el sistema comparando el iris de la persona contra la información que tiene almacenada en la base de datos de los iris del propietario del smartphone (Fig. 3).

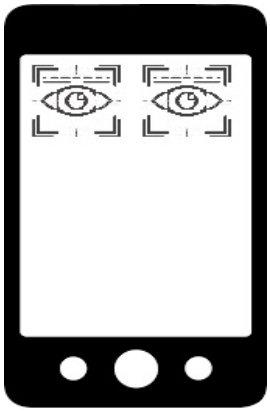


Figura 3 – Escáner de iris en un Smartphone.

3.3. Escáner de Huellas Digitales

Las huellas dactilares son características exclusivas del ser humano. Son las formas caprichosas que adopta la piel que cubre las yemas de los dedos. Están constituidas por rugosidades que forman salientes y depresiones. Las salientes se denominan crestas papilares y las depresiones surcos interpapilares. En las crestas se encuentran las glándulas sudoríparas. El sudor que éstas producen contiene aceite, que se retiene en los surcos de la huella, de tal manera que cuando el dedo hace contacto con una superficie, queda un residuo de ésta, lo cual produce un facsímil o negativo de la huella (Tolosa & Giz). Este es un sistema muy utilizado por los usuarios de smartphones que poseen esta tecnología, lo único que hay que hacer es colocar la huella dactilar en el sensor y el sistema realizara la verificación utilizando las huellas digitales ingresadas previamente por la persona (Fig.4).

Figura 4 – Escáner de huellas digitales en un Smartphone.

4. Conclusiones y/o Recomendaciones

En el presente trabajo se presentó los principales sistemas biométricos que actualmente están uso y desarrollo por los diferentes fabricantes de smartphones. Al igual que en muchos otros campos, la tecnología sigue avanzando tanto en la mejora de las técnicas utilizadas en los sistemas ya existentes como en el desarrollo de nuevas técnicas. Esto es consecuencia de que la seguridad es cada vez más esencial para los usuarios.

Se llego a la conclusión de que los tres sistemas mencionados son muy buenos, pero que sobresalen el escáner de huellas digitales y de iris, porque utilizar rasgos físicos del individuo que son casi imposibles de falsificar. Esto permite a los usuarios sentirse seguros de que ninguna otra persona podrá acceder fácilmente a su información. La mayoría de personas realizan las actividades en sus smartphones y se han dado miles de casos en que, mediante la vulneración de la seguridad del dispositivo, los usuarios han tenido grandes pérdidas económicas o de integridad personal.

Referencias

Etchart, G., Luna, L., Leal, R., Benedetto, M. G., & Alvez, C. E. (2012). Sistema adecuado a estándares de reconocimiento de personas mediante el iris. *XIV Workshop de Investigadores en Ciencias de la Computación*.

Gonzalez, M. (2010). *Reconocimiento de iris,in Tesis de licenciatura*. Universidad de Barcelona, Barcelona, España. Obtenido de http://www.maia.ub.es/~sergio/linked/marcosgonzalezmemoria.pdf

Jaime-Ortega, V., Uribe-Hernández, M., & Hernández-Domínguez, C. (2015). *Researchgate*. Obtenido de Sistemas Biométricos en los Dispositivos Móviles: https://www.researchgate.net/publication/282253580_Sistemas_Biometricos_en_los_Dispositivos_Moviles

López, P. A. (2010). *Seguridad Informática*. Editex.

Osorio Cortés, J. A., Aguirre Medina, F. A., & Escobar, J. (2010). Sistemas de Seguridad basados en biometría. *Scientia Et Technica*, 98–102.

Terissi, L., Cipollone, L., & Baldino, P. (2006). Sistema de Reconocimiento de Iris. *Revista Argentina de Trabajos Estudiantiles*.

Tolosa, C., & Giz, Á. (s.f.). *Departamento de Sistema Informáticos*. Obtenido de http://www.dsi.uclm.es/personal/MiguelFGraciani/mikicurri/Docencia%20/Bioinformatica/web_BIO/Documentacion/Trabajos/Biometria/Trabajo%20Biometria.pdf

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 11/10/2018
Aceitação/Acceptance: 09/12/2018

Seguridad Informática en el Uso de los Nuevos Equipos Tecnológicos

Jhonny Flores[1], Teresa Guarda[1,2], Lilian Molina[1]

flores_dark10@outlook.es, tguarda@gmail.com, lmolina@upse.edu.ec

[1] Universidad Estatal Península de Santa Elena, Santa Elena, Ecuador – UPSE, Av. La Libertad, Santa Elena, Ecuador

[2] Algoritmi Centre, Minho University, Guimarães, Portugal

Pages: 32–38

Resumen: En este artículo se detalla como en los hogares el uso de equipos tecnológicos ayudan a optimizar procesos en las actividades diarias, al realizar una transacción se maneja mucha información la cual es frágil a la hora de ser enviada al internet, ya que por este medio la sustracción de datos se efectúa por las diversas vulnerabilidades que los piratas informáticos encuentran en los equipos. Las personas tienen poco conocimiento sobre seguridad, sus datos son almacenados en la nube y un aparato físico (disco duro), muchas veces administrados por las empresas que ofrecen estos servicios, las leyes de la legislación informática en la constitución de muchos países permitirán la seguridad de la información con el apoyo de herramientas y técnicas establecidas. El objetivo es brindar información para que las personas tengan conocimiento sobre seguridad y como tener precaución en sus equipos.

Palabras clave: seguridad informática, equipos tecnológicos, tecnología de la información.

Computer Security in the Use of New Technological Equipment

Abstract: In this article we detail how in households the use of technological equipment helps to optimize processes in daily activities, when carrying out a transaction, a lot of information is handled, which is fragile when it is sent to the Internet, since by this means the Data subtraction is done by the various vulnerabilities that hackers encounter on computers. People have little knowledge about security, their data is stored in the cloud and a physical device (hard disk), often administered by companies that offer these services, the laws of computer legislation in the constitution of many countries will allow security of information with the support of established tools and techniques. The objective is to provide information so that people have knowledge about safety and how to use caution in their equipment.

Keywords*:* computer security, technological equipment, information technology.

1. Introducción

Muchas de las actividades que se realizan de forma habitual en los países desarrollados dependen en mayor o menor medida del uso de sistemas y de redes en el ambiente tecnológico. El desarrollo de nuevas tecnologías e equipos informáticos permiten mantener la ejecución de transacciones de comercio electrónico, administración electrónica, comunicación entre distintos puntos de la red, y video llamadas o videoconferencias por medio del internet, el termino informática y redes de computadoras ha crecido de forma exponencial hasta el punto de estar inmerso en todos los ámbitos laborales, profesionales e incluso en los hogares desde un mensaje de texto hasta una facturación electrónica. (Vieites, 2011).

Los sistemas informáticos pueden verse afectados por fallas dentro de la seguridad de la compañía que desarrollo el programa informático, pero la información es el factor más importante, el cual debe estar respaldado y con normas específicas de seguridad; si el fallo es de hardware pueden ser remplazados; si es de software del equipo puede volverse a instalar o restaurar; si omitió el paso de respaldar la información puede causar pérdidas a la empresa; si es en el hogar se perderán claves, correos, fotos , entre otros; el principal factor de un fallo de seguridad es el ser humano (Purificación, 2010).

El problema al usar los servicios tecnológicos e informáticos, es que no se tienen los cuidados necesarios por parte del cliente, los usuarios confían en las empresas y esperan que no exista ningún problema; para ingresar a los sistemas debe existir un control de acceso. En la actualidad los usuarios están protegidos y pueden denunciar las situaciones de abusos con sus datos, en cada país existen leyes sobre la privacidad de la información.

Según el estudio realizado por la Policía Nacional, Interpol y el centro de respuesta a Incidentes Informáticos de Ecuador, con el soporte de entidades afines de América Latina, indica que el 85% de los ataques a los programas informáticos son causados por errores de los usuarios; el 58% de personas deja sus teléfonos móviles que contienen información sensible en sus vehículos o lugares de trabajo; el 60% utiliza la misma contraseña en dispositivos de trabajo y de uso diario; el 35% ha ingresado a los correos recibidos por curiosidad por falta de conocimiento del tema; el 59% almacena sus datos de trabajo en la nube; y el 80% de intimidaciones en las redes sociales" (El Telégrafo, 2016).

El articulo está estructurado en 4 secciones que detallan el uso de sistemas informáticos en las transacciones diarias que realizan las personas y no tienen un control donde ingresan información personal; la segunda sección se presentan los peligros del uso de sistemas; por su vez en la tercera sección se describen las buenas prácticas de seguridad informática a nivel personal; en la cuarta sección se encuadra la ley CEFEMD y los riegos potenciales; para finalizar las se presentan las conclusiones y recomendaciones que permiten tener cierto nivel de seguridad en los dispositivos personales que se utiliza a diario.

2. Peligros en el uso de equipos Informáticos

Existen diversas definiciones acerca de delito informático, tomando en consideración cada uno de ellos podemos definir que delito informático es un acto de sustracción de

información con un fin específico para apropiarse de los datos de los usuarios o denegar sistemas empresariales.

En el Consejo de Europa en un convenio sobre Ciberdelincuencia se definió un concepto global sobre los delitos informáticos en los cuales hace mención, que los datos deben ser accesibles a la personas autorizadas, la integridad para que los datos sean reales almacenados en una base de datos y siempre estén disponibles cuando el usuario los solicite (Consejo Europe, 2001).

La tecnología evoluciona días tras día, si piensas que la tecnología (nuevos dispositivos que se diseñan o perfeccionan) puede corregir tus inconvenientes de seguridad (archivos, imágenes, claves), está claro que no comprendes los problemas, ni los beneficios que brinda la tecnología, por la razón de que las empresas innovan de acuerdo a sus conveniencias y estándares internos, que pueden afectar a los usuarios. Debes invertir más en seguridad que en otras prioridades, ya que vas ser blanco fácil para un hacker de los cuales tenemos 4 principios básicos que implican estos casos (Herrera Rosello, 2014):

- Amenaza: se refiere a un fenómeno que atenta a la seguridad de un lugar que almacene los datos por ejemplo: sismos, lluvia, etc.;
- Vulnerabilidad: es un error por medio del cual los hackers ingresan puede ser en el hardware, software o equipo informático;
- Riesgo: son las probabilidades de que sucedan las amenazas contra el equipo físico o el software;
- Seguridad: son las medidas que buscan proteger los recursos informáticos los cuales son 3: datos que ingresan los usuarios, hardware parte física, software para la parte lógica.

En la actualidad existen sistemas, dispositivos electrónicos, que nos permiten optimizar procesos para una determinada actividad, en este conjunto está incluido hardware, software, información, usuarios o familias representados en la figura 1.

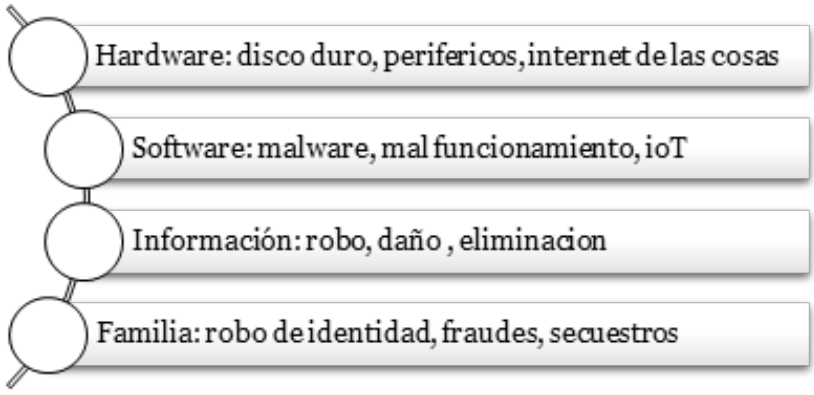


Figura 1 – Categorías de Peligros Informáticos.

Plantear un caso es difícil por la razón de buscar las pruebas en un determinado problema, son delitos que no llevan mucho tiempo en ser ejecutados, no se necesita

estar físicamente, con un computador desde cualquier parte se puede acceder, trabajan de manera anónima.

3. Prácticas de control para equipos informáticos

Si nos referimos a la protección de equipos personales en nuestros hogares, oficinas, dispositivos móviles, cada una de estas instrucciones se debe conocer de forma práctica y teórica, para que las personas tengan precaución de las acciones que realizan en internet.

Lo primero es actualizar el antivirus que viene por defecto o instalar uno nuevo; para los parches de seguridad del sistema se debe tener habilitado la actualización automática aunque muchas veces cause molestias es necesario; visitar páginas web seguras verificando en los navegadores el candadito de color verde en los URL de la parte superior; si va a realizar alguna transacción en internet debe asegurarse de estar en la página correcta; no responder mensajes electrónicos y enviar información personal; actualizar las contraseñas y ubicar letras, números, caracteres para que no las descifren fácilmente; no informas a otras personas las páginas que más visitamos; si recibes correo de una dirección desconocida no lo visualice; no atender llamadas de números desconocidos; no ingresar cd, pendrive en sus computadores si no son de ustedes; usar diferentes contraseñas en sus computadores y no instalar programas que no son necesarios o de fuentes no seguras; el antivirus me permite bloquear las ventanas emergentes y de publicidad que traen algunas páginas para evitar el ingreso de scripts ejecutables (Lopez Vallejo, 2017).

El tema de seguridad es muy importante, se tiene un computador para realizar cualquier tipo de trámites, sirve como una herramienta de la vida diaria, el uso de la tecnología se ha diversificado existen teléfonos, tabletas, laptops, etc., y es allí donde aparecen los peligros. La edad en que las personas usan la tecnología es muy baja, en la actualidad desde los 6 años de edad ya manejan información que se registra en el internet y no tienen ningún cuidado en su uso, las actividades más comunes son el acceso a las redes sociales, correos electrónicos, compras en línea., el robo de contraseña es el más común ya que no tenemos ni el más mínimo cuidado, seguimos ubicando nombre típicos de mascotas, apellidos, nombre, entre otros (Policia Ecuador, 2017).

Si nos vamos ámbito social se debe tener un control parental para los hijos, pues son propensos en las redes sociales a ser engañador al aceptar amigos desconocidos y darles información de tarjetas de créditos, la principal causa es por la compra de juegos, con eso lo seducen y pierden todo su dinero, otro motivo es ingresar a páginas que no son seguras, los agentes externos también pueden causar un daño, en los navegadores los ataques más comunes son spam uso de correo electrónico basura, adware, difunden publicidad en internet, phishing es lo más común se basa en él envió de correos electrónicos similares a los originales con publicidad o archivos impregnados con código malicioso.

4. Leyes informáticas Contra la Protección de información.

A medida que transcurre el tiempo ocurren muchos tipos de casos informáticos a los cuales se les va asignando a una determinada categoría, en el ecuador los registrados según la (Ley No. 2002-67, 2002).

la Pornografía infantil que es muy difícil de ubicar a los infractores por que se redirigen las IP; Violación del derecho a la intimidad si una persona publica fotos, videos en redes sociales sin tu autorización; Revelación ilegal de información de bases de datos, esto involucra tanto a las empresas, usuarios y terceros que actúan en el robo de información; Interceptación de comunicaciones, si se descubre a las personas que captan el tráfico en la red para beneficio propio; Pharming y Phishing, destinatarios de corres electrónicos; Fraude informático en la empresa que labora o intercediendo con otros usuarios; Ataque a la integridad de sistemas informáticos; Delitos contra la información del estado reservada legalmente por los organismos competentes; Acceso no autorizado a un programa informático o de telecomunicaciones por parte de terceros (Policia Ecuador, 2017).

En Ecuador existe la ley de comercio electrónico con sus definiciones en cada uno de sus artículos, que se detallan en la Tabla 1 (Ley No. 2002-67, 2002).

Leyes	**Descripción**
Las infracciones informáticas	Las personas que violen todo tipo de seguridad con información protegida por algún sistema, tendrá prisión de seis meses a un año con una multa económica.
Violación de los deberes de los funcionarios públicos	Funcionario público que dañe, manipule todo tipo de información de los sistemas o datos encomendados a terceros prisión de 3 a 6 años.
Delitos contra la fe pública	Falsificaciones electrónicas alteran o simulan mensajes.
Infracciones informáticas	Borre base de datos o dañe programas prisión de 3 a 5 años con una multa económica.
De la sustracción de datos o sistemas	Los que utilicen un sistema para apropiarse de un bien ajeno. Prisión de 5 años con multa económica.
Estafa y otras defraudaciones	Crear falsas empresas para otorgar créditos. Prisión de 5 años con multa económica.

Tabla 1 – Definición de algunas leyes (Fuente (Ferruzola & Cuenca, 2014)).

Las leyes pueden ser modificadas por el estado y los ciudadanos deben cumplirlas, deben ser conocidas para que no existan conflictos, pero si algunas personas no las respetan deben atenerse a los cargos que se les impongan, la tecnología avanza se requiere que las personas estén protegidas por que toda la información está circulando a través de la red.

5. Conclusiones y Recomendaciones

La seguridad es un tema muy extenso, siempre se busca crear el mejor programa para proteger a los usuarios tanto en las empresas como en los hogares, pero así como se

desarrollan cosas positivas existen también los que se encargan de descifrar el programa creado, existen muchas personas dedicadas al robo de información (la información circula en internet en lo cual no hay nada 100% confiable solo restricciones que cada uno crea).

Se publica información en las redes sociales, ingresamos a correos de publicidad, computadoras desactualizadas, las compras en línea es algo que avanza significativamente he ahí el cuidado de su dinero al ingresar los datos de su tarjeta de crédito y datos personales, en referencia a ello es útil verificar las paginas seguras de las empresas. Se encontró efectivamente que el uso de sistemas informáticos (software), antivirus en las computadoras nos dan un cierto porcentaje de seguridad al momento de navegar en el internet y descargar archivos.

El conocer sobre tecnología permite utilizar los mejores programas y saben cuidar su información para saber dónde la publican, haciendo una recapitulación del conjunto de ideas anteriormente expuestas, siempre se debe preguntar o investigar cuando se valla a acceder, descargar cualquier tipo de información en internet.

Se puede Mejorar el acceso y utilización de las programas, el conocimiento, utilizando nuevas tecnologías desarrolladas, verificar la información que se comparte y atención al robo de datos aunque no es algo tan visible.

Entre algunas características tenemos: Actualizar tu sistema: Periódicamente y con tu red personal es más seguro; Verificar los datos que se envían por correos electrónicos: Publicidad es por lo general el recurso más visto no abras esos correos que no conoces; Crear contraseñas seguras y cambiarlas periódicamente: No coloques tus nombre procura usar letras, números, símbolos para hacerlas más difícil de descifrar; Utilizar antivirus: Es importante bloquea navegación insegura y no deja descargar archivos infectados; No dejar abierto sesiones: Cierra tus sesiones a menos que sea tu computador personal y tenga contraseña principal; Tener activado tu firewall: Para que las personas intrusas tengan dificultad de ingreso; No descargar programas innecesarios: Por curiosidad a veces se descargan cosas innecesarios evite realizar eso; Realizar copias de seguridad: De información importante en caso de perder tu computador; Nunca digas tus datos personales: No existe el amigo confiable cuando se trata de dinero o cualquier otra razón; En redes públicas no hagas transacciones: Este tipo de redes no se sabe quién las administra y que hace; Controla tu privacidad de las redes sociales: No publiques o compartas más de los necesarios; En general son reglas básicas que debes tener en consideración para así poder tener un mejor control de tus datos.

Referencias

Consejo Europe. (2001). *Convenio sobre la Ciberdelincuencia*. European Treaty Series.

El Telégrafo. (2016). *En Ecuador el 85% de los delitos informáticos ocurre por descuido del usuario*. Quito: El Telegrafo.

Ferruzola, E., & Cuenca, H. (2014). Cómo responder a un Delito Informatico. *Revista Ciencia Unemi*, 43–50.

Herrera Rosello, R. (2014). Breve analisis y algunas observaciones al delito. *Revista de Investigación Juridica de Estudiantes*, 1–12.

José, E. G. (2001). *Auditoría en informática* (Segunda ed.). Mexico: Editorial: McGraw-Hill.

Ley No. 2002-67. (2002). *ley de comercio electronico, firmas electronicas y mensajes de datos*. Ecuador, San Francisco de Quito: Publicaciones Congreso Nacional del Ecuador.

Lopez Vallejo, M. R. (2017). Hacking ético. Vulnerabilidad de Sistemas Operativos en el acceso por contraseña. *Revista Publicando*, 31–51.

Policia Ecuador. (2017). *Delitos informáticos establecidos en el COIP y como prevenirlos*. Quito: Policia Ecuador.

Purificación, A. L. (2010). *Seguridad Informatica* (Primera ed.). España: Editorial: Editex.

Vieites, A. G. (2011). *enciclopedia de la seguridad informatica* (Vol. 6). Madrid, España: Grupo Editorial RA-MA.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 15/09/2018
Aceitação/Acceptance: 22/11/2018

Educação a Distância: as características do líder aplicada ao papel do tutor no processo de ensino aprendizagem

Arcângelo dos Santos Safanelli[1], Dalton Francisco de Andrade[2], Juliana Brito[3], Luiz Salgado Klaes[4], Léia Mayer Eyng[5], Vânia Ribas Ulbricht[6]

safanelli.arcangelo@gmail.com[1], dalton.andrade@ufsc.br[2], juligbrito@hotmail.com[3], l.klaes@ufsc.br[4], leiamayer@gmail.com[5], vrulbricht@gmail.com[6]

[1] Universidade Federal de Santa Catarina, Programa de Pós-Graduação em Engenharia de Produção, Florianópolis, Brasil

[21] Universidade Federal de Santa Catarina, Programa de Pós-Graduação em Engenharia de Produção, Florianópolis, Brasil

[3] Instituto Federal de Pernambuco, Tecnologia em Gestão Ambiental, Recife, Brasil

[4] Universidade Federal de Santa Catarina, Departamento de Administração, Florianópolis, Brasil

[5] Universidade Federal de Santa Catarina, Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento, Florianópolis, Brasil

[63] Universidade Federal de Santa Catarina, Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento, Florianópolis, Brasil

Pages: 39–47

Resumo: Este artigo contextualiza a Educação a Distância (EaD) no Brasil, destacando a sua evolução histórica o processo pedagógico nas distintas fases e cenário que a EaD se insere. Com o objetivo analisar quais as características necessárias de liderança para um tutor de polo EaD, o estudo utilizou da pesquisa bibliográfica e a análise qualitativa acerca do tema proposto. Percebeu-se como resultado, que quando mencionamos os diferentes elementos envolvidos na avaliação do processo de EaD assim destacados: professor-tutor, supervisor na comunicação e o auxílio da aprendizagem dos alunos, e trajetória desses atores junto a EaD, ratificam a necessidade e preferência por estabelecer vínculos que demonstram comprometimento entre os pares, em prol de um objetivo comum. Desse modo, o resultado obtido projeta que o professor-tutor da EaD bem como os demais atores, devem contemplar competências, habilidades e atitudes no processo de mediação de aprendizagem coerente e que valorize as potencialidades dos discentes da EaD de modo a integrá-los nas novas plataformas proporcionadas pela EaD atualmente, visando ampliar a interação dos discentes em relação às múltiplas ferramentas disponibilizadas pelos ambientes acadêmicos virtuais de aprendizagem.

Palavras-chave: Educação a Distância (EaD); Tutor; Liderança.

Distance Education: the characteristics of the leader applied to the role of the tutor in the process of teaching learning

Abstract: This article contextualizes Distance Education (DE) in Brazil, highlighting its historical evolution the pedagogical process in the different phases and scenario that the DE is inserted. With the objective of analyzing the necessary leadership characteristics a tutor of DE pole, the study used bibliographic research and qualitative analysis about the proposed theme. As a result, when we mentioned the different elements involved in evaluating the DE, the following are highlighted: teacher-tutor, supervisor in communication and the aid of student learning, and the trajectory of these actors with DE, ratify the need and preference by establishing links that demonstrate commitment among peers, for the sake of a common goal. In this way, the obtained result projects that the profesor-tutor of the DE as well as the other actors, must contemplate competences, abilities and attitudes in the process of mediation of coherent learning and that it values the potentialities of the DE students in order to integrate them in the new platforms offered by the DE currently, aiming to broaden the interaction of the students in relation to the multiple tools made available by the academic virtual learning environments.

Keywords*:* Distance Education (DE). Tutor. Leadership.

1. Introdução

O desenvolvimento da Educação a Distância (EaD) no Brasil, segue inicialmente uma experiência mundial, ressaltadas algumas diferenças. Na EaD brasileira, há a oferta de cursos pelo modelo de correspondência num primeiro momento, todavia, o uso de mídias - rádio e a televisão - , serão incorporadas a EaD e utilizadas em larga escala, obtendo eficácia e efetividade, tornando os conteúdos criativos, bem antes da introdução da Internet.

Historicamente a EaD teve várias denominações como educação por correspondência (Reino Unido); nos Estados Unidos como estudo em casa ou estudo independente; já na Austrália foi chamada de estudos externos. Ainda na França, foi conhecido como telensino ou ensino a distância e teleducação em Portugal. Por fim, Alemanha denominou de estudo ou ensino a distância, e educação a distância na Espanha.

Independente das múltiplas nomenclaturas relatadas pelo mundo, a educação a distância é também um espaço de pessoas. Isso remonta a idéia de que o sucesso da EaD, perpassa a necessidade de criação de instituições e pessoas capacitadas que estabeleçam uma relação de diálogo entre professor e aluno, mediados por materiais didáticos adequados.

Assim, o artigo tem por objetivo analisar quais as características necessárias de liderança para um tutor de polo EaD.

Refletir sobre este papel delegado ao tutor é compreender como este influencia o comportamento de outras pessoas no que refere ao processo de ensino-aprendizagem, sem especificações e detalhes do motivo ou dos processos de liderança. O tutor possibilita a cooperação, promove a interação, auxilia na minimização dos conflitos, mas ao mesmo tempo busca maximizar a criatividade integrativa legítima dos discentes envolvidos.

Neste cenário, o papel do tutor é direcionado a estimular a motivação das pessoas, fazendo com que elas, neste ambiente EaD, tenham objetivos comuns e os percebam como significativos quando menciona-se o processo de formação, mediação e transmissão de conhecimento.

2. Método

A presente pesquisa foi realizada a partir da revisão bibliográfica de uma abordagem qualitativa, por meio da análise de artigos científicos, livros e teses com aprofundamento nos temas relacionados às características no processo pedagógico de ensino-aprendizagem na modalidade EaD e sua evolução ao longo das décadas nos seus diferentes cenários e fases, notadamente pela avaliação dos diferentes elementos envolvidos nesse processo: o professor-tutor, supervisor na comunicação e auxílio da aprendizagem dos alunos, a sua forma de interação ao longo da trajetória em EaD, as formas de estudos e elaboração de ferramentas utilizadas no processo de ensino e aprendizagem, tendo-se em vista uma análise focada no comprometimento entre as equipes de tutores e os discentes na construção pedagógica enfatizando a valorização da figura humana. Por fim, a divulgação dos resultados alcançados por esse estudo através da elaboração de trabalhos científicos resultantes da participação direta dos autores em diferentes instituições de ensino tecnológico à distância e posterior avaliação da interação dos alunos em relação às diferentes ferramentas utilizadas nos ambientes acadêmicos virtuais, tais como os fóruns de discussões e dúvidas, fóruns avaliativos, envios de arquivos, chats de dúvidas e sugestões e as aulas presenciais como suporte ao encontro presencial entre tutores e alunos.

3. Referencial Teórico

3.1. Educação a Distância

A EaD quando se refere a uma padronização conceitual não é unânime. Em 1967, Dohmen, apontava que a EaD, como algo sistemático e organizado, em que o auto-estudo é destacado, mediante o uso de um material de apoio, acompanhamento e supervisão (BERNARDO, 2009).

De forma polêmica, "o conceito de Peters em 1973, dá ênfase a metodologia da EaD, afirmando que "a Educação a Distância é uma forma industrializada de ensinar e aprender" (BERNARDO, 2009, p.85).

> *Educação/ensino a distância é um método racional de partilhar conhecimento, habilidades e atitudes, através da aplicação da divisão do trabalho e de princípios organizacionais, tanto quanto pelo uso extensivo de meios de comunicação, especialmente para o propósito de reproduzir materiais técnicos de alta qualidade, os quais tornam possível instruir um grande número de estudantes ao mesmo tempo, enquanto esses materiais durarem. É uma forma industrializada de ensinar e aprender (MOORE, 1973 apud VASCONCELOS, 2010).*

Com esta fala, Moore (1973 apud VASCONCELOS, 2010), esclarece que as ações do professor e o desenvolvimento da comunicação com os alunos deverá ser facilitada, com o uso de meios e métodos instrucionais. Já observava o autor, que a intervenção do professor deveria ocorrer a partir da ação do aluno sempre mediada por meios impressos, eletrônicos, mecânicos entre outros.

O conceito de Holmberg em (1977 apud VASCONCELLOS, 2010), aponta para as múltiplas formas de estudos em que a EaD apresenta. Na sua conceitualização, Holmberg

ratifica a contínua e imediata supervisão com tutores, porém em sala de leituras e em um espaço físico local.

Em contraponto a questão da presença física entre tutor e aluno, 14 anos depois, em 1991, Keegan ressalta que a EaD faz uso do planejamento, direção e instrução que uma Keegan organização de ensino possui. Conceitualmente, sua definição passar a ser compreendida como:

> *a separação física entre professor e aluno, que a distingue do ensino presencial, comunicação de mão dupla, onde o estudante beneficia-se de um diálogo e da possibilidade de iniciativas de dupla via com possibilidade de encontros ocasionais com propósitos didáticos e de socialização. (VASCONCELOS, 2010).*

De encontro ao proposto por Keegan em 1991, Chaves ratifica (1999 apud VASCONCELOS, 2010), a separação física e o uso de tecnologias de telecomunicação. Para o autor o "ensinante e o aprendente estão separados no tempo e no espaço". Esta separação deveria ser preenchida pela telecomunicação por meio da transmissão de dados, voz e imagens.

Paralelo a esta conceituação, o mercado de trabalho tem solicitado dos indivíduos, uma aquisição constante de conhecimento, ampliando habilidades e competências. Estas novas demandas ratificam a necessidade da modalidade EaD, de modo a explorar este nicho e atender os anseios da comunidade.

Todavia, requer apontar que esta nova demanda, aponta para uma forma diferencial de construção do conhecimento, de saberes. Na EaD, o *modus operanti* baseia-se na ampliação do conhecimento, na troca de informações, no rearranjo dos materiais instrucionais, de modo a extrapolar a estrutura física de uma sala de aula e o sinal do relógio.

Se antes a educação a distância, era reconhecida em uma época em que o correio e o rádio se destacam como suportes e veículos para os cursos das mais diversas áreas, o avanço dos meios de comunicação e da informática computacional, a ampliação e o desenvolvimento de *softwares* educativos, apresentam uma forma diferenciada para a modalidade, ampliando a possibilidade de ensino e a aprendizagem.

Para Salvucci, Lisboa e Mendes (2012, p.51):

> *a didática do ensino virtual deve contemplar a multidimensionalidade dos recursos e ferramentas tecnológicas e os procedimentos adequados prevêem: - Ênfase na autonomia do aluno, Exploração das possibilidades do material didático, Domínio das ferramentas, Conhecimento prévio dos processos de interação e mediação, Disponibilidade e interesse para a comunicação diferenciada das fontes de informação.*

Assim, Pallo e Pratt (2002), acenaram em seus estudos acerca aprendizagem a distância, mediada por mídia computacional, que são fundamentais:

> *[...]as interações entre os próprios estudantes, as interações entre professores e os estudantes e a colaboração na aprendizagem que resulta de tais interações. Em outras palavras, a formação de uma comunidade de alunos por meio da qual o conhecimento é transmitido e os significados sejam criados*

> *conjuntamente, prepara o terreno para bons resultados na aprendizagem. (PALLO; PRATT, 2002, p. 27.)*

Por fim, conceituar a EaD nos requer compreender com detalhismo o processo de formação, os limites e possibilidades das Tecnologias da Informação e Comunicação (TICs), porém não perdendo de vista o quanto é possível configurar e reconfigurar de modo dinâmico a rede.

3.2. A Educação a Distância, a Tutoria e a Liderança

Muitos autores vêm criticando a expressão EaD, pois ela apontaria para uma distância que, do ponto de vista educacional ou psicológico, não existe necessariamente quando fazemos EaD. Tori (2010), por exemplo, fala em uma Educação sem Distância. Uma expressão alternativa, cada vez mais utilizada, é Educação Aberta e a Distância. Outros autores, como Silva e Santos (2006) e Silva, Pence e Zuim (2010), preferem a expressão Educação On-line, que sinalizaria a interatividade possibilitada pelas novas TICs. Algumas variações em inglês são online learning, virtual learning, networked learning e web-based learning. A expressão e-learning é também bastante utilizada, em geral para representar a EaD corporativa, que se faz em instituições que não têm o ensino como missão primária.

Contudo, a educação a distância, deve considerar, conforme Faria (2010), que o estudante é sujeito ativo do processo tornando-se protagonista na construção do próprio conhecimento e para isso, este precisa estar disposto a buscar as informações, com estímulo e motivação.

Um elemento importante neste processo de construção do conhecimento e da ratificação da cultura EaD, cabe a presença do tutor, neste processo de mediação. Bruno e Lemgruber (2010) falam de docência coletiva, e de possibilidades que vão da fragmentação neotecnicista à parceria cooperativa, e que tudo depende da concepção de educação que se tem. Lembram que

> *o tutor a distância é também um docente e não simplesmente um animador ou monitor neste processo, e muito menos um repassador de pacotes instrucionais. Este profissional, como mediador pedagógico do processo de ensino e de aprendizagem, é aquele que também assume a docência e, portanto, deve ter plenas condições de mediar conteúdos e intervir para a aprendizagem. Por isso, na prática, o professor-tutor é um docente que deve possuir domínio, tanto tecnológico quanto didático (BRUNO; LEMGRUBER, 2010, p. 75).*

É visível que a equipe de tutores possui um papel fundamental neste processo de educacional, pois é visto como um sujeito protagonista que participa de modo ativo na prática e na construção pedagógica. Independente de estar presencial ou a distância, sua contribuição se dá de modo efetivo na construção do processo de ensino aprendizagem e por conseqüência no processo de acompanhamento e avaliação do projeto pedagógico dos cursos.

Assim, segundo Shulman (1995 apud LITWIN, 2001), é ressaltado o domínio do conteúdo como algo imprescindível e elemento essencial para o exercício da função. Paralelo a isso, o domínio do conteúdo está aliado à necessidade de dinamismo, visão crítica e global.

Neste cenário, o tutor deve considerar aspectos relacionados a liderança como elementos indispensáveis à sua conduta no processo de ensino e aprendizagem. As transformações que ocorrem no segmento educacional, impõe uma nova realidade para as instituições, exigindo das mesmas, inovadoras formas de agir, estruturas mais flexíveis e valorização do capital humano. No entanto, para se conseguir efetivar o processo de mudança nas instituições de ensino, é necessário que as pessoas que delas fazem parte estejam comprometidas com a mudança. E, neste caso, o papel do tutor que possui liderança é fundamental, uma vez que ele é co-responsável pela valorização da participação efetiva dos discentes, motivando-as e inspirando-as do processo de aprendizagem.

"A ligação aluno-professor ainda é, no imaginário pedagógico, uma dominante, o que torna a tutoria um ponto-chave em um sistema de ensino a distância (MAIA, 1998 apud NISKIER, 1999, p.391). Em função disto, é indispensável que as instituições desenvolvam planos de capacitação de seu corpo de tutores. Um programa de capacitação de tutores deve, no mínimo, prever três dimensões como apontado por Shulman (1995 apud LITWIN, 2001):

- capacitação no domínio específico do conteúdo;
- capacitação em mídias de comunicação; e
- capacitação em fundamentos da EaD e no modelo de tutoria.

Para Niskier (1999, p.320), o papel e a capacitação do tutor deve extrapolar as competências e habilidades, visando a inclusão de atitudes. Deve-se considerar os aspectos didático-pedagógico voltados para acompanhar, motivar, orientar e estimular o aprendizado do discente. "Seu trabalho torna-se mais qualificado com o conhecimento, com profundidade, do material didático produzido pelo professor, muitas vezes, podendo até participar, com sugestões, de sua produção e/ou de sua reformulação, sobremodo se estiver inteirado do projeto político-pedagógico do curso".

Assim, os tutores qualificados, são para Bennis (1996), os líderes que desenvolvem estratégias em aprender, em adquirir conhecimentos e experiências. Assim, seu mundo vai tornando-se mais amplo e complexo, sendo que a complementaridade entre educação formal e informal é extremamente importante para a formação do líder.

O verdadeiro líder tutor, forma-se de maneira espontânea, sem se propor racionalmente para se tornar um líder. E, de acordo com Bennis (1996), a questão não é tornar-se líder, é tornar-se você mesmo, usar-se completamente – todas as suas capacidades, talento e energia – para que sua visão se manifeste.

Desse modo, o tutor atual deve ampliar as dimensões apontadas por Litwin (2001) e para ratificar este apontamento, é fundamental a colaboração de Covey (1994, p.48), ao apontar que valores, idéias, normas e ensinamentos "enobrecem, elevam, realizam, legitimam e inspiram as pessoas".

Para Covey (1994) os princípios/características que devem nortear o exercício da tutoria no atual cenário pode ser assim apontados:

a. os tutores estão **continuamente aprendendo**. Neste sentido, são pessoas que desenvolvem novas habilidades e interesses a partir da curiosidade e estão abertas a adquirir novos conhecimentos, pois são humildes o suficiente para admitir opiniões diferentes e divergentes das suas;

b. os tutores **constantemente voltados para o serviço do processo educacional**. Compreendem a vida como uma missão e não como uma carreira. O sucesso desses líderes está diretamente relacionado com o sucesso das pessoas que estão ao seu redor, com o todo, havendo uma atuação harmônica entre as pessoas;
c. os tutores possuem atitude otimista, e seu espírito é entusiasta. Essas pessoas **irradiam energia positiva** e contagiam todos a sua volta. Assim, percebe-se que o ambiente é influenciado pela postura que as pessoas assumem perante os fatos e situações da vida;
d. os tutores **acreditam nas outras pessoas**, mesmo que o potencial delas ainda esteja latente. Como conseqüência, essa postura cria um ambiente favorável ao crescimento e à oportunidade, podendo despontar potencialidades que não haviam se configurado ainda;
e. os tutores baseados em princípios conciliam aspectos físicos e mentais, desenvolvendo ambas as habilidades físicas e cognitivas. Por isso, **suas vidas são equilibradas**. Este equilíbrio refere-se também ao reconhecimento de seu próprio valor, porém, sem a necessidade de mostrar-se aos outros. Pessoas baseadas em princípios, "vivem sensatamente no presente, planejam cuidadosamente o futuro, e adaptam-se com flexibilidade às circunstâncias mutáveis". (Covey, 1994, p. 11);
f. os tutores sabem saborear a vida, pois possuem segurança, ao invés de dependerem de fatores externos. Portanto, **encaram a vida como uma aventura** e fazem ricas as suas vidas, pois não lhes faltam vontade de aprender;
g. a **sinergia** é uma outra característica dos tutores, eles devem ser catalisadoras de mudanças, melhoram qualquer situação em que se envolvem e trabalham de forma tão árdua quanto inteligente. Além disso, sabem trabalhar em equipe, relacionam-se de forma sincera, contribuindo para o processo criativo na solução de problemas.
h. complementando uma vida equilibrada, os tutores orientadas por princípios **exercitam-se pela auto-renovação**, representada pela congruência das quatro dimensões da personalidade humana: física, mental, emocional e espiritual.

Com a observância dos princípios acima descritos, a função essencial do tutor voltada para orientação didática tornar se menos árdua, pois grande parte dos alunos que não possui hábito de autoestudo, nem autodisciplina, característicos dessa modalidade de educação serão acolhidos de modo coerente, orientados e estimulados por meio do eles tem de potencial, e não pelo que apresentam como distorções diante do modelo EaD. É preciso que o tutor seja flexível, perseverante, persistente e sensível para lidar com eventuais dificuldades encontradas por parte dos alunos (ALVES; NOVA, 2003).

Por fim, o quadro de tutores previstos para o processo de mediação pedagógica deve especificar a relação numérica estudantes/tutor capaz de permitir interação no processo de aprendizagem. Como Santos e Silva (2009) apontava a construção do conhecimento deve considerar que a interação é um processo importante seja no modelo presencial ou na EaD. Cada pessoa é um ser social, relacional e participante de um processo único e histórico. Contudo, a este processo de ensino-aprendizagem é requisitado que aquele que ensina e o que aprende, estabeleçam uma relação entre eles.

4. Conclusões

Com as alterações ocorridas no cenário educacional nas últimas décadas, e a ampliação do modelo de ensino a distância no segmento das instituições de ensino superior, homem e neste caso, especificamente o tutor, passou a assumir uma posição primordial e elementar dentro da estrutura educacional da EaD. Não resta dúvida que no processo de ensino aprendizagem, a mediação realizada pelos tutores deve ser ratificada, qualificada e legitimada, pelas instituições da EaD.

Este processo ocorre conforme o modelo de gestão das IES envolvidas, e desse modo algumas estão preocupadas com a qualificação do seu capital humano e outras nem tanto. Todavia, este tema relacionado especificamente a análise da gestão, cabe um olhar detalhado, em outro momento.

Cabe ressaltar com o estudo ora apresentado, observou que ainda de modo incremental está acontecendo no cerne das organizações um processo global de desenvolvimento da EaD, o qual propicia e resulta no desenvolvimento das pessoas, do chamado "capital intelectual", neste caso os tutores, foco do estudo.

Percebeu-se que aos tutores foi direcionado uma formação de competências e técnicas que significativamente interferem no modelo de ensino proposto. Porém, o estudo, apresentou que o desenvolvimento de atitudes juntos aos tutores, e estes focados na perspectiva da liderança, legítima facilitadores integrados aos objetivos organizacionais, bem como ao modelo de EaD hoje ratificado.

Por conseqüência, a organização que ampliará o olhar sobre o tutor da EaD, como líder no seu segmento, e nas atribuições de suas funções, tenderá a se tornar uma organização com maior competitividade frente a outras organizações, o que torna isso um diferencial no atual cenário de crise.

Assim, a liderança vem recebendo, cada vez mais, um papel de destaque nas organizações, que no decorrer do século XX privilegiaram o gerenciamento. Da mesma forma, uma nova era no desenvolvimento das organizações vem surgindo, é a era do conhecimento, onde a liderança não se restringe aos cargos ocupados pela gestão, mas ampliou-se às pessoas capazes orientar, de disseminar o conhecimento, o espírito de liderança e de equipe.

Hoje, entretanto, espera-se das organizações da EaD o respeito e a valorização da figura humana, tutor ou professor. E, neste contexto, cabe a gestão, o desenvolvimento do profissional, respeitando-o enquanto indivíduo, o seu espaço substancial dentro da racionalidade instrumental da organização.

Referências

ALVES, L.; NOVA, C. (2003). *Educação a Distância:* Uma Nova Concepção de Aprendizagem e Interatividade. São Paulo: Futura.

BENNIS, Warren. (1996). *A Formação do Líder*. São Paulo: Atlas.

BERNARDO, V. (2015). *Educação a distância*: fundamentos. Universidade Federal de São Paulo UNIFESP. Disponível em: http://www.virtual.epm.br/material/tis/enf.htm Acesso em: 14 set.

BRUNO, A. R.; LEMGRUBER, M. S. (2010). Docência na educação online: professorar e (ou) tutorar? In: BRUNO, A. R. et al. *Tem professor na rede*. Juiz de Fora: UFJF.

COVEY, Stephen R. (1994). *Liderança baseada em princípios*. Rio de Janeiro: Campus.

FARIA, J. G. (2015). *A construção do conhecimento em cursos a distância*. 2010. Disponível em: <http:// pt.scribd.com/doc/26856323/A-CONSTRUCAO-DO-CONHECIMENTO-EM-CURSOS-DE-EAD>. Acesso em: 10 out.

LITWIN, E. (Org.). (2001). *Educação a Distância:* Temas para Debate de uma Nova Agenda Educativa. Porto Alegre: Artmed.

NISKIER, A. (1999). *Educação a Distância:* A Tecnologia da Esperança. São Paulo: Loyola.

PALLO, Rena M.; PRATT, Keith. (2002). *O aluno virtual:* um guia para trabalhar com estudantes on- -line Porto Alegre: Artmed.

SALVUCCI, M.; LISBOA, M.; MENDES, Nelson C. (2015). Educação a Distância no Brasil: Fundamentos legais e implementação. *ASSOCIAÇÃO BRASILEIRA DE EDUCAÇÃO A DISTÂNCIA*, s.n. , vol. 11, 2012. Disponível na internet: < http://www.abed.org.br/revistacientifica/Revista_PDF_Doc/2012/artigo_04_v112012.pdf>. Acesso em: 10 out.

SANTOS, R. M.; SILVA, P. (2009). A didática da EAD virtual. In: *Anais do II Seminário Educação, Comunicação, Inclusão e Interculturalidade*. EDaPECI.

SILVA, M. A.; SANTOS, E. (Org.). (2006). *Avaliação da aprendizagem em educação on-line:* fundamentos, interfaces e dispositivos, relatos de experiências. Loyola.

SILVA, M.; PESCE, L.; ZUIN, A. (Org.). (2010). *Educação on-line:* cenário, formação e questões didático metodológicas. Rio de Janeiro: Wak.

TORI, R. (2010). *Educação sem distância:* as tecnologias interativas. São Paulo: Senac SP.

VASCONCELOS, S. P. G. (2010). *Educação a Distância*: histórico e perspectivas. Universidade do Estado do Rio de Janeiro (UERJ). Disponível em:http://www.filosofia.org.br/viiifelin/19htm . Acesso em: 08 jan.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 23/10/2018
Aceitação/Acceptance: 26/12/2018

Minería de datos educativa para identificar la relación entre cociente intelectual, estilos de aprendizaje, inteligencia emocional e inteligencias múltiples de estudiantes de ingeniería

Pedro Arevalo-Marin[1], Yonder Cabrera-Piedra[1], Miguel Cabrera-Sarango[1], Jonathan Caicedo-Rodriguez[1], Andres Camacho-Veliz[1], Sergio Cartuche-Morocho[1], Cristian Castillo-Aguilar[1], Daniel Castillo-Betancourt[1], Jason Chamba-Tigre[1], Alex Condoy-Carrion[1], Monica Coronel-Cardenas[1], Jorlan Elizalde-Cando[1], Elvis Freire-Feijoo[1], Adriana Gomez-Jara[1], Johnny Gonzalez-Guaman[1], Manuel Mora-Medina[1], Andres Morocho-Cumbicus[1], Maria-Jose Ojeda-Bazaran[1], Nelson Puchaicela-Godoy[1], Guido Quezada-Alvarez[1], Nayo Salinas-Minga[1], Javier Sarango-Espinoza[1], Luis Chamba-Eras[1], Milton Labanda-Jaramillo[1], Edison Coronel-Romero[1], Jose-Luis Granda[1] y Maria Roman-Sanchez[2]

{pdarevalom, jycabrerap, macabreras, jicaicedor, dacamachov, sdcartuchem, cecastillo, adcastillob, jachambat, arcondoyc, mncoronelc, jielizaldec, ecfreiref, acgomezj, jfgonzalezg, mamoram, admorochoc, mjojedab, nrouchaicelag, ggquezadaa, nfsalinasm, jasarangoe, lachamba, miltonlab, edisoncor, jose.l.granda}@unl.edu.ec, elizabeth.roman@educacion.gob.ec

[1] Grupo de Investigación en Tecnologías de la Información y Comunicación (GITIC), Carrera de Ingeniería en Sistemas, Facultad de Energía, Universidad Nacional de Loja, Av. Pío Jaramillo Alvarado y Reinaldo Espinosa, EC110110, Loja, Ecuador.

[2] Unidad Educativa Fiscomisional "Nuestra Señora del Rosario", Ministerio de Educación, EC110350, Catamayo, Ecuador.

Pages: 48–63

Resumen: El objetivo principal de este trabajo es identificar por medio de la minería de datos educativa (MDE), la relación que existe entre el cociente intelectual (CI), inteligencias múltiples (IM), inteligencia emocional (IE) y estilos de aprendizaje (EA) de 282 estudiantes universitarios de ingeniería, que será un instrumento objetivo para la toma de decisiones en el contexto educativo. Esta investigación fue exploratoria y descriptiva. La investigación exploratoria permitió el acercamiento al fenómeno objeto de estudio, en cambio, la investigación descriptiva por medio de la MDE permitió identificar la relación entre el CI, IM, IE y EA. La población estudiada tiene un CI normal (media=108.7), encontrándose además que, la IM verbal (79.1 %) y visual espacial (64.2 %) deben desarrollarse o potenciarse, mientras que, la IM lógico matemática (64.9 %) predomina. Además, la IE claridad (46.5 %) en conjunto con el EA activo-reflexivo (68.8 %) sobresalen en el grupo de estudio.

Palabras-clave: Minería de Datos Educativa; Inteligencia Artificial; Inteligencias Múltiples; Estilos de Aprendizaje; Inteligencia Emocional

Educational data mining to identify the relationship between IQ, learning styles, emotional intelligence and multiple intelligences of engineering students

Abstract: The aim of this work is to identify through the educational data mining (EDM), the relationship that exists between the intellectual quotient (IQ), learning styles (LS), emotional intelligence (EI) and multiple intelligences (MI) of 282 engineering university students, which will be an objective instrument for decision-making in the educational context. The research was exploratory and descriptive. The exploratory investigation allowed the approach to the phenomenon object of study, however, the descriptive investigation by means of the EDM allowed to identify the relationship between the IQ, MI, EI and LS. The population studied has a normal IQ (mean = 108.7), and it is also found that verbal MI (79.1%) and visual spatial MI (64.2%) must be developed or enhanced, while mathematical logical MI (64.9%) predominates. In addition, the IE clarity (46.5%) in conjunction with the active-reflexive EA (68.8%) stand out in the study group.

Keywords*:* educational data mining; artificial intelligence; multiple intelligences; learning styles; emotional intelligence

1. Introducción

Alan Turing en el año de 1950 planteó la pregunta “Can machines think?”, cuyo objetivo fue proponer a la comunidad científica la teoría de que algún día las máquinas podrían imitar la inteligencia humana y que sea indistinguible por el hombre (Aliseda, 2013; Gershenson, 2013; Turing, 1950). Partiendo de esa teoría, nace la inteligencia artificial (IA), un área de investigación multidisciplinaria dentro de las ciencias de la computación (CC) y con diversos campos de aplicación como la medicina, geología, economía, educación, entre otros. Actualmente, en la educación, la IA cumple un papel preponderante para mejorar o potenciar los procesos de enseñanza-aprendizaje en los diversos escenarios, que van desde la educación presencial, mixta o virtual y que se ven beneficiados por el valor agregado que la IA le brinda a la educación en todos sus niveles. Muchos artefactos a nivel de software y/o hardware se construyen con el objetivo de apoyar a la educación y con ello brindar apoyo personalizado y adaptado a las necesidades individuales o grupales en el proceso de aprendizaje. Desde los sistemas tutores inteligentes (STI) (Machin Torres, 2018) hasta los salones de clases inteligentes (SaCI) (Aguilar et al., 2016; Chamba-Eras & Aguilar, 2017), la aplicación práctica de la IA es una muestra fehaciente de que los resultados en el aprendizaje son favorables y pueden ser usados como herramientas de apoyo al profesor y directivos. Por otro lado, tomando como base la estadística, las bases de datos, sistemas de información y las CC, se encuentra la minería de datos (MD), cuyo objetivo general es extraer información de un conjunto de datos y transformarla en una estructura comprensible para su uso posterior. El comprender y aplicar los conceptos de la MD en el contexto educativo se le denomina minería de datos educativa (MDE) (Montoya Lince, Vargas Bonilla, & Contreras Olivares, 2015; Sin & Muthu, 2015) y que actualmente es una línea de investigación con mucho futuro y que muchos grupos de investigación a nivel mundial apuestan todos sus esfuerzos (Agudo-Peregrina et al., 2017; Lemos dos Santos, Cechinel, Carvalho Nunes,

& Ochoa, 2017). En algunos países en vías de desarrollo como el Ecuador, encontramos algunos escenarios de aprendizaje en la educación superior, donde el profesor al trabajar con una asignatura no considera características individuales de sus alumnos, sino que preparan actividades o recursos generales, que son impartidas como clases magistrales o planificadas como actividades autónomas personales o en grupo, donde se asume que todos deberán aprender de la misma forma y ritmo, pero que en la realidad y está demostrado científicamente, que cada individuo aprende, siente, construye modelos de la realidad y se emociona de manera diferente (Ortiz Torres, E; Aguilera, E.; González Pérez, 2010).

El objetivo principal de este trabajo es identificar la relación entre las variables cociente intelectual, estilo de aprendizaje, inteligencia emocional e inteligencias múltiples, y con ello, obtener un primer modelo para la toma de decisiones en la adecuación de escenarios de aprendizaje para cada alumno por parte del profesorado de Carreras de ingeniería.

2. Revisión de literatura

2.1. Cociente intelectual (CI)

En la investigación psicológica científica, Galton utilizó medidas en términos de aptitudes psicofísicas y determinó la distribución normal de la inteligencia, que sigue un modelo de curva normal en el que el mayor porcentaje de la población se sitúa en torno a la media y porcentajes sensiblemente inferiores en los extremos superior e inferior. Por tal motivo, se trata de identificar las conductas inteligentes mediante test de inteligencias (Galway, 2014) que se aplican a una muestra muy grande de personas, con el fin de obtener los patrones "normales" o baremos en que esas conductas o capacidades se adquieren. A este tipo de inteligencia se le denomina inteligencia psicométrica. El CI es el índice habitual empleado para clasificar intelectualmente a las personas; consiste en el cociente, multiplicado por cien, entre la edad mental y la edad cronológica de la persona (Benavides, Maz, Castro, & Blanco, 2004; Molero, Sáiz, & Esteban, 1998).

2.2. Inteligencias múltiples (IM)

Gardner manifiesta que no existen los métodos precisos para medir la inteligencia pura de los individuos, por lo tanto, propone su teoría donde cada individuo posee múltiples inteligencias, las cuales se constituyen como herramientas que se desarrollan en espacios vivos y estimulantes para el aprendizaje. Las IM propuestas por Gardner son: lingüísticas, matemáticas, kinestésicas, musicales, naturalistas, espaciales, interpersonales e intrapersonales (Gardner, 1998).

2.3. Inteligencia emocional (IE)

Las relaciones que se tiene en el ámbito personal y social, está ligada al nivel de IE, mientras mayor sea el grado de IE, el individuo presenta un mejor rendimiento académico, mayor grado de relación familiar y social, disminuye los niveles de depresión y conflictos intrapersonales. La inteligencia emocional implica la habilidad de percibir,

valorar y expresar emociones con precisión, acceder y generar sentimientos para facilitar el pensamiento, comprender emociones y conocimiento emocional y para regular emociones que promuevan el crecimiento intelectual y emocional (Mayer & Salovey, 1997).

2.4. Estilos de aprendizajes (EA)

Los EA son los rasgos cognitivos, afectivos y fisiológicos que sirven como indicadores relativamente estables, de cómo los alumnos perciben interacciones y responden a sus ambientes de aprendizaje (Sternberg, 1999). Los EA en estudiantes universitarios constituyen características muy importantes que deben estar presentes por el profesorado cuando se vayan a elaborar y ejecutar estrategias didácticas, ya que representan fenómenos complejos de comunicación inter e intrapersonal en cada uno de los estudiantes (Aguilera Pupo & Ortiz Torres, 2009). Aunque los estudiantes posean uno o más EA, se reconoce la flexibilidad y las posibilidades de poder cambiar o reajustar el estilo para el logro de un aprendizaje más eficiente (Sternberg, 1999), por lo que es importante hacer un seguimiento del progreso que se ha obtenido luego de aplicar las estrategias didácticas a los estudiantes (Ortiz Torres & Aguilera Pupo, 2005).

2.5. Minería de datos educativa (MDE)

La minería de datos (MD) es un campo de las CC, cuyo objetivo es descubrir patrones en grandes volúmenes de conjuntos de datos. Usa técnicas de la IA, aprendizaje automático, estadística y sistemas de bases de datos. El proceso de MD consiste en extraer información de un conjunto de datos y transformarla en una estructura comprensible para su uso posterior (Núñez Cardenas, Hernández Palacios, Tomás Mariano, & Felipe Redondo, 2013). Dentro de la MD se emplean tareas predictivas y descriptivas, las primeras, se orientan a predecir valores de salida de una variable objetivo, que se realiza a partir de un conjunto de variables predictoras y existen tareas de clasificación y regresión, las segundas, se orientan a acciones descriptivas como describir patrones o datos, se aplican cuando el conjunto de datos no es robusto para aplicar tareas predictivas, y encontramos tareas de clustering/segmentación y asociación (Almendrales Jiménez, 2016). En el contexto educativo, encontramos a la MDE que permite responder preguntas sobre ¿qué sabe realmente un estudiante? y ¿cómo está aprendiendo? De esta forma, la MDE permite descubrir información útil que ayuda a los profesores y autoridades de las instituciones interesadas en determinar la manera más pertinente para guiar a sus estudiantes, maximizando su aprendizaje (Montoya Lince et al., 2015) y también, clasificar a los mismos en grupos en función de sus necesidades de orientación y seguimiento (Bogarín Vega, Romero Morales, & Cerezo Menéndez, 2015; Del Río & Pineda, 2016).

3. Trabajos relacionados

En la Tabla 1, se mencionan los 22 trabajos relacionados con el objeto de estudio y que permiten sustentar académicamente la investigación en la sección 6, al identificar por medio de la MDE la relación entre las variables.

Trabajo	Objeto de estudio	Hallazgo
(Maureira, Flores, & Ravanal, 2015)	CI	No existe diferencia del CI entre varones y mujeres
(González-Arias & Alucema, 2015)	CI	Universitarios con mayor CI que particulares, y posiblemente varía entre carreras
(Oliveira Rosseti, Sant'Ana Rabelo, F. Almeida de Sá Leme, Pacanaro, & Bernd Guntert, 2009)	CI	Los varones tienen mayor CI
(Ariza-Hernández, 2017)	IE	Existe relación significativa entre rendimiento académico y la IE
(Enríquez Anchondo, 2011)	IE	Conocer los índices de IE y la influencia en un contexto cultural
(Adorjan, 2014)	IM	Evitar deserción, planificando actividades vinculadas a las IM
(Paredes Barragán, 2008)	EA	Personas con diferentes EA permiten resolver problemas más complejos
(Felder & Silverman, 1988)	EA	La mayoría de estudiantes de ingeniería son visuales, sensoriales, inductivos y activos
(Ramírez León & Rosas Espín, 2014)	EA	Existen características de EA que permiten agrupar estudiantes
(Durán & Costaguta, 2007)	MDE	Se identificaron los EA por medio de algoritmos de agrupamiento
(Almendrales Jiménez, 2016)	MDE	Se identificaron los EA que influyen en el rendimiento académico
(Olague Sánchez & Torres Ovalle, 2009)	MDE	Patrones de aprendizaje de los EA en un LMS
(Bogarín Vega et al., 2015)	MDE	Identificar la ruta de aprendizaje en un LMS
(Montoya Lince et al., 2015)	MDE	Patrones de interacción en un LMS
(Núñez Cardenas et al., 2013)	MDE	Descubrir los EA dominantes
(Coello Blanco, 2016)	MDE	Descubrir los EA dominantes en estudiantes de informática
(Gabino & Salguero, 2017)	MDE	Detección de EA por medio de reglas de decisión
(Machin Torres, 2018)	MDE	Se desarrolla un software combinando IM y EA
(Aldás Jarrín, 2010)	IE e IM	Relación entre la IE y la personalidad
(Gusqui Cayo, 2017)	IE	Los varones expresan los sentimientos, las mujeres regulan sus emociones
(Aldás Salazar, 2017)	IM	Las IM más desarrolladas en ingeniería son: lógica matemática, interpersonal y lingüística
(Bravo Mancero & Urquizo Alcivar, 2016)	IE e IM	No existe relación entre el razonamiento lógico y los factores de la IE

Tabla 1 – Trabajos relacionados al objeto de estudio y sus hallazgos.

4. Metodología

4.1. Objetivo

Identificar la relación entre el cociente intelectual, estilo de aprendizaje, inteligencia emocional e inteligencias múltiples de estudiantes de ingeniería.

4.2. Proceso y métodos

La investigación fue de tipo exploratoria y descriptiva. La investigación exploratoria permitió el acercamiento al fenómeno objeto de estudio (búsqueda bibliográfica, identificación, selección de cuestionarios, planificación de la experimentación), en cambio, la investigación descriptiva permitió describir la realidad del objeto de estudio. Los investigadores responsables del estudio fueron técnicos y profesores del Laboratorio de Inteligencia Artificial, en conjunto, con una experta en pedagogía.

4.3. Participantes

La población matriculada en la Carrera de Ingeniería en Sistemas (CIS) de la Universidad Nacional de Loja (Ecuador), según el reporte del sistema de gestión académico[1] en el período abril – septiembre del 2018 es 325 estudiantes, distribuidos en primero, segundo, tercero, cuarto, quinto, sexto y noveno ciclo. Se usó un muestreo no probabilístico por conveniencia, es decir se trató de involucrar a la mayor cantidad de matriculados que asistieron en la semana de aplicación de los instrumentos (04-08/06/2018), obteniendo 282 observaciones (86.8 %).

4.4. Instrumentos y procedimientos

Los instrumentos fueron[2]:

- Consentimiento informado: permitió registrar la participación voluntaria en la investigación.
- Cuestionario demográfico: permitió obtener la edad, género, lugar de nacimiento, ciclo y paralelo.
- Cuestionario para el CI: se usó el propuesto por RAVEN, compuesto por 39 preguntas (Galway, 2014).
- Cuestionario para las IM: se usó el propuesto por Howard Gardner, compuesto por 35 preguntas (Fernández González & Mihura de Rosa, 2015).
- Cuestionario para el EA: se usó el propuesto por Felder y Silverman, compuesto por 44 preguntas (Paredes Barragán, 2008).
- Cuestionario para la IE: se usó la Escala Rasgo de Metaconocimientos sobre Estados Emocionales en su versión para español, compuesto por 24 preguntas (Oliva, Antolín, Pertegal, Ríos, & Parra, 2004).
- Metodología para la MDE: CRISP-DM (Hernández Cedano, 2015; Montoya Lince et al., 2015).

[1] https://docentes.unl.edu.ec/ [Fecha de acceso: 03 de junio del 2018]
[2] Los instrumentos en formato PDF se encuentran en el repositorio: https://github.com/lachamba/IALab18

- Tarea descriptiva de la MDE: se usaron las reglas de asociación (Apriori) y el agrupamiento SimpleKmeans (Coello Blanco, 2016; Funes, 2008).
- Software para MDE: Weka (Hall et al., 2009).

Para alcanzar el objetivo planteado, se planificó el siguiente procedimiento:

1. Fase 1: aplicación manual de los instrumentos (5 cuestionarios en un cuadernillo) para recoger la información en un banco de datos (dataset).
2. Fase 2: ejecución de CRISP-DM.

5. Resultados

5.1. Fase 1

Se obtuvo 286 cuadernillos de la recolección manual, que fueron procesados usando hojas de cálculo dinámicas para obtener el dataset de 286x20 en formato CSV, posteriormente, se transformó el dataset en formato Weka (ARFF)[3]. En la Tabla 2 se resume los estadísticos del dataset usando Weka.

Variable	Tipo	Estadísticos
Edad	Numeric	Min: 16; Max: 33; Mean: 21.365; StdDev:2.804
Genero	Nominal	Count (Femenino): 65 Count (Masculino): 217
Lugar	Nominal	Count (Loja): 160 Count (Zamora): 16
Ciclo	Nominal	Count (Noveno): 67 Count (Tercero): 48
Paralelo	Nominal	Count (A): 157 Count (B): 106 Count (C): 19
Cociente intelectual	Numeric	Min: 87; Max: 135; Mean: 108.734; StdDev:9.1
InteligenciaMultipleVerbal	Nominal	Count (No tiene): 223 Count (Tiene): 53 Count (Sobresaliente): 6
InteligenciaMultipleLogicoMatematica	Nominal	Count (No tiene): 99 Count (Tiene): 86 Count (Sobresaliente): 97
InteligenciaMultipleVisualEspacial	Nominal	Count (No tiene): 181 Count (Tiene): 77 Count (Sobresaliente): 24
InteligenciaMultipleKinestesicaCorporal	Nominal	Count (No tiene): 210 Count (Tiene): 58 Count (Sobresaliente): 14

[3] El dataset en formato CSV, ARFF, así como las hojas de cálculo dinámicas se encuentran en el repositorio: https://github.com/lachamba/IALab18

Variable	Tipo	Estadísticos
InteligenciaMultipleMusicalRitmica	Nominal	Count (No tiene): 128 Count (Tiene): 94 Count (Sobresaliente): 60
InteligenciaMultipleIntrapersonal	Nominal	Count (No tiene): 134 Count (Tiene): 83 Count (Sobresaliente): 65
InteligenciaMultipleInterpersonal	Nominal	Count (No tiene): 157 Count (No tiene): 75 Count (Sobresaliente): 50
InteligenciaEmocionalAtencion	Nominal	Count (Adecuado): 139 Count (Excelente): 24 Count (Mejorar): 119
InteligenciaEmocionalClaridad	Nominal	Count (Adecuado): 29 Count (Excelente): 131 Count (Mejorar): 122
InteligenciaEmocionalReparacion	Nominal	Count (Adecuado): 43 Count (Excelente): 64 Count (Mejorar): 175
EstiloAprendizajeActivoReflexivo	Nominal	Count > (Equilibrio entre AR): 109 Count > (Equilibrio entre RA): 85
EstiloAprendizajeSensitivoIntuitivo	Nominal	Count > (Equilibrio entre SI): 107 Count > (Más SI): 101
EstiloAprendizajeVisualVerbal	Nominal	Count > (Más VV): 102 Count > (Equilibrio entre VV): 80
EstiloAprendizajeSecuencialGlobal	Nominal	Count > (Equilibrio entre SG): 110 Count > (Más SG): 73

Tabla 2 – Resumen estadístico de las variables del objeto de estudio.

5.2. Fase 2

De acuerdo a CRISP-DM se muestran los resultados:

Paso A. Entendimiento del modelo de datos: de acuerdo al dataset obtenido en la fase 1, conformado de 20 variables, se seleccionaron en Weka las variables para la aplicación de la técnica de agrupamiento y las reglas de asociación.

Paso B. Preparación de datos: se procedió a transformar las variables cuantitativas (edad y cociente intelectual) a variables cualitativas, para ello se usó el filtro (*Discretize*) para atributos no supervisados en Weka, configurando 4 rangos (*bins*) y el *useEqualFrequency* en *True* para obtener rangos con el mismo número de frecuencias.

Paso C. Modelado (reglas de asociación y agrupamiento): con las variables estandarizadas, se procedió a identificar las reglas de asociación usando el algoritmo "*A priori*" con las configuraciones por defecto de Weka, posteriormente, se realizó el agrupamiento usando el algoritmo *SimpleKmeans*. Las Tablas 3 y 4 muestran los resultados obtenidos.

Reglas de asociación (Apriori)
1. CocienteIntelectual='(106.5-112.5]' InteligenciaMultipleInterpersonal=0 47 ==> InteligenciaMultipleVerbal=0 44 <conf:(0.94)> lift:(1.18) lev:(0.02) [6] conv:(2.46)
2. CocienteIntelectual='(106.5-112.5]' InteligenciaMultipleVisualEspacial=0 InteligenciaMultipleKinestesicaCorporal=0 46 ==> InteligenciaMultipleVerbal=0 43 <conf:(0.93)> lift:(1.18) lev:(0.02) [6] conv:(2.41)
3. InteligenciaMultipleIntrapersonal=0 InteligenciaMultipleInterpersonal=0 InteligenciaEmocionalAtencion=Adecuado 46 ==> InteligenciaMultipleVerbal=0 43 <conf:(0.93)> lift:(1.18) lev:(0.02) [6] conv:(2.41)
4. InteligenciaMultipleLogicoMatematica=0 InteligenciaMultipleInterpersonal=0 62 ==> InteligenciaMultipleKinestesicaCorporal=0 57 <conf:(0.92)> lift:(1.23) lev:(0.04) [10] conv:(2.64)
5. InteligenciaMultipleLogicoMatematica=0 InteligenciaMultipleVisualEspacial=0 InteligenciaMultipleInterpersonal=0 46 ==> InteligenciaMultipleKinestesicaCorporal=0 42 <conf:(0.91)> lift:(1.23) lev:(0.03) [7] conv:(2.35)
6. InteligenciaMultipleVisualEspacial=0 InteligenciaMultipleKinestesicaCorporal=0 InteligenciaMultipleInterpersonal=0 InteligenciaEmocionalAtencion=Adecuado 46 ==> InteligenciaMultipleVerbal=0 42 <conf:(0.91)> lift:(1.15) lev:(0.02) [5] conv:(1.92)
7. InteligenciaMultipleKinestesicaCorporal=0 InteligenciaMultipleIntrapersonal=0 InteligenciaMultipleInterpersonal=0 InteligenciaEmocionalReparacion=Adecuado 46 ==> InteligenciaMultipleVerbal=0 42 <conf:(0.91)> lift:(1.15) lev:(0.02) [5] conv:(1.92)
8. CocienteIntelectual='(106.5-112.5]' InteligenciaMultipleVisualEspacial=0 56 ==> InteligenciaMultipleVerbal=0 51 <conf:(0.91)> lift:(1.15) lev:(0.02) [6] conv:(1.95)
9. InteligenciaMultipleVerbal=0 InteligenciaMultipleLogicoMatematica=0 InteligenciaMultipleInterpersonal=0 56 ==> InteligenciaMultipleKinestesicaCorporal=0 51 <conf:(0.91)> lift:(1.22) lev:(0.03) [9] conv:(2.38)
10. InteligenciaMultipleKinestesicaCorporal=0 InteligenciaMultipleMusicalRitmica=0 InteligenciaMultipleInterpersonal=0 55 ==> InteligenciaMultipleVerbal=0 50 <conf:(0.91)> lift:(1.15) lev:(0.02) [6] conv:(1.92)

Tabla 3 – Ejecución del algoritmo de reglas de asociación.

6. Discusión

Dentro de la información demográfica, el rango de edad de los estudiantes de CIS se encuentra entre 16-33 años, y las ciudades de Loja (56.7 %) y Zamora (5.7 %), son los lugares de nacimiento más frecuentes de los estudiantes. El género masculino representa el 76 % y el femenino el 24 % de los estudiantes de la CIS que participaron del estudio. Tomando como base los estadísticos de la Tabla 2, y considerando la escala de Raven (González-Arias & Alucema, 2015), se identifica que la media del *cociente intelectual (CI)* es 108.7, valor que se encuentra en el intervalo 85–115, por lo tanto, los estudiantes de la CIS tienen un *CI normal* corroborando lo manifestado en los trabajos relacionados en la Tabla 1 (González-Arias & Alucema, 2015; Oliveira Rosseti et al., 2009). Además, En las inteligencias *múltiples verbal y visual espacial*, el valor "No tiene" se presenta en 223 (79.1 %) y 181 (64.2 %) casos respectivamente; siendo estas dos IM las que necesitan ser desarrolladas, a través de actividades académicas adecuadas (para la *verbal*: reforzar la lectura y escritura de artículos/libros científicos/técnicos; para la *visual espacial*: reforzar el modelado en 3D, crear mapas conceptuales y juegos serios) (Secretaría de Educación, 2004). Por otro lado, la *inteligencia emocional claridad* presenta 131 (46.5

Agrupamiento (SimpleKMeans)

Initial starting points (random):
Cluster 0:
‘\’(19.5-21.5]\”,’\’(106.5-112.5]\”,0,0,0,0,2,0,0,Mejorar,Mejorar,Adecuado,13BAR,13BSI,57AVV,13ASG
Cluster 1: ‘\’(21.5-23.5]\”,’\’(-inf-102.5]\”,0,0,0,0,0,2,2,Adecuado,Adecuado,Excelente,13AAR,57ASI,911AVV,13BSG
Cluster 2:
‘\’(19.5-21.5]\”,’\’(102.5-106.5]\”,0,0,0,0,1,1,2,Mejorar,Adecuado,Adecuado,57AAR,13ASI,57AVV,13ASG
Final cluster centroids:

		Cluster#		
Attribute	Full Data	0	1	2
	(282.0)	(137.0)	(89.0)	(56.0)
Edad	‘(-inf-19.5]’	‘(-inf-19.5]’	‘(23.5-inf)’	‘(19.5-21.5]’
CocienteIntelectual	‘(106.5-112.5]’	‘(106.5-112.5]’	‘(-inf-102.5]’	‘(102.5-106.5]’
InteligenciaMultipleVerbal	0	0	0	0
InteligenciaMultipleLogicoMatematica	0	0	2	2
InteligenciaMultipleVisualEspacial	0	0	0	0
InteligenciaMultipleKinestesicaCorporal	0	0	0	0
InteligenciaMultipleMusicalRitmica	0	0	0	1
InteligenciaMultipleIntrapersonal	0	0	2	1
InteligenciaMultipleInterpersonal	0	0	1	0
InteligenciaEmocionalAtencion	Adecuado	Mejorar	Adecuado	Adecuado
InteligenciaEmocionalClaridad	Mejorar	Mejorar	Adecuado	Adecuado
InteligenciaEmocionalReparacion	Adecuado	Adecuado	Adecuado	Adecuado
EstiloAprendizajeActivoReflexivo	13AAR	13BAR	13AAR	57AAR
EstiloAprendizajeSensitivoIntuitivo	13ASI	13ASI	57ASI	13ASI
EstiloAprendizajeVisualVerbal	57AVV	13AVV	57AVV	57AVV
EstiloAprendizajeSecuencialGlobal	13ASG	13ASG	13ASG	13ASG

Clustered Instances:
0 137 (48%)
1 89 (32%)
2 56 (20%)

Tabla 4 – Ejecución del algoritmo de agrupamiento.

%) casos en el valor “Excelente”, siendo la IE que más predomina en los estudiantes (saben comprender sus estados emocionales). La *inteligencia emocional reparación* presenta 175 (62.1 %) casos en el valor “Mejorar”, con ello, es la IE que más se debe desarrollar o potenciar en los estudiantes de la CIS (deben aprender a regular sus estados emocionales), coincidiendo con los trabajos relacionados (Ariza-Hernández, 2017; Gusqui Cayo, 2017). La *inteligencia emocional atención* presenta 139 (49.3 %) casos en el valor “Adecuado”, es decir que los estudiantes de la CIS pueden sentir y expresar sus sentimientos de manera adecuada. Así mismo, se identifica que el EA que más predomina en 194 estudiantes (68.8 %) de la CIS es el “equilibrio entre el activo y reflexivo” corroborando lo manifestado en el trabajo de (Felder & Silverman, 1988), por lo tanto, en los estudiantes de la CIS existen preferencias de trabajos en equipo y de afrontar nuevos desafíos, y a la vez, existen estudiantes que les gusta trabajar en base a la experiencia y analizar con detenimiento las acciones a realizar (Secretaría de Educación, 2004).

Para identificar el primer conjunto de relaciones entre las variables, se consideró las tres primeras reglas de asociación obtenidas por medio de Weka (ver Tabla 3): **R1)** si el

CI de los estudiantes se encuentra en el rango de 106.5 a 112.5 (CI normal) y no tienen desarrollada la *inteligencia múltiple interpersonal*, entonces los estudiantes no tienen desarrollada la *inteligencia múltiple verbal*; **R2)** si el CI de los estudiantes se encuentra en el rango de 106.5 a 112.5 (*CI normal*) y no tienen desarrollada la *inteligencia múltiple visual espacial* y tampoco la *kinestésica corporal*, entonces los estudiantes no tienen desarrollada la *inteligencia múltiple verbal*; y **R3)** si los estudiantes no tienen desarrollada la *inteligencia múltiple intrapersonal e interpersonal* y la *inteligencia emocional atención* es adecuado, entonces los estudiantes no tienen desarrollada la *inteligencia múltiple verbal*. Con las reglas descritas, se puede observar que existen variables relacionadas como las IM, IE y CI que permiten identificar que la *inteligencia múltiple verbal* es la que se encuentra menos desarrollada en los estudiantes de la CIS. Si se potenciarán las inteligencias vinculadas, posiblemente se activaría o desarrollaría la *inteligencia múltiple verbal*.

Por otro lado, para identificar el segundo conjunto de relaciones entre las variables, se consideró los tres clústeres obtenidos por medio de Weka (ver Tabla 4), así: **clúster 0** (48 %), la edad de los estudiantes es hasta los 19 años, cuyo CI se encuentra en el rango de 106 a 112, no tienen desarrolladas ninguna de las IM, la *inteligencia emocional atención* y *claridad* deben mejorar, la *inteligencia emocional reparación* es *adecuado*, tienen *equilibrio* entre el *estilo de aprendizaje reflexivo-activo*, *sensitivo-intuitivo*, *visual-verbal*, *secuencial-global*; **clúster 1** (32 %), la edad de los estudiantes es mayor de los 23 años, cuyo CI se encuentra en el rango hasta 102, no tienen desarrolladas las *inteligencias múltiples verbal*, *visual espacial*, *kinestésica corporal*, *musical rítmica*, son *sobresalientes* en la *inteligencia múltiple lógico matemática* e *intrapersonal*, tienen la *inteligencia múltiple interpersonal*, la *inteligencia emocional atención*, *claridad y reparación* es *adecuado*, tienen *equilibrio* entre el *estilo de aprendizaje activo-reflexivo y secuencial-global*, son *más sensitivo-intuitivo y visual-verbal*; y **clúster 2** (20 %), la edad de los estudiantes se encuentra en el rango de 19 a 21 años, cuyo CI se encuentra en el rango de 102 a 106, no tienen desarrolladas las *inteligencias múltiples verbal*, *visual espacial*, *kinestésica corporal*, *interpersonal*, son *sobresalientes* en la *inteligencia múltiple lógico matemática* y tienen la *inteligencia múltiple musical rítmica e intrapersonal*, la *inteligencia emocional atención, claridad y reparación* es *adecuado*, tienen equilibrio entre el *estilo de aprendizaje sensitivo-intuitivo* y *secuencial-global*, son *más activo-reflexivo y visual-verbal*.

Por destacar, que los estudiantes de la CIS (dos de los clústeres), tienen la *inteligencia múltiple lógico matemática* característica de los ingenieros (Aldás Salazar, 2017), en contraste, en los tres clústeres se evidencia que la *inteligencia múltiple verbal* es la que requiere ser potenciada y desarrollada (ver Tabla 2); corroborando con los trabajos relacionados (Coello Blanco, 2016; Gabino & Salguero, 2017).

7. Conclusiones y trabajos futuros

Podemos concluir como una de las principales contribuciones del trabajo es que mediante el uso de las técnicas de reglas de asociación y agrupamiento se ha identificado dos conjuntos de relaciones entre las variables, cociente intelectual, inteligencias múltiples, inteligencia emocional y estilos de aprendizaje de los estudiantes de ingeniería. Por otro lado, una de las limitantes del trabajo es que la experimentación no se realizó con

toda la población de estudiantes (325) de la CIS, además, para validar los hallazgos de este trabajo, se podría usar otros cuestionarios existentes en la literatura científica por cada una de las variables. Este trabajo es una primera aproximación para identificar la relación entre las variables del objeto de estudio, y al ser un sistema muy dinámico, si se repitiera el experimento, los resultados pueden variar, por tal motivo, se recomienda realizar más experimentos en otros estudiantes de ingenierías vinculadas al estudio y así poder tener valores más aproximados. Todo el material usado en la investigación se encuentra de acceso libre, disponible para la comunidad científica en https://github.com/lachamba/IALab18.

Como líneas futuras, se tiene planificado por medio de la teoría de agentes inteligentes, construir en su primera versión la arquitectura del Agente *EstudianteCIS* que puede interactuar en una arquitectura de SMA en el contexto de un salón de clases inteligente (SaCI) (Aguilar et al., 2016). Además, se requiere aplicar técnicas didácticas y pedagógicas para reforzar las inteligencias múltiples con falencias, por ejemplo, construyendo un agente conversacional basado en PLN para potenciar la inteligencia lingüística. Así mismo, se está trabajando en obtener modelos básicos académicos de aprendizaje automático para resolver problemas de regresión, clasificación y agrupamiento usando las 286 observaciones o agregando más datos de otras ramas de la ingeniería u otras ciencias (Zhu, 2015). Finalmente, se pretende a corto plazo usar la computación emergente para identificar nuevos agrupamientos presentes en la población de estudio y trabajar de forma multidisciplinar, involucrando a profesionales de la psicología, educación e IA, esto permitirá dar una base teórica y metodológica a toda la investigación (Barba-Guamán, Valdiviezo-Diaz, & Aguilar, 2018).

Agradecimientos

Este trabajo forma parte de las actividades académicas y de investigación del Grupo de Investigación GITIC, agradecemos al profesorado, autoridades y alumnado de la Carrera de Ingeniería en Sistemas de la Universidad Nacional de Loja (UNL).

Contribuciones de autoría

Luis Chamba-Eras y Jose-Luis Granda concibieron el estudio, la minería de datos educativa y redactaron la versión inicial y final del artículo. Pedro Arevalo-Marin, Yonder Cabrera-Piedra, Miguel Cabrera-Sarango, Jonathan Caicedo-Rodriguez, Andres Camacho-Veliz, Sergio Cartuche-Morocho, Cristian Castillo-Aguilar, Daniel Castillo-Betancourt, Jason Chamba-Tigre, Alex Condoy-Carrion, Monica Coronel-Cardenas, Jorlan Elizalde-Cando, Elvis Freire-Feijoo, Adriana Gomez-Jara, Johnny Gonzalez-Guaman, Manuel Mora-Medina, Andres Morocho-Cumbicus, Maria-Jose Ojeda-Bazaran, Nelson Puchaicela-Godoy, Guido Quezada-Alvarez, Nayo Salinas-Minga, Javier Sarango-Espinoza han realizado conjuntamente distribuidos en equipos de tres personas la argumentación teórica y científica del objeto de estudio (estilos de aprendizaje, inteligencias múltiples, inteligencia emocional y cociente intelectual), la recolección de los datos, y la descripción de datos. Así mismo, Milton Labanda-Jaramillo, Edison Coronel-Romero y Maria Roman-Sanchez revisaron la redacción del texto y aportaron comentarios en la revisión de literatura sobre el objeto de estudio,

metodología y discusión y otros elementos para la mejora el artículo y aprobaron la versión final para su publicación.

Referencias

Adorjan, A. (2014). *Un enfoque de Inteligencias Múltiples y Competencias aplicado a la enseñanza inicial de la Programación*. Universidad ORT Uruguay.

Agudo-Peregrina, Á. F., Caeiro-Rodríguez, M., Conde, M. Á., Cruz-Benito, J., Delgado Kloos, C., Menchaca, I., … Robles Gómez, A. (2017). SNOLA: creando una Red sobre Analíticas de Aprendizaje en España. In *IV Congreso Internacional sobre aprendizaje, innovación y competividad* (pp. 69–73).

Aguilar, J., Chamba-Eras, L., & Cordero, J. (2016). *Specification of a smart classroom based on agent communities. Advances in Intelligent Systems and Computing* (Vol. 444).

Aguilera Pupo, E., & Ortiz Torres, E. (2009). Las investigaciones sobre los estilos de aprendizaje y sus modelos explicativos. *Revista Estilos de Aprendizaje*, *4*(2), 1–19.

Aldás Jarrín, J. C. (2010). *El desarrollo de las Inteligencias múltiples y la formación por competencias en el primer nivel de la Carrera de Medicina de la Facultad de Ciencias de la Salud de la Universidad Técnica de Ambato*. Universidad Técnica de Ambato.

Aldás Salazar, D. S. (2017). *Las Inteligencias Múltiples y su incidencia en el logro de competencias genéricas y específicas en los estudiantes de la Carrera de Ingeniería Industrial de la Universidad Técnica de Ambato, período académico Septiembre 2012-Febrero 2013*. Escuela Politécnica del Ejército.

Aliseda, A. (2013). ¿Inteligencia mecánica? *Ciencia*, *64*(4), 10–17.

Almendrales Jiménez, G. (2016). *Estudio sobre estilos de aprendizaje mediante Minería de Datos como apoyo a la gestión directiva de la institución educativa Joaquín Cárdenas Gómez del Municipio de San Carlos*. Universidad Pontificia Bolivariana.

Ariza-Hernández, M. L. (2017). *Influencia de la inteligencia emocional y los afectos en la relación maestro-alumno, en el rendimiento académico de estudiantes de educación superior*. *Educación y Educadores*.

Barba-Guamán, L., Valdiviezo-Diaz, P., & Aguilar, J. (2018). Gestión emergente de espacios colaborativos de aprendizaje. *Revista Ibérica de Sistemas e Tecnologias de Informação*, *4*(15), 271–281.

Benavides, M., Maz, A., Castro, E., & Blanco, R. (2004). *La Educación de Niños con Talento en Iberoamérica*.

Bogarín Vega, A., Romero Morales, C., & Cerezo Menéndez, R. (2015). Aplicando minería de datos para descubrir rutas de aprendizaje frecuentes en Moodle. *Revista de Educación Mediática y TIC*, *5*(1), 73–92.

Bravo Mancero, P., & Urquizo Alcivar, A. M. (2016). Razonamiento lógico abstracto e inteligencia emocional: trayectorias en la formación de estudiantes universitarios. *Sophía, Colección de Filosofía de La Educación*, *2*(21), 179.

Chamba-Eras, L., & Aguilar, J. (2017). Augmented Reality in a Smart Classroom - Case Study: SaCI. *Revista Iberoamericana de Tecnologias Del Aprendizaje*, *12*(4), 165–172.

Coello Blanco, L. (2016). Uso de algoritmos de agrupamiento en la determinación de los estilos de aprendizaje. *Congreso Universidad*, *5*(4), 30–42.

Del Río, C. A., & Pineda, J. A. (2016). Predicting academic performance in traditional environments at higher-education institutions using data mining : A review. *Ecos de La Academia*, *4*, 1–24.

Durán, E. B., & Costaguta, R. N. (2007). Minería de datos para descubrir estilos de aprendizaje. *Revista Iberoamericana de Educación*, *42*(2), 6.

Enríquez Anchondo, H. A. (2011). *Inteligencia Emocional Plena : Hacia un Programa de Regulación Emocional Basado en la Conciencia Plena*. Universidad de Málaga.

Felder, R., & Silverman, L. (1988). Learning and teaching styles in engineering education. *Engineering Education*, *78*(7), 674–681.

Fernández González, A., & Mihura de Rosa, D. (2015). Inteligencias Múltiples. *E-Motion. Revista de Educación, Motricidad e Investigación*, (4), 6–17.

Funes, A. (2008). *Agrupamiento Conceptual Jerárquico Basado en Distancias. XII Workshop de ...*. Universidad Politécnica de Valencia.

Gabino, R., & Salguero, A. (2017). Herramienta para detección de estilos de aprendizaje en estudiantes de educación superior. *Revista Tecnológica ESPOL – RTE*, *30*(3), 106–121.

Galway, S. (2014). Test de Inteligencia. In *Historia, Definición y Teorías de los Test de Inteligencia* (pp. 135–164).

Gardner, H. (1998). *Inteligencias múltiples*. Paidós Barcelona.

Gershenson, C. (2013). ¿Cómo medir la inteligencia de las máquinas? *Ciencia*, *64*(4), 18–23.

González-Arias, M. I., & Alucema, L. B. (2015). Formación universitaria e inteligencia en carreras de ingeniería y pedagogía. Revisitando un antiguo debate. *Formacion Universitaria*, *8*(6), 13–22.

Gusqui Cayo, L. H. (2017). *Inteligencia Emocional en Estudiantes de la Facultad de Ingeniería de la Universidad Nacional de Chimborazo. Riobamba. Marzo-Julio 2017*. Universidad Nacional de Chimborazo.

Hall, M., Holmes, G., Pfahringer, B., Witten, I., Hall, M., Frank, E., ... Witten, I. H. (2009). The WEKA data mining software: An update. *ACM SIGKDD Explorations Newsletter*, *11*(1), 10–18.

Hernández Cedano, J. Á. (2015). *Modelo de Minería de Datos para Identificación de Patrones que Influyen en el Aprovechamiento Académico*. Instituto Tecnológico de la Paz.

Lemos dos Santos, H., Cechinel, C., Carvalho Nunes, J. B., & Ochoa, X. (2017). An Initial Review of Learning Analytics in Latin America. In *Twelfth Latin American Conference on Learning Technologies (LACLO)* (pp. 1–9).

Machin Torres, I. T. (2018). Diseño del componente generación de tutoría para el aprendizaje de la programación orientada a objetos. *Revista ESPACIOS*, *39*(10), 1–15.

Maureira, F., Flores, E., & Ravanal, M. (2015). Capacidad intelectual de estudiantes de educación física de una universidad privada de Santiago de Chile. *Psiquiatría Universitaria*, *11*(4), 395–401.

Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence? In *Emotional development and emotional intelligence: Educational Implications* (pp. 3–31). New York: Harper Collins.

Molero, C., Sáiz, E., & Esteban, C. (1998). Revisión histórica del concepto de inteligencia: Una aproximacion a la inteligencia emocional. *Revista Latinoamericana de Psicología*, *30*, 11–30.

Montoya Lince, A., Vargas Bonilla, J., & Contreras Olivares, L. Y. (2015). Educational Data Mining (EDM) para la Determinación de Comportamientos en Estudiantes de Ingeniería del Modelo UDE@. In *Encuentro Internacional de Educación en Ingeniería* (pp. 1–12).

Núñez Cardenas, F. de J., Hernández Palacios, R., Tomás Mariano, V. T., & Felipe Redondo, A. M. (2013). Identificación de Estilos de Aprendizaje en Alumnos Universitarios de Computación de la Huasteca Hidalguense mediante Técnicas de Minería de Datos. *Ciencia Huasteca Boletín Científico de La Escuela Superior de Huejutla*, *1*(2), 1–7.

Olague Sánchez, J. R., & Torres Ovalle, S. (2009). Aplicación De Técnicas De Minería De Datos Y Sistemas De Gestión De Contenidos De Aprendizaje Para El Desarrollo De Un Sistema Informático de Aprendizaje de Programación de Computadoras. In *X Congreso Nacional de Investigación Educativa* (pp. 1–15).

Oliva, A., Antolín, L., Pertegal, M. A., Ríos, M., & Parra, A. (2004). Escala para la Evaluación de la Expresión, Manejo y Reconocimiento de Emociones. In *Instrumentos para la Evaluación de la Salud Mental y el Desarrollo Positivo Adolescente y los Activos Que lo Promueven* (pp. 144–159).

Oliveira Rosseti, M., Sant'Ana Rabelo, I., F. Almeida de Sá Leme, I., Pacanaro, S. V., & Bernd Guntert, I. (2009). Evidências de validade das Matrizes Progressivas Avançadas de Raven em universitários. *Psico-USF*, *14*(2), 177–184.

Ortiz Torres, E; Aguilera, E.; González Pérez, A. (2010). Los Estilos De Aprendizaje , La Superdotación Intelectual Y El Talento En Estudiantes Universitarios. *Revista Estilos de Aprendizaje*, *5*(5), 1–21.

Ortiz Torres, E., & Aguilera Pupo, E. (2005). Los estilos de aprendizaje de los estudiantes universitarios y sus implicaciones didácticas en la educación superior. *Revista Pedagogía Universitaria*, *X*(5), 1–9.

Paredes Barragán, P. (2008). *Una Propuesta de Incorporación de los Estilos de Aprendizaje a los Modelos de Usuario en Sistemas de Enseñanza Adaptativos*. Universidad Autónoma de Madrid.

Ramírez León, Y. del V., & Rosas Espín, D. (2014). Aplicación de la teoría de estilos de aprendizaje al diseño de contenidos didácticos en entornos virtuales. *Etic@net*, *2*(14), 8.

Secretaría de Educación, P. (2004). Manual de estilos de aprendizaje. *Material Autoinstruccional Para Docentes y Orientadores Educativos*.

Sin, K., & Muthu, L. (2015). Application of Big Data in Education Data Mining and Learning Analytics – a Literature Review. *ICTACT Journal on Soft Computing*, *05*(04), 1035–1049.

Sternberg, R. J. (1999). *Thinking styles*. Cambridge university press.

Turing, A. M. (1950). Computing Machinery and Intelligence. *MIND*, *59*(236), 433–460.

Zhu, X. (2015). Machine Teaching: An Inverse Problem to Machine Learning and an Approach Toward Optimal Education. In *Twenty-Ninth AAAI Conference on Artificial Intelligence* (pp. 4083–4087).

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 26/09/2018
Aceitação/Acceptance: 25/11/2018

Modelo Learning Vectors na Avaliação Formativa no AVA Moodle: de Emoticons a GIFs Animados

Gilvandenys Sales[1], Monck C. de Albuquerque[2], Ricky P. Paz[1], Eliana A. M. Leite[3]

denyssales@ifce.edu.br, {monckcna, rickypaz, elimoreiraead}@gmail.com

[1] Instituto Federal do Ceará - IFCE, Av 13 de Maio, 2081, 60.040-215, Fortaleza/Ceará, Brasil

[2] Instituto Federal da Bahia – IFBA, Estrada Vicinal para Tenda, s/n, 46.900-970, Seabra/Bahia, Brasil

[3] Universidade do Minho, Largo do Paço, 4704-553, Braga, Portugal

Pages: 64–76

Resumo: Este artigo descreve a concepção da inserção de GIFs como escala de avaliação formativa associada à menções qualitativas no Moodle. Tomou-se por fundamentação o Modelo LV com seus Vetores-Aprendizagem e escala iconográfica. Às menções qualitativas tipo: Muito Bom, Bom, Regular, Fraco, Não Satisfatório e Neutro, associaram-se GIFs. Foi feita uma validação na ferramenta Fóruns de discussão simples. Com a utilização do poder das cores e ícones animados, ou Gifs, a serem aplicados pelos Professores/Tutores em suas intervenções pedagógicas, espera-se tornar a sala de aula virtual em espaço mais afetivo e motivador e a avaliação em processo mais dinâmico.

Palavras-chave: Avaliação formativa; Modelo Learning Vectors; AVA Moodle; Menções qualitativas; GIFs.

Learning Vectors Model in Formative Assessment LMS Moodle: from emoticons to animated GIFs

Abstract: This paper describes the conception of the insertion of GIFs as a formative evaluation scale associated to the qualitative mentions in Moodle. The Model LV with its Learning-Vectors and Iconographic Scale was taken as a basis. To the qualitative mentions type: Very Good, Good, Fair, Poor, Not Satisfactory and Neutral, associated GIFs. A validation was performed on the Simple Discussion Forums tool. Using the power of animated colors and icons, or Gifs, to be applied by the Teachers / Tutors in their pedagogical interventions, it is hoped to make the virtual classroom in a more affective and motivating space and the evaluation in a more dynamic process.

Keywords*:* Formative assessment; Learning Vectors Model; LMS Moodle; Qualitative statements; GIFs.

1. Introdução

O modelo pedagógico de ensino presencial, Educação a Distância (EaD) ou híbrido, tem incluído cada vez mais comunidades virtuais ou espaços colaborativos online, desde aplicativos de troca de mensagens, vídeos a Ambientes Virtuais de Aprendizagem (AVA) com o propósito de dar suporte ao aluno, por serem acessíveis de qualquer local por um dispositivo conectado à internet, e ao adequar-se ao ritmo de aprendizagem de cada um, porém quando o aluno é avaliado e recebe a nota de participação de uma atividade nestes espaços virtuais, o processo acaba se tornando similar ao de uma sala de aula presencial ou mesmo pior, por se tratar de uma relação apenas entre usuário e máquina, mesmo que do outro lado esteja um profissional gerenciando suas ações. Ainda é um desafio em cursos com suporte online a implantação de mecanismos de avaliação realmente formativa disponibilizados nos AVA que possibilitem um "processo comunicativo de negociações" capazes de auxiliar nas intervenções pedagógicas por parte dos Professores/Tutores de forma mais significativa e afetiva (Santos, 2010). O foco da ação tutorial em AVA, deveria ser a avaliação e o constante acompanhamento do aluno como um "processo formativo que envolve subjetividades e tomadas de decisões contínuas" (Leite *et. al* , 2015, p.1232), acrescente-se a afetividade.

O design de interação das interfaces dos AVA, que priorizam espaços de emissão e recepção de mensagens ou arquivos, não facilitam a formação de vínculos afetivos, necessários muitas vezes ao processo de aprendizagem. Para Valente (2002) o "estar junto virtual" traz a sensação de pertencimento àquele espaço de aprendizagem evitando a solidão virtual e possível evasão, que constitui outro grande desafio em EaD. Nos AVA sobressaem os aspectos quantitativos: contabilização do número de acessos, notas nas diversas ferramentas síncronas ou assíncronas, além de dados e gráficos estatísticos. Nos AVA os "elementos subjetivos construídos de forma partilhada na interação, que poderiam ser inferidos sobre a qualidade da aprendizagem do aluno, não se revelam facilmente quando se consultam esses indicadores" (Sales, 2017).

Este trabalho apresenta a evolução da ferramenta e modelo de avaliação formativa, denominado Modelo Learning Vectors (Modelo LV) (Sales, 2010), que faz uso de métricas não-lineares na avaliação contínua dos alunos e auxilia na redução da sobrecarga de trabalho dos Professores/Tutores. O Modelo LV tem princípios fundamentados na interação e vai além de escalas numéricas ao envolver a dimensão qualitativa, a subjetividade e a possibilidade da multiplicidade de critérios avaliativos, o que justifica o modelo matemático vetorial com seus Vetores-Aprendizagem.

Ao fazer uso dos indicadores de aprendizagem do Modelo LV, Cacais, Sales e Caldas (2018) propuseram um sistema para avaliação de desempenho corporativo que utiliza elementos de jogos como motivação em suas métricas para quantificar e qualificar os resultados obtidos em empresas. Albuquerque (2018) apresenta uma proposta de avaliação da presença social aplicando Lógica Fuzzy em fóruns do AVA Moodle e possibilita um maior refinamento das métricas do Modelo LV. Silva (2017) propôs o Alert-LV, uma ferramenta capaz de monitorar e emitir ao aluno constantes feedbacks, aplicando um sistema de alertas, tendo por base uso de agentes. Marinho *et al.* (2016) propõem a criação do Glossário LV e o implantaram para avaliar a ferramenta Glossário do AVA Moodle. Goncalves, Cunha e Sales (2016) concebem o Fórum P&R LV após

modelarem matematicamente o Fórum Perguntas e Respostas do AVA Moodle e proporem uma aplicação em cursistas envolvendo a metodologia ativa Instrução por Pares associado ao Modelo LV.

Descreve-se a concepção e modificação da escala iconográfica de ícones do Modelo LV original para a escala contendo Graphics Interchange Format (GIF), Gifs animados, na mediação da avaliação. O uso dos Gifs no ambiente virtual ao fazer uma combinação de cores e animação ajuda a por o aluno em estado de alerta e pode transformar o AVA em um ambiente mais afetivo.

Na Seção a seguir trata-se do Modelo LV e sua concepção, na seção seguinte apresentam-se as diversas ferramentas do Moodle e suas interfaces com o Modelo LV implementado e integrado aos Gifs, na sequência é posto uma validação do Plugin LV e, por fim, as considerações finais.

2. O Modelo LV

O modelo de avaliação Learning Vectors (Modelo LV) aplica-se na efetivação de uma avaliação formativa e diagnóstica e foi concebido para facilitar o acompanhamento de alunos em Ambientes Virtuais de Aprendizagem (AVA) por parte de Professores-Tutores e, assim, reduz sua sobrecarga de trabalho. Sua metodologia de ação é fundamentada na interação. Na especificação da solução tecnológica fez-se uso de vetores e de equações trigonométricas por ser um modelo matemático que possibilita múltiplas vertentes (Sales, 2010, 2017; Sales, Barroso, & Soares, 2012). O Modelo LV de avaliação envolve aluno e Professor-Tutor em constante interação, dialogia, aprendizagem colaborativa e troca de feedbacks (Fig. 1).

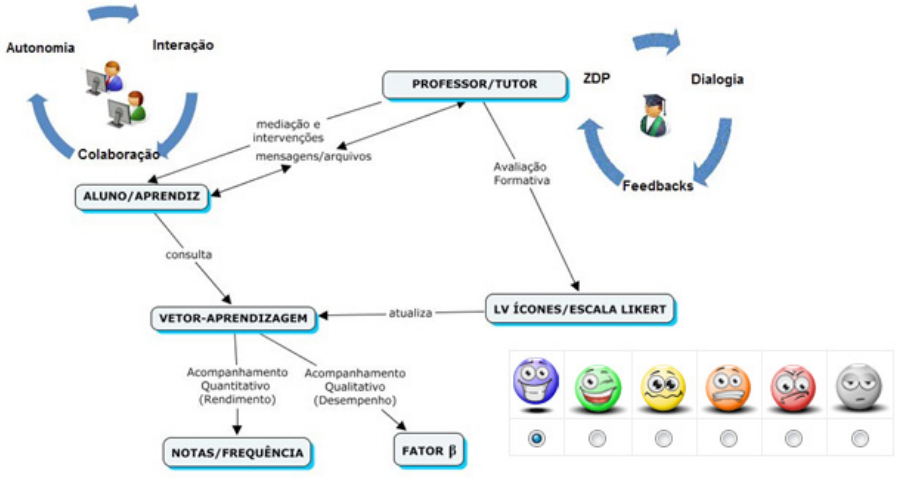


Figura 1 – Modelo Tecno-Metodológico de Avaliação Formativa LV. Fonte: Adaptado de Sales (2010)

O modelo matemático de vetores para a especificação do Modelo LV apóia-se no fato da bidimensional inerente a seus componentes e na possibilidade de associar isto à multiplicidade de critérios avaliativos. O Modelo LV fundamenta-se em teorias interacionistas sistêmicas, na avaliação como contínua formação que permite regulações das aprendizagens e tem por base os sistemas dinâmicos não-lineares, tendo apoiado-se no modelo psicomatemático Meta Learning (Losada, 1999; Losada & Heaphy, 2004; Fredrickson & Losada, 2005). Crê-se que a aprendizagem de forma presencial, ou com suporte virtual, inclui uma dinâmica muitas vezes imprevisível e cheio de incertezas, assemelha-se a algo quântico.

Com o desenvolvimento do plugin a partir do Modelo LV, devidamente implementado no AVA Moodle, pode-se avaliar: o Fórum de discussão, salas de Chat, Tarefas e os Wikis, também foi possível importar notas de Quizzes e gerenciar a frequência do aluno e oferecer uma avaliação contínua e formativa. Para tanto, nessa comunicação mediada é utilizado os Vetores-Aprendizagem (Fig. 2), onde os alunos visualizam suas notas.

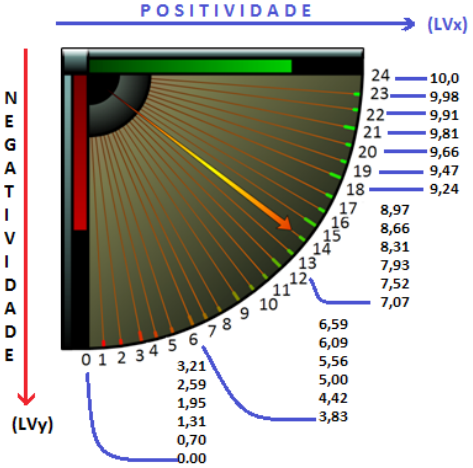


Figura 2 – Representação do Vetor-Aprendizagem com Passos correlacionados aos avanços angulares do vetor. Fonte: Adaptado de Sales (2010)

A projeção horizontal do vetor expressa a nota da atividade (LV_x) e a positividade de desempenho do aluno, já a projeção vertical relaciona-se à negatividade do desempenho (LV_y). Os 25 possíveis valores de ângulos e notas associadas, representadas no Vetor-Aprendizagem, e os avanços angulares ou passos (múltiplos de 3,75°), são apresentadas na Tabela 1.

Uma escala de menções qualitativas: Muito Bom, Bom, Regular, Fraco, Não Satisfatório e Neutro, devidamente categorizadas, associada a uma escala iconográfica, LV Ícones atualmente transformados em LV Gifs, é utilizada pelo Professor/Tutor para avaliar as mensagens ou arquivos enviados pelos alunos (Fig. 3) durante as atividades no AVA.

Passos/ Nota	Ângulo/		Passos/ Nota	Ângulo/		Passos/ Nota	Ângulo/		Passos/ Nota	Ângulo/		Passos/ Nota	Ângulo/	
0	0,00°	**0,00**	**5**	18,75°	**3,21**	**10**	37,50°	**6,09**	**15**	56,25°	**8,31**	**20**	75,00°	**9,66**
1	3,75°	**0,70**	**6**	22,50°	**3,83**	**11**	41,25°	**6,59**	**16**	60,00°	**8,66**	**21**	78,75°	**9,81**
2	7,50°	**1,31**	**7**	26,25°	**4,42**	**12**	45,00°	**7,07**	**17**	63,75°	**8,97**	**22**	82,50°	**9,91**
3	11,25°	**1,95**	**8**	30,00°	**5,00**	**13**	48,75°	**7,52**	**18**	67,50°	**9,24**	**23**	86,25°	**9,98**
4	15,00°	**2,59**	**9**	33,75°	**5,56**	**14**	52,50°	**7,93**	**19**	71,25°	**9,47**	**24**	90,00°	**10,0**

Fonte: Adaptado de Sales (2010)

Tabela 1 – Possíveis Valores da Variação Angular e Correspondentes Notas

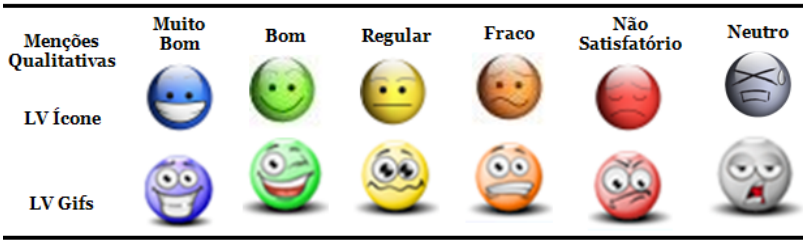


Figura 3 – Escala de Menções Qualitativas: de LV Ícones a LV Gifs

O Modelo LV expressa o nível de desempenho do aluno no curso por meio do Fator β (Sales, Barroso, & Soares, 2012), tal fator é um indicador qualitativo não-linear que envolve as subjetividades a partir dos itens de apreciação de uma escala de menções. O Fator β pode indicar qualitativamente o nível de desempenho do aluno e é expresso pela dimensão bipolar Positividade (P)/Negatividade(N), matematicamente representada pela taxa P/N (Fig. 4).

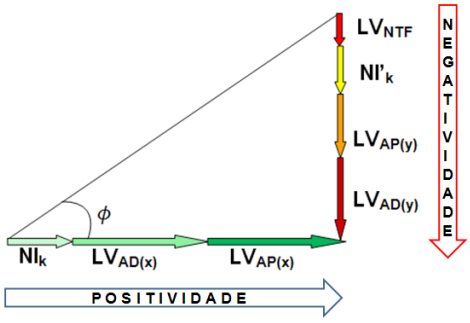


Figura 4 – Fator β: Positividade X Negatividade. Fonte: Sales (2010)

Toma-se como Positividade (P), o somatório das projeções horizontais dos Vetores-Aprendizagem, acrescido do somatório do Número de Interações Positivas (NI_k),

devidamente ponderadas e categorizadas com menções qualitativas “Muito Bom- LV Gif Azul” (peso 3), “Bom - LV Gif Verde” (peso 2) ou “Regular - LV Gif Amarelo” (peso 1). E, como Negatividade (N), toma-se o somatório das projeções verticais dos vetores, acrescido do módulo do LV_{NTF} (LV Número Total de Faltas), mais o somatório do Número de Interações Negativas (NI'_k), devidamente ponderadas e categorizadas com menções qualitativas “Fraco - LV Gif Laranja” (peso 1) ou “Não Satisfatório - LV Gif Vermelho” (peso 2).

3. Ferramentas LV Síncronas e Assíncronas e suas Interfaces no AVA Moodle

Apresentam-se nesta seção as diversas ferramentas onde foram implementadas o Modelo LV com suas interfaces e funcionalidades.

3.1. A Ferramenta Fórum LV

Em Fórum LV as avaliações são salvas automaticamente ao selecionar o LV Gif, portanto, sem a necessidade de rolar a página até seu final para salvar alterações (Fig. 5).

Figura 5 – Interface do Fórum LV

A cada mensagem avaliada, o Vetor-Aprendizagem avança de acordo com a configuração de mensagens mínimas atribuídas, pode-se configurar o Fórum LV para que sejam avaliadas 1, 2, 6 ou 10 mensagens para resultar na nota 10 fica a critério dos objetivos traçados para a discussão. Para o caso de duas mensagens mínimas a serem avaliadas, e sendo estas postadas pelo aluno e avaliadas pelo Professor/Tutor ambas com LV Gif Azul resultará em nota 10 (Tabela 2).

Fórum LV									
1ª Mensagem			2ª Mensagem					3ª Mensagem em diante	
LV Gif	Nota							+2 passos	- 2 passos
	7,07		10,0	9,81	9,24	8,31	7,07		
	5,56	+	9,81	9,24	8,31	7,07	5,56	+	
	3,83		9,24	8,31	7,07	5,56	3,83		
	1,95		8,31	7,07	5,56	3,83	1,95		
	0,0 --		7,07	5,56	3,83	1,95	0,00		

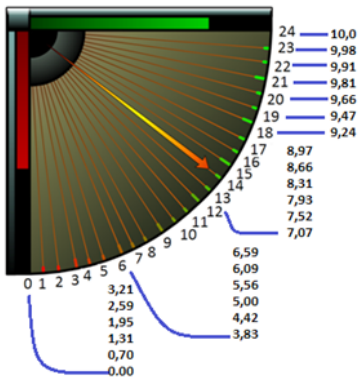


Tabela 2 – Ordem de Mensagens e Notas do Vetor-Aprendizagem em Fórum LV configurado para duas mensagens mínimas a serem avaliadas

Por exemplo, caso se configure o Fórum LV para que o Professor/Tutor avalie no mínimo duas mensagens do cursista para totalizar a nota 10, e, tendo sido o mesmo avaliado com LV Gif Verde na primeira mensagem e LV Gif amarelo na segunda mensagem, o Vetor-Aprendizagem resultará em nota 8,31 o que não impede do aluno enviar mais mensagens para serem avaliadas, neste exemplo, suponha que uma terceira mensagem do aluno tenha sido avaliada com LV Gif azul, verde ou amarelo, seu Vetor-Aprendizagem avançará dois passos, indo do Passo 17 ao 19 e finalizará a nota em 8,97.

3.2. A Ferramenta Tarefa LV

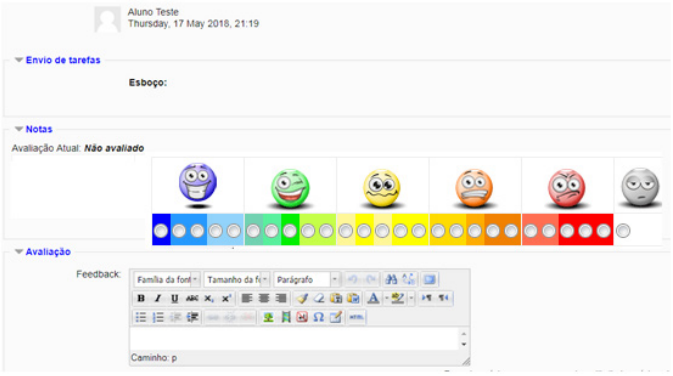


Figura 6 – Escala ampliada de LV Gifs para as 25 Notas em Tarefa LV

Para Tarefas LV a Escala de Gifs foi estendida permitindo uma avaliação de múltiplos critérios (Fig. 6).

A relação LV Gif versus Notas em Tarefa LV (Tabela 3) relaciona a cada cor principal do LV Gif outras cinco totalidades, estabelecendo assim, uma matriz bidimensional de avaliação com dois critérios complementares: "Domínio de Conteúdo" e "Pontualidade/ Apresentação".

Tarefa LV						
Domínio de Conteúdo		Pontualidade/Apresentação				
		Muito-Bom	Bom	Regular	Fraco	Não Satisfatório
	Muito-Bom	10,0	9,98	9,91	9,81	9,66
	Bom	9,47	9,24	8,97	8,66	8,31
	Regular	7,93	7,52	7,07	6,59	6,09
	Fraco	5,56	5,00	4,42	3,83	3,21
	Não Satisfatório	2,59	1,95	1,31	0,70	0,00
	Neutro	Nenhuma nota será atribuída				

Tabela 3 – Critérios bidimensionais de Avaliação em Tarefa LV

3.3. A Ferramenta Wiki LV

Figura 7 – Avaliação do trabalho final produzido no grupo sem a necessidade de se avaliar o histórico individual na ferramenta Wiki LV

Na ferramenta Wiki LV pode-se avaliar somente o trabalho final produzido no grupo, não se verificando as contribuições individuais no histórico da ferramenta Wiki (Fig. 7).

Observe-se a escala completa de notas uma vez que avalia-se apenas o produto final construído pelos alunos. Por sua vez o Wiki pode ser configurado quanto ao número mínimo de contribuições individuais (0, 1, 2, 3, 4 ou 5) a serem avaliadas no histórico desta ferramenta (Fig. 8).

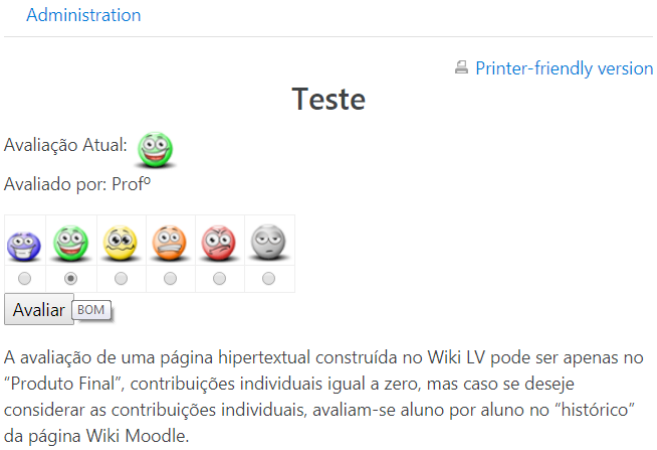


Figura 8 – Avaliação do trabalho final produzido no grupo com a necessidade de se avaliar o histórico individual na ferramenta Wiki LV

Observe-se a escala reduzida de notas no produto final construído, uma vez que essa nota será acrescida das notas nas contribuições individuais, ocasião em que se avaliam aluno por aluno no "histórico" da página Wiki Moodle (Fig. 9).

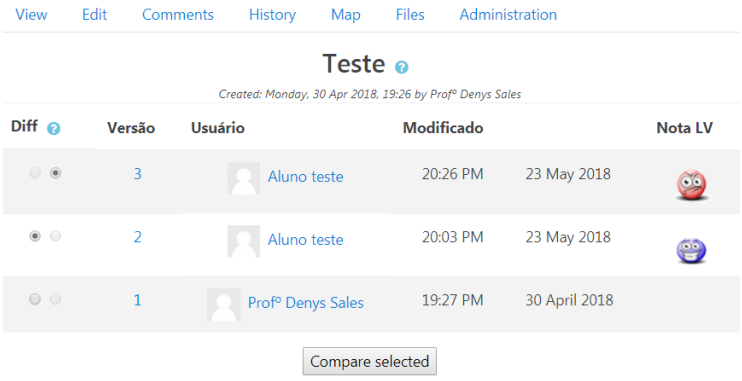


Figura 9 – Avaliação em histórico do Wiki LV

Para uma configuração de quatro contribuições individuais apresentam-se os valores de notas (Tabela 4).

Wiki LV configurado para 4 contribuições individuais				
Produto Final (PF)			Contribuições Individuais (CI)	
LV Gif	Nota		LV Gif	Passos (Avanço/ Recuo)
	5,0			+ 4
	3,8	+		+ 3
	2,6			+ 2
	1,3			+ 1
	0			-1
	-			-

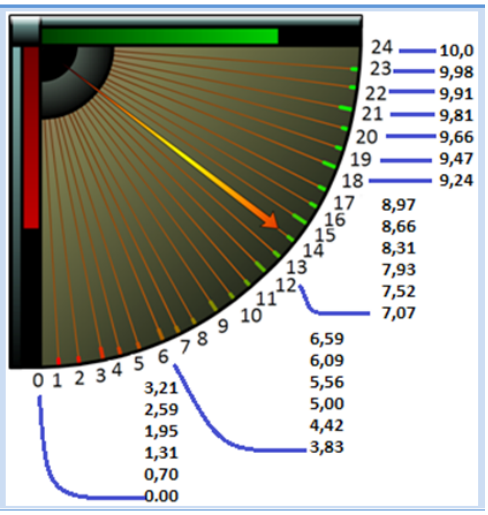


Tabela 4 – Notas no Wiki LV configurado para um mínimo de 4 contribuições individuais

Suponha que se atribua LV Gif Verde ao Produto Final (nota 3,83, passo N° 6) e que se avaliem as quatro contribuições individuais como: LV Gif Verde (avanço de 3 passos), LV Gif Amarelo (avanço de 2 passos), LV Gif Azul (avanço de 4 passos) e LV Gif Laranja (avanço de 1 passos), sua nota resulta a partir de 3,8 com um avanço total de 3+2+4+1 = 10 passos em 8,66 (passo N° 16).

3.4. A Ferramenta Chat LV

Chat LV						
1ª Mensagem			2ª a 6ª Mensagem		7ª Mensagem em diante	
LV Gif	Nota		LV Gif	Passos Avanço	LV Gif	Passos (Avanço/ Recuo)
	2,59			+ 4		+ 1
	1,95	+		+ 3	+	+ 1
	1,31			+ 2		+ 1
	0,70			+ 1		+ 1
	0,00			0		- 1
	-			0		0

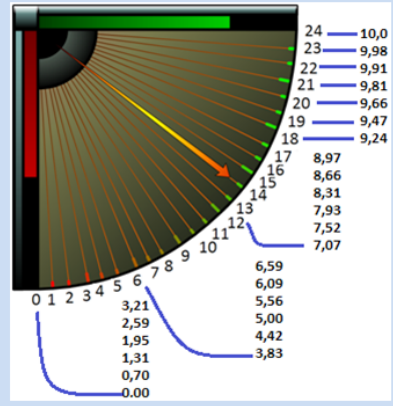


Tabela 5 – Número de mensagens avaliadas e Avanços angulares do Vetor-Aprendizagem na ferramenta Chat LV configurado para seis mensagens mínimas a serem avaliadas

No Chat LV para o Professor/Tutor avaliar online exige-se uma dinâmica à proporção que o diálogo ocorre de forma síncrona, mas a avaliação pode ser feita também após o encerramento da sessão de chat uma vez que a ferramenta grava toda a conversa. Pode-se configurar esta ferramenta para um mínimo de 3 ou 6 mensagens mínimas postadas por parte do aluno/cursista a serem devidamente avaliadas pelo Professor/Tutor (Tabela 5).

Fazendo uso da Tabela 6 verifica-se que, se o aluno tirar LV Gif Azul em sua primeira mensagem, nota 2,59 ou passo nº 4, ele precisará que seu Vetor-Aprendizagem avance mais 20 passos, ou o correspondente a avaliação por parte do Professor/Tutor de cinco LV Gifs Azuis, ou qualquer correspondência de outros LV Gifs que resultem em 20 avanços angulares unitários de tal forma a atingir o 24º passo ou nota 10.

3.5. O Plugin de Notas LV e sua Validação

Todas estas funcionalidades foram postas em um Plugin, denominado Notas LV, desenvolvido em PHP pode ser instalado no LMS Moodle, este *plugin* apresenta um espaço para se configurar e lançar as notas das atividades presenciais e a distância, permite também a importação das notas dos *quizzes* e apresenta um mapa de notas, denominado 'Avaliação Processual', contendo todas as notas, ausências em alguma atividade e a classificação dos alunos qualitativamente pelo Fator β, tal mapa de notas pode apresentar-se com todos os alunos ou de forma individualizada. Por fim, ao se clicar em 'Avaliação Processual', o sistema retorna as notas de cada aluno e sua situação final no curso. Para garantir a privacidade, cada aluno visualiza apenas suas notas, já o Professor/Tutor pode visualizar as notas de todos os alunos e ainda pode exportar o mapa de notas em pdf (FIGURA 10).

Nome: Didática e Docência (MEF_2018.1)
Descrição: Semestre 2018.1
β Médio: 3.33

Nome	NF	NC	NW	NQ	ND	NP	Média	NTF (hA / %)	Fator β	LV Ícone	PF	MF	Situação
Aluno 1	0.58	0	0	0	0.58	0	0.6	43 / 53.75%	0.09		-	0.6	RF
Aluno 2	8.01	0	0	0	8.01	0	8	0 / 0%	2.55		-	8	AM
Aluno 3	7.17	0	0	0	7.17	0	7.2	7 / 8.75%	1.85		-	7.2	AM
Aluno 4	8.27	0	0	0	8.27	0	8.3	0 / 0%	3.43		-	8.3	AM
Aluno 5	7.52	0	0	0	7.52	0	7.5	7 / 8.75%	2.7		-	7.5	AM
Aluno 6	1.39	0	0	0	1.39	0	1.4	43 / 53.75%	0.19		-	1.4	RF
Aluno 7	8.57	0	0	0	8.57	0	8.6	0 / 0%	5.65		-	8.6	AM
Aluno 8	8.43	0	0	0	8.43	0	8.4	0 / 0%	5.75		-	8.4	AM
Aluno 9	7.37	0	0	0	7.37	0	7.4	0 / 0%	2.01		-	7.4	AM
Aluno 10	8.28	0	0	0	8.28	0	8.3	0 / 0%	3.43		-	8.3	AM
Aluno 11	8.14	0	0	0	8.14	0	8.1	0 / 0%	3.45		-	8.1	AM
Aluno 12	8.81	0	0	0	8.81	0	8.8	0 / 0%	8.9		-	8.8	AM

Figura 10 – Mapa de notas originária do Plugin LV

O mapa de Notas LV diferencia do Notas Moodle pela inclusão da coluna NTF, que apresenta o número total de faltas do cursista em alguma das atividades presenciais e/ou a distância proposta, a coluna representativa do desempenho qualitativo do cursista, Fator β, bem como a coluna da situação final do aluno: Aprovado (AM), Reprovado (R),

Reprovado por falta (RF). Aplicado numa turma composta por 12 alunos do curso de Licenciatura de Física na disciplina de Metodologia do Ensino do Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE), pode-se verificar a funcionalidade da ferramenta Fórum, o curso durante o semestre de 2018.1 contou com um total de 10 fóruns de discussão versando sobre vários temas inerentes a disciplina.

Percebe-se nos resultados finais expressos na Figura 10, que os Alunos 7, 8 e 12, cujas médias foram: 8,6; 8,4 e 8,8 e que apresentam o LV Gif Azul são os alunos de maior Fator β, respectivamente são: 5,65; 5,75 e 8,9, bem acima do Fator β Médio da turma que é 3,33. Percebe-se também que o Aluno 8 com média inferior ao Aluno 7 tem maior Fator β, isso revela a não linearidade desta métrica e justifica-se pelo fato desse aluno ter tido melhor desenvoltura nas aulas, qualitativamente ele é melhor que o Aluno 7. Os alunos 1 e 6 tem situação final reprovados por faltas (RF), o fato do plugin capturar as ausências dos alunos nas atividades, coluna NTF na Figura 11, facilita e reduz a sobrecarga de trabalho do Professor/Tutor.

4. Conclusões

Com a concepção e desenvolvimento desta escala de avaliação que utiliza o poder das cores e Gifs animados a serem utilizados pelos Professores/Tutores em suas intervenções pedagógicas nos AVA, espera-se tornar a sala de aula virtual em espaço mais afetivo e motivador com uma avaliação mais próxima de uma avaliação formativa e reguladora de aprendizagens, assim como, que oportunize a qualquer instante pela consulta da métrica qualitativa Fator β a devida avaliação diagnóstica.

Com a aplicação na turma de alunos pode-se perceber um maior engajamento revelado por suas constantes participações nos fóruns e pela expectativa de qual cor de Gif o professor iria categorizar suas mensagens. O Plugin foi desenvolvido para ser compartilhado livremente no endereço http://bit.ly/2ANXA1b e como trabalhos futuros pretende-se fazer a inclusão dos alunos no processo de avaliação, habilitando a avaliação por pares com a inserção do Modelo LV e seus GIFs, assim, por em prática um design participativo e dar ao aluno um papel ativo e gestor de sua aprendizagem. Sugere-se também uma pesquisa sobre os efeitos dos GIFs na aprendizagem.

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) - Brasil - Código de Financiamento 001.

Referências

Albuquerque, M. C. N. de (2018). *Fórum LV Fuzzificado: Avaliando a presencialidade em Ambientes Virtuais*. Dissertação (Mestrado em Computação). Instituto Federal do Ceará - Campus Fortaleza.

Cacais, M. G., Sales, G. L., & Caldas, W. L. (2018). A Gamified System for Task Monitoring and Performance Evaluation Using the Learning Vectors (LV) Model. *Journal of Internet Technology and Secured Transactions*, v. 7, p. 605–614.

Fredrickson, B. L., & Losada, M. (2005). Positive Affect and the Complex Dynamics of Human Flourishing. *American Psychologist*, v. 60, nº 7, p. 678 – 686.

Goncalves, A. J., Cunha, J. L. L., & Sales, G. L. (2016). Concepção do Fórum P&R LV: avaliação formativa da aprendizagem. *RENOTE. Revista Novas Tecnologias na Educação*, v. 14, p. 1-10.

Leite, E. A. M., Machado A., Sales G. L., Zednik H., & Maia S. M. (2015). Distance Education Evaluation: An Analysis of the β Factor from LV Model Subjectivity. In: Rocha A., Correia A., Costanzo S., Reis L. (eds) *New Contributions in Information Systems and Technologies. Advances in Intelligent Systems and Computing,* vol 353. Springer, Cham. Disponível em: https://goo.gl/jiAxdK Acessado em: 21 Jun 2018.

Losada, M. (1999) The complex dynamics of high performance teams. *Mathematical and Computer Modelling*, v. 30, n. 9, p. 179-192.

Losada, M., & Heaphy, E. (2004) The role of positivity and connectivity in the performance of business teams: A nonlinear dynamics model. *American Behavioral Scientist*, vol. 47, nº 6, 740-765.

Marinho, M. J. F., Cunha, J. L. L., Goncalves, A. J., & Sales, G. L. (2016). Glossário Hipertextual Colaborativo LV. *RENOTE. Revista Novas Tecnologias na Educação*, v. 14, p. 1-11.

Santos, E. (2010). Educação online para além da EAD: um fenômeno da cibercultura. In. SILVA, Marco; PESCE, Lucila; ZUIN, Antônio (orgs). *Educação online: cenário, formação e questões didático-metodológicas*. Rio de Janeiro: wak Ed.

Sales, G. L. (2010). *Learning Vectors (LV): um modelo de avaliação da aprendizagem em EaD online aplicando métricas não-lineares*. Tese Doutorado. Departamento de Engenharia de Teleinformática. Universidade Federal do Ceará. 239f. Disponível em: < https://goo.gl/2ULNJU > Acessado em: 08 Maio 2018.

Sales, G. L.(2017). *Learning Vectors: Avaliação Formativa em EaD Online*. Recife: Imprima Soluções Gráficas.

Sales, G. L., Barroso, G. C., & Soares, J. M. (2012). Learning Vectors (LV): A Procedural Assessment Model with Non-Linear Measurement of Learning in Distance Education Online. *Brazilian Journal of Computers in Education*, p. 60. Available at: <http://www.br-ie.org/pub/index.php/rbie/article/view/1368>. Date accessed: 20 june 2018.

Sales, G. L., Albuquerque, M. C. N., Leite, E. A. M., & Paz, R. P. (2018). Dez Anos do Modelo Learning Vectors na Avaliação Formativa no AVA Moodle: de Emoticons a Gifs Animados In: *Tecnologias da Educação: passado, presente, futuro. Coleção História da Educação*.1 ed.Fortaleza: UFC, v.1, p. 63-78.

Silva, F. A. G. da (2018). *Alert-LV: Um sistema de monitoramento e apoio à tutoria para suporte à avaliação em ambientes virtuais de aprendizagem*. Dissertação (Mestrado em Ciência da Computação). Instituto Federal do Ceará - Campus Fortaleza.

Valente, J. A. (2002). A espiral da aprendizagem e as tecnologias da informação e comunicação: repensando conceitos. In: JOLY, M. C. (Ed.). *Tecnologia no ensino: implicações para a aprendizagem*. São Paulo: Casa do Psicólogo, p.15-37.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 08/10/2018
Aceitação/Acceptance: 19/12/2018

Alert-LV: Um Sistema de Monitoramento e Apoio à tutoria para suporte à avaliação Formativa em AVA

Francisco Adaias Gomes da Silva, Gilvandenys Leite Sales

adaiasgomes314@gmail.com, denyssales@ifce.edu.br

[1] Instituto Federal do Ceará, Av 13 de Maio, 2081, 60 040 - 215, Fortaleza/Ceará, Brasil

Pages: 77–89

Resumo: Acompanhar as acões do aluno em cursos EaD de forma contínua e utilizar-se de mecanismos automatizados para lidar com esse processo torna-se cada vez mais necessário. Para tanto, apresenta-se o sistema de emissão de alertas, Alert-LV, e os resultados obtidos ao validá-lo. Fundamentado no Modelo *Learning Vectors* e recomendações do *Learning Analytics,* a pesquisa ocorreu em quatro etapas, em que se obteve o desempenho dos alunos antes e depois da utilização do sistema Alert-LV. Foram extraídos resultados satisfatórios das taxas de evasão e elevacão na participacão e desempenho dos alunos no período das análises. Constatou-se que os frequentes alertas de desempenho emitidos ao aluno durante o curso, bem como as notificações apontando o andamento de suas atividades, são fatores imprescindíveis para a motivação, melhoria e apoio da autogestão da aprendizagem, o que consequentemente pode evitar o quadro negativo da evasão em EaD.

Palavras-chave: Sistema de emissao de alertas; Avaliação formativa; Modelo Learning Vectors; AVA Moodle.

Learning Vectors Model in Formative Assessment LMS Moodle: from emoticons to animated GIFs

Abstract: Monitor the actions of the student in e-learning courses continuously and be used for automated mechanisms to deal with this process becomes increasingly necessary. For this, the alert issuing system, Alert-LV, is presented and the results obtained when validating it. Based on the Learning Vectors Model and recommendations from the Learning Analytics, the research was conducted in four stages, in which student performance was obtained before and after the use of the Alert-LV system. Satisfactory results were obtained of the dropout rates and increase in student participation and performance in the analysis period. Frequent performance alerts issued to the student during the course, as well as notifications indicating the progress of their activities, are essential factors for the motivation, improvement and support of self-management of learning, which can therefore avoid the negative evasion in e-learning.

Keywords*:* Alert issuing system; Formative evaluation; Learning Vectors Model; LMS Moodle.

1. Introdução

Em Educação a Distância (EaD), um ponto importante a se destacar, seria estabelecimento do "estar junto virtual", definido por Valente (1999) como o sentimento, a sensação de proximidade e pertencimento ao grupo de aprendizagem que passa a existir no meio virtual a partir das interações realizadas entre os indivíduos atuantes naquele meio. Há uma real possibilidade de o aprendiz obter conhecimento e motivação, ao sentir proximidade com seu tutor e com os colegas na realização das atividades no ambiente virtual de aprendizagem, independentemente da distância que os separa geograficamente (Tori, 2010). "Muitas das frustrações na EaD ocorrem devido à demora do tutor em responder aos alunos nas interações propostas. O aluno que não recebe indicações, respostas e correções considera que não tem o apoio de seu tutor" (Mercado, 2007).

Feedbacks rápidos e adequados mantem o estudante motivado e consciente de seus deveres e, isso, o aproxima de seu tutor. Por outro lado, se o tutor não permitir essa aproximação, não há como motivar e guiar adequadamente o aluno que encontrará problemas como, a má administração do tempo; desatenção a prazos; silêncio e orfandade online, entre outros (Mercado, 2007). Um bom sistema de monitoramento do processo educacional do aluno em EaD, faz toda diferença no sucesso e no desenvolvimento dos cursos nessa modalidade, e para isso, utiliza-se de sistemas computacionais eficientes de acompanhamento e avaliação da aprendizagem (Moore & Kearsley, 2007).

Utilizando-se de forma adaptada a dimensão psicológica entre Positividade (P) e Negatividade (N) de desempenho (Sales, 2010, 2018, Sales, Barroso, & Soares, 2012), foi concebido um modelo de avaliação processual fundamentada em menções qualitativas e mediadas por linguagem iconográfica denominado Modelo *Learning Vectors* ou simplesmente Modelo LV (Sales, 2010). Buscar o sentimento de proximidade e de acompanhamento constante, vem a ser necessário na aprendizagem baseada por sistemas. Para tanto um sistema de alertas automatizados, que emita constantes feedbacks, e seja capaz de reduzir a distância comunicacional entre tutor e aluno nos cursos que se valem de salas de aulas virtuais, como um mecanismo alternativo, poderá facilitar o processo avaliativo e reduzir o percentual de evasão (Silva, 2018).

Esse artigo tem como objetivo apresentar a metodologia de aprendizagem baseada em alertas utilizando o sistema Alert-LV (Silva, 2018), que possui relação com o ambiente virtual Moodle e foi fundamentado na métrica qualitativa de avaliação, em associação ao Modelo LV. Este sistema fica encarregado de monitorar eventos na base de dados do Moodle e emitir os alertas caso o desempenho do aluno ou do tutor seja abaixo do esperado. Para tanto, utiliza-se do acompanhamento na tabela de notas, dos logs do sistema e dos ícones característicos do Modelo LV que denotam o desempenho dos alunos. Na Seção a seguir trata-se do Modelo LV e sistemas de alertas, na seção seguinte apresenta-se a arquitetura do Sistema Alert-LV, na sequência é posto uma validação do Alert-LV e, por fim, as considerações finais.

2. O Modelo LV e Sistemas de Alerta

O Modelo LV (Sales, 2010, 2018) é um sistema de avaliação que preza pela colaboração e interação entre os atores envolvidos em cursos na modalidade Educação a Distância: Professores/Tutores, Alunos e Ambiente Virtual de Aprendizagem (AVA). Também

apresenta métricas alinhadas a um sistema de avaliação quali-quantitativo e permite direcionar o processo de avaliação na direção formativa. O Modelo LV utiliza-se de ícones para atribuir valores, com os quais se propõe transmitir a ideia de afetividade e proximidade (Sales, 2010, 2018, Sales, Barroso, & Soares, 2012). Subjacente a esses atributos está o modelo matemático de mensuração, que associado a esses ícones, devem ficar bem explicados no contrato didático estabelecido no início do curso, assim ficam mais compreensíves as regras da avaliação. Associado a esse modelo tem-se os Vetores-Aprendizagem (Fig. 1), representado por um ente matemático denominado vetor que possui módulo fixo de 10 unidades e direção variável. Esse vetor se move em sentido anti-horário, cujos ângulos de avanço relacionam-se a uma nota atribuida pelo Professor/Tutor.

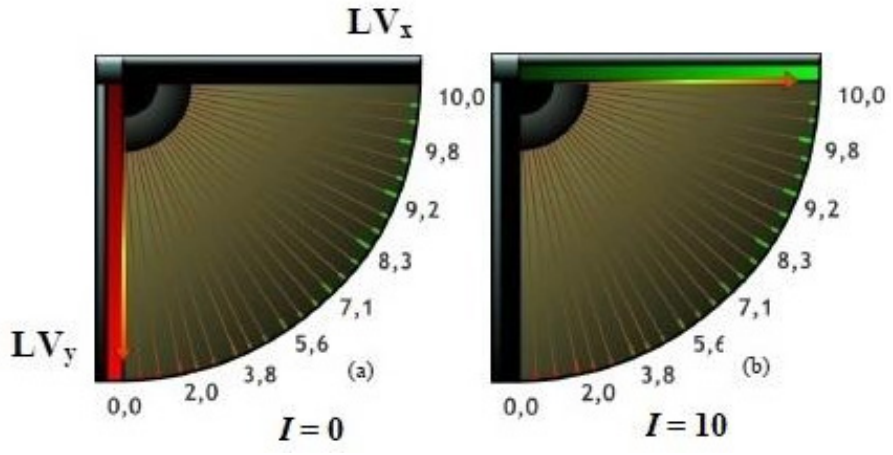


Figura 1 – Vetor-Aprendizagem com escalas de notas do Modelo LV
Fonte: Adaptado de Sales (2010)

As mensagens e/ou arquivos postados pelos alunos durante um curso passam a ser avaliadas pelo Professor/Tutor por meio da escala icônica, denominada LV Ícones (Fig. 2).

Muito Bom	**Bom**	**Regular**	**Fraco**	**Não Satisfatório**	**Neutro**
Reflexões aprofundadas	Boas reflexões	Razoáveis reflexões	Reflexões vazias de conteúdo	Aluno que assume posição passiva ou ainda que envia arquivos anexados sem originalidade	Mensagens/Arquivos anexados que não agregam notas

Fonte: Adaptado de Sales, Barroso e Soares (2012)

Figura 2 – Escala LV Ícones de categorização das mensagens

Esses ícones são atribuidos pelo Professor/Tutor às ações dos alunos, sejam nas ferramentas de interação, ou de portfólio de arquivos recebidos, e correlacionam-se a valores numéricos de notas. Uma outra métrica associada ao Modelo LV é o Fator β, que é

um indicador qualitativo não-linear para suporte no acompanhamento da aprendizagem do aluno, adaptado a partir da dimensão psicológica Positividade/Negatividade (Taxa P/N) do modelo de avaliação de desempenho denominado *Meta Learning* (Modelo ML), proposto tanto para avaliação de equipes quanto para indivíduos (Losada & Heaphy, 2004). Esse fator expressa o desempenho do aluno em: Muito Alto, Alto, Médio, Baixo e Muito Baixo, com valores médios classificados segundo Tabela 1.

Escala	**Fator β**	**LV Ícone**
Muito Alto	$\beta \geq 3{,}78$	Azul
Alto	$2{,}62 \leq \beta < 3{,}78$	Verde
Médio	$0.90 \leq \beta < 2{,}62$	Amarelo
Baixo	$0{,}30 \leq \beta < 0{,}90$	Laranja
Muito Baixo	$0 \leq \beta < 0{,}30$	Vermelho

Fonte: Adaptado de Sales (2010)

Tabela 1 – Escala qualitativa de especificação do Fator β

A relação do Fator β com os LV Ícones (Tabela 1) possibilita uma avaliação diagnóstica e serve de alerta para o desempenho do cursista, desde que seja constantemente monitorado pelo Professor/Tutor, portanto, o desenvolvimento de sistema de alertas automatizados no acompanhamento e avaliação formativa que possibilite uma maior comunicação com os alunos faz-se necessário. A utilização de técnicas de interação e participação e uso dos feedbacks promove integração, permite orientação adequada dos processos e atividades em estudo, e destaca o papel de cada sujeito na aprendizagem (Moore & Kearsley, 2007).

Na pesquisa de Kampff (2009) foram utilizadas as ferramentas RapidMiner, WizRule e Weka, para verificação e organização das ações e relações dos alunos no ambiente de aprendizagem, e na estruturação das informações para a tomada de decisões com a Mineração dos Dados. Utilizando as ferramentas, a autora propôs o desenvolvimento de uma arquitetura de sistemas de alertas pré-definidos em AVA e outras notificações geradas a partir dos dados coletados na técnica de mineração, sua intenção foi mapear comportamentos que poderiam levar a evasão ou reprovação dos alunos em EaD (Kampff, 2009). Em que pese os seus grandes benefícios, uma desvantagem identificada e que dificulta a inserção da ferramenta no cotidiano do professor seria a utilização da Mineração de Dados por profissionais que não detêm o domínio dessa técnica.

Silveira (2001) utilizando-se de uma plataforma denominada Eletrotutor, desenvolvida a partir do Framework Jade, elaborou três modelos base para a proposta do sistema: modelo organizacional, modelo de agente e um modelo de cooperação. O referido sistema pertence a uma classe de sistemas de Ambientes Inteligentes Distribuídos de Aprendizagem e aborda assuntos de Física, precisamente Eletricidade abrangendo assuntos com Eletrodinâmica e a Lei de Ohm.

O trabalho de Musa (2001) foi voltado para a realização das avaliações nas atividades a distância. O foco da pesquisa seria contribuir com uma nova metodologia de avaliação. Propõe a criação de um Sistema de Ensino na Web, denominado Sistema Tapejara.

O sistema utilizava de alertas que permitiam especificar situações especiais que geravam avisos essenciais para realizar as tomadas de decisões. Para Musa (2001), a função desses sistemas, seria alertar possíveis ocorrências de problemas nas atividades dos alunos no curso, e esses avisos seriam emitidos aos tutores dos problemas ocorridos, e tomar as medidas necessárias para resolvê-los.

Almeida (2013) aponta as diversas ferramentas e os múltiplos recursos existentes nos AVA como datas, horário e frequência de acessos, quantidade de atividades submetidas e as necessárias para haver acompanhamento, resultados de testes on-line, participação nos fóruns de discussão, a quantidade de mensagens interacionistas realizadas. Todas as atividades são registradas e podem ser acessadas posteriormente a título de verificação e auditoria. Porém a problemática identificada é a falta de organização dessas informações, pois isto dificulta a tomada de ações dos personagens da educação. Propõe ainda definir quais os indicadores que podem contribuir para a coleta de informações relevantes para monitorar os usuários. Fazendo assim uso do ambiente de aprendizagem Tidia-AE, propondo uma ferramenta para disponibilizar em tela essas informações coletadas nesse AVA (Almeida, 2013).

3. Arquitetura do Sistema Alert-LV

A arquitetura proposta inclui o desenvolvimento de uma interface em Java utilizando o Framework Jade (JADE, 2016) para a realização da geração dos alertas que serão utilizados pelo Professor/Tutor, e por meio dela são configurados (Fig. 3).

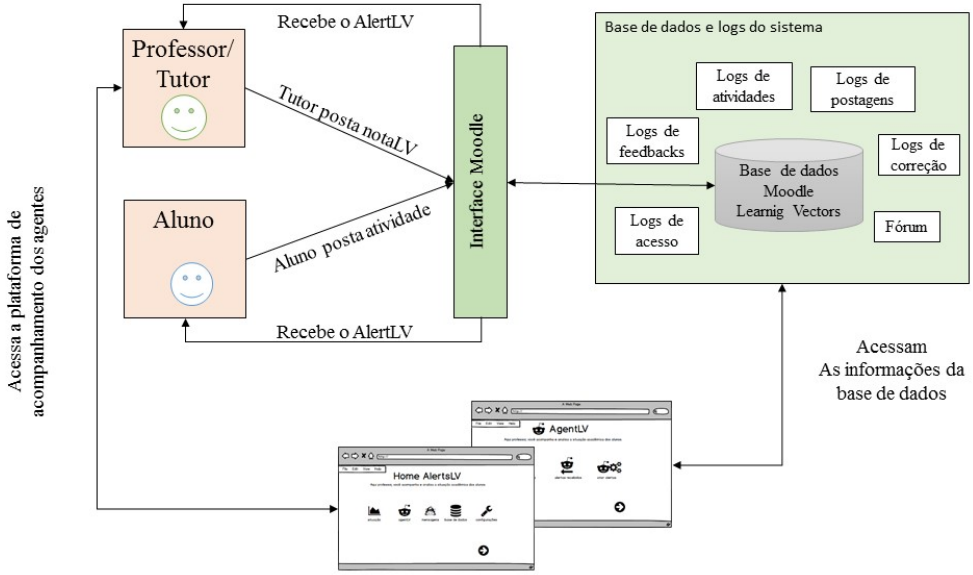


Figura 3 – Arquitetura do sistema Alert-LV. Fonte: Adaptado de Silva (2018)

Os sujeitos interagem com a interface do Moodle, que posteriormente realiza comunicação com o banco de dados afim de obter os dados de logs e as notas inseridas nas atividades

realizadas pelos alunos. Ao acessar o sistema de alertas o Professor/Tutor gerenciará os alertas que posteriormente serão enviados aos alunos a partir das informações da base de dados do Moodle relacionadas aos eventos de atividades realizadas no ambiente, assim emite os alertas para os alunos caso estes sejam necessários.

Os alertas são configurados e emitidos a partir das interações dos usuários. A base de dados utilizada será o banco de dados adotado pelo ambiente Moodle. Nessa base de dados ficam armazenadas as informações de logs no sistema, as últimas alterações realizadas, e quais as postagens, tanto do professor, como do aluno. Além de registrar as notas e comentários emitidos durante o curso. A interface de recebimento dos alertas emitirá e registrará as mensagens que serão geradas pelas escolhas do Professor/Tutor e pelos dados existentes na base de dados Moodle. O monitoramento e configuração dos dados que são coletados dos bancos de dados do Moodle, é realizado pelo Alert-LV. O sistema realiza a coleta das informações nos registros de logs do Moodle, bem como, informações referentes aos registros de atividades e notas. O sistema identifica padrões de acessos, datas de entrada no sistema, realização de tarefas, a quantidade de envio de mensagens, entre outros dados e informações importantes para o tutor.

Também identifica o desempenho dos estudantes, isso quando o tutor aplica a análise de alguma ação, seja ela participação em chats, fóruns, Wikis, dentre outras atividades solicitadas. Cabe ao sistema analisar se o Fator β do aluno está de acordo com o período de atuação dele e de toda a turma, e assim emitir alertas individuais e/ou coletivos. O usuário aluno tem como papel principal realizar o acesso a plataforma Moodle conforme disponibilidade das Tarefas LV, participar do Fórum LV, do Chat LV e da Wiki LV. Posteriormente, cabe ao aluno averiguar e acompanhar suas notas e receber as mensagens de feedbacks a ele direcionado.

Com base nisso, o usuário Professor/Tutor, realiza a inserção das atividades no ambiente, acompanha as postagens dos alunos e emite uma nota, por meio do sistema de avaliação *Learning Vectors* e/ou recebe o retorno do sistema caso o aluno não tenha recebido nota, ou tenha acontecido algum outro problema. Em resposta, o professor recebe e acompanha o desempenho da turma como um todo analisando o coeficiente de aprendizagem ou Fator β (Leite *et al.*, 2015).

A Base de dados utilizada será o banco de dados adotado pelo ambiente Moodle. Nessa base de dados ficam armazenadas as informações de logs no sistema, as últimas alterações realizadas, e quais as postagens tanto do professor como do aluno, além de registrar as notas e comentários emitidos durante o curso. O Alert-LV fica responsável por coletar informações na base de dados do Moodle incluindo logs no sistema, ausências, mensagens emitidas, verifica o período de disponibilização de cada atividade no AVA, identificando as informações do Fórum LV (AF), Tarefa LV (AT), Wiki LV (AW), e Chat LV (AC), para saber quais foram realizadas, como foi a participação dos alunos, bem como upload e download de arquivos no sistema, ainda verifica os prazos de entrega e as pendências. Assim, certifica-se para quais alunos será emitida uma resposta de retorno ou alerta.

O alerta ao analisar o desempenho do aluno e a nota a ele atribuída, ativa automaticamente o sistema de detecção de eventos e então o alerta é emitido na forma de mensagem *popups* automáticas na máquina do aluno. O trabalho aqui proposto desenvolveu mensagens de alertas classificadas em cinco categorias diferentes, cada uma com seu grau de criticidade e relevância. Assim como o Modelo LV, as mensagens foram colocadas em uma escala

com uma paleta de cores similares, com o intuito de transmitir o grau de criticidade que a informação repassada possui em positivas, negativas ou neutras. Abaixo representam-se os tipos de mensagens e quais as cores que representam essas mensagens (Tabela 2) (Silva, 2018).

Alertas	Mensagem
Mensagens de cor vermelha - notas em estado crítico, ou o desempenho abaixo do esperado	1 - Você está sendo relapso com suas atividades, você não postou a atividade e suas notas estão abaixo do esperado!
	2 - Você perdeu o prazo de submissão da atividade, que pena, entre em contato com seu tutor para justificar sua ausência!
	3 - Atenção! Faltam 3,2, ou 1 dias para encerrar a aula, o prazo para postagem está acabando corra ou entre em contato com seu tutor para justificar sua ausência!! (Clique aqui para acessar o sistema Moodle)
	4 - Atenção! Você está com o Fator β bem abaixo do esperado, seu desempenho está em nível crítico. Entre em contato com seu tutor urgentemente par saber o que está acontecendo!
Mensagem de cor laranja - presencialidade e participação	1 - Atenção a sua presencialidade! Faz dias que você não acessa o ambiente, também identificamos ausência nas postagens das Atividades LV. Procure ser mais assíduo! Entre em contato com seu tutor!
	2 - Ops! Identificamos um conteúdo que não foi acessado! Recomendamos que você o verifique posteriormente, para não se prejudicar
	3 - A atividade LV (.....) Foi a menos acessada na semana! Isso pode prejudicar o Fator β da turma!
	4 - Atenção turma ajudem os colegas com dificuldades, identificamos ..... Alunos com baixo número de acessos!
Mensagens de cor amarela - notificações gerais e dúvidas	1 - Seu Professor/Tutor entrou no ambiente. Possui dúvidas? Converse com seu tutor ou com algum colega!
	2 - Não esqueça de fazer os exercícios. São importantes para seu aprendizado!
	3 - Um aluno entrou no ambiente. Analise a situação acadêmica dele, e contate-o informando seu desempenho.
	4 - A atividade da aula desta semana já está disponível, entre no ambiente, poste comentários. Adoramos sua participação!
Mensagem de cor verde - Motivação, incentivo	1 - Estamos contentes com seu desempenho e sua presencialidade, continue assim, isso eleva a aprendizagem e demonstra empenho da turma!
	2 - Que bom! Suas notas estão cada vez melhores!
	3 - Ops! Seu Fator β melhorou, porém ainda está abaixo de 0,90. Entre em contato com seu tutor urgentemente para saber o que está acontecendo.
	4 - Que bom! Ficamos contentes por você está voltando depois de tanto tempo fora.
Mensagens de cor azul - parabenizações ou bonificações	1 - Sua nota na atividade foi muito boa! Parabéns!
	2 - Parabéns você concluiu o módulo esperamos que tenha gostado!
	3 - Ufa!!! Seu Fator β melhorou, está no nível normal, porém nada de corpo mole, vamos melhorá-lo.
	4 - Muito bem, adoramos que você tenha comentado a postagem de seu colega, isso contribui para a sua aprendizagem e a dos demais.

Tabela 2 – Categorias de mensagens de alertas

3.1. Design do protótipo do Alert-LV

O sistema Alert-LV foi criado com o intuito de promover presencialidade e aproximar ainda mais Aluno e Tutor (Silva, 2018). A proposta engloba o desenvolvimento de um sistema JAVA que aguce a interação entre professor/aluno/AVA. Na tela inicial do sistema de Alertas Alert-LV (Fig. 4), a tela Home, o Professor/Tutor tem disponível os ícones direcionais: Situação, AgentLV, Mensagens, Base de Dados, Configurações, cada ícone redireciona para a próxima tela.

Figura 4 – Tela Inicial Alert-LV. Fonte: Adaptado de Silva (2018)

O ícone 'situação' traz os dados referentes a situação do aluno e isso inclui os prazos de entrega das atividades. Na funcionalidade 'inserir alunos' o Professor/Tutor pode inserir alunos no sistema, bem com escolher quais alunos receberão alertas naquele período. Já o 'excluir aluno' permite ao professor remover o aluno, que foi antes adicionado no sistema, e que precisa ser removido por motivos adversos. O professor ao acessar a tela de edição de AlertsLV obtém como recurso um calendário ao qual marcará o período de acompanhamento das atividades realizadas no ambiente, e com base nesse período os alertas serão emitidos. Ainda nessa tela o professor edita os alertas acrescentando o nome dos alunos que farão parte da turma, bem como referenciar e direcionar se o alerta vai para o respectivo aluno e a data de criação do alerta, o motivo ao qual o evento do alerta será desencadeado, e o texto motivacional a ser inserido no sistema.

4. Metodologia e Validação do Alert-LV no ambiente Moodle

Realizou-se uma pesquisa quali-quantitativa com alunos de cursos de educação a distância oferecidos pelo Instituto Federal do Ceará (IFCE). O objetivo foi acompanhar uma turma no ambiente Moodle e investigar e experienciar a aplicação de um sistema automatizado de envio de mensagens de alertas vinculado a um modelo de avaliação processual, fundamentada em menções qualitativas e mediadas por linguagem iconográfica, a fim de melhorar a participação e desempenho dos mesmos na realização das atividades, bem como diminuir o número de discentes evadidos. Ao todo participaram da análise 15 alunos. O processo de análise foi realizado em 4 etapas onde o Professor/Tutor dentro do ambiente de aprendizagem ora se torna personagem principal, ora coadjuvante.

Na primeira etapa, optou-se por deixar apenas o tutor acompanhar a turma, emitindo as mensagens e notificações comumente aplicados no decorrer dos cursos por ele

ministrados. Na segunda etapa o Tutor além de emitir as mensagens comumente enviadas, passou a enviá-las com as cores e categorias as quais as mensagens do Alert-LV foram categorizadas. Este passo possibilitou a verificação da necessidade do agente automatizado em auxiliar o Professor/Tutor, bem como verificar a atuação do docente no sistema realizando ainda essa tarefa. Permitiu também observar as reações dos alunos com as mensagens modificadas e categorizadas.

Na terceira etapa o Professor/Tutor não mais agirá com o papel de agente, e sim haverá uma terceira pessoa fazendo esse papel. De forma que, essa pessoa responsável terá acesso similar, ou melhor, permissão semelhante às do Professor/Tutor. O 'agente' ficará emitindo as mensagens para os alunos de acordo com o que foi especificado nas funções do sistema automatizado e o Professor/Tutor emitirá as mensagens comumente aplicadas no decorrer do curso. O propósito desta etapa é verificar o desempenho do tutor e o dos alunos com a inserção deste agente no curso.

Na quarta etapa, o Professor/Tutor continuará o seu papel, já o 'agente', realizará o envio das mensagens dentro e fora do ambiente virtual de aprendizagem. Aqui o comportamento dos alunos será posto à prova. Serão verificadas a presença, ou melhor o alcance dos alertas emitidos. O propósito aqui foi identificar se os alunos entram no ambiente e recebem os alertas, ou se necessitam que a ferramenta tenha alcance além do AVA, evidenciando a necessidade de outros meios comunicacionais para se obter a atenção e atrair os alunos para a sala de aula virtual.

4.1. Aplicando a 1º etapa da análise

Como dito antes, na primeira etapa o tutor acompanhará a turma sem a intervenção dos alertas ou das características presentes neles. As mensagens emitidas são as mesmas que o professor comumente aplica no decorrer dos cursos por ele ministrados. Essas mensagens são os feedbacks que os Professores/Tutores emitem comumente durante o processo. Como objeto de análise, foi estipulado um período de observação dos comportamentos dos alunos, o que inclui: a quantidade de feedbacks emitido, o total de alunos presentes no curso, além da participação. Foram comparadas as duas primeiras semanas da disciplina em estudo. Os dados obtidos foram os seguintes: Referente a quantidade de feedbacks emitidos, na primeira semana e na segunda semana. Foram analisados os comportamentos do tutor, do aluno e do 'sistema de alertas'. Ao verificar as contribuições e interações realizadas pelo Tutor, verificou-se suas atitudes em emitir as mensagens de aviso "normais", ou seja, emitiu apenas as mensagens minimamente "necessárias" para o acompanhamento do curso. Enquanto que o emissor de alertas ficou apenas observando sem muitas interferências (Tabela 3).

	Fórum LV		**Tarefa LV**	
Atores envolvidos	Semana 1	Semana 2	Semana 1	Semana 2
Professor	8	8	7	8
Aluno	14	9	7	8
Sistema de alertas	-	-	-	-

Tabela 3 – Quantidade de postagens/intervenções

Dois alunos entregaram as atividades 1 três semanas depois e não participaram do fórum. Durante a primeira semana 8 alunos estavam ausentes, e em nenhum momento realizaram log de acesso ao sistema. Esse cenário sucedeu para a segunda semana, onde apenas 7 alunos haviam acessado o ambiente. Os outros permaneciam completamente ausentes do ambiente.

4.2. Aplicando a 2° etapa da análise

Na segunda etapa o Tutor passou a enviar as mensagens atribuindo-lhes cores e categorias semelhantes às mensagens do Alert-LV. Na análise observou-se um ligeiro acréscimo no número de participantes em comparação com as duas semanas anteriores. Compreende-se tal fato relacionando-o à mudança na abordagem e na emissão das mensagens do professor para os alunos. Nessas mensagens foi enfatizado a importância da presencialidade dos mesmos, por meio das mensagens com conteúdo iconográfico. Onze alunos participaram do Quiz, sendo realizadas 14 tentativas. Este passo possibilitou verificar a necessidade do sistema automatizado auxiliar o Professor/Tutor, bem como identificar a atuação do docente no sistema.

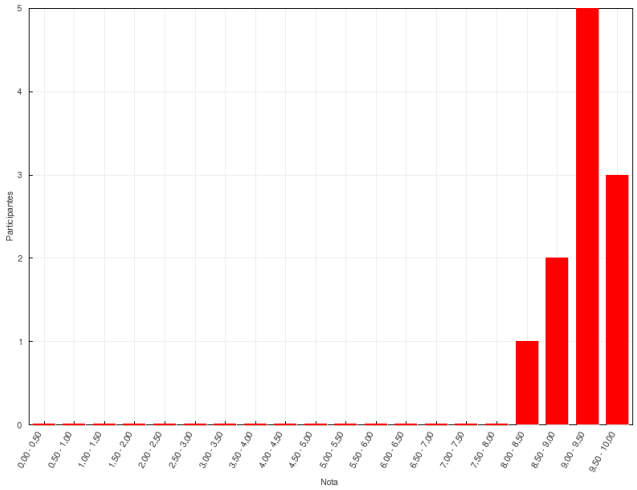


Figura 5 – Resultado da atividade

Extraindo como resultados as reações positivas dos alunos com as mensagens modificadas e categorizadas (Fig. 5), verificou-se que as notas dos alunos na atividade inclinaram-se a resultados mais satisfatórios. Percebe-se que os alunos se surpreenderam, ao receberem as mensagens de alerta categorizadas, e a ênfase no conteúdo das mesmas, denotando urgência, caso o desempenho não fosse favorável, ou mesmo parabenizando-os ao obterem notas satisfatórias na atividade.

4.3. Aplicando a 3° etapa da análise

Na terceira etapa o Professor/Tutor não mais agiu como o agente, foi substituído por uma terceira pessoa, ou seja, um 'sistema independente' fazendo esse papel. O novo sistema

obteve acesso e permissões similares às do Professor/Tutor. Foram emitidas mensagens constantes, e os alunos foram muito mais observados, transmitindo a sensação de vigilância automática. Muitas novas mensagens foram emitidas pelo 'Sistema de alertas', em proporção contrária à do Professor/Tutor (Tabela 4).

Atores envolvidos	**FórumLV**		**TarefaLV**	
	Semana 1	Semana 2	Semana 1	Semana 2
Professor	15	18	8	8
Aluno	20	13	8	8
Sistema de alertas	33	27	22	28

Tabela 4 – Quantidade de postagens/intervenções

Foi verificado que o Professor/Tutor por si só não poderia agir sobre todas as dúvidas dos alunos cursistas, assim como não conseguiria responder a tantas mensagens nos fóruns, nem identificar quantas postagens foram enviadas fora do prazo avaliativo pelos alunos. Com o apoio do 'Sistema de alertas' foi possível melhor monitorar a frequência de acesso dos alunos no ambiente, suas postagens de arquivos e interações nas ferramentas síncronas e assíncronas do Moodle.

4.4.Aplicando a 4° etapa da análise

Na quarta etapa, foram coletados os endereços de e-mails e telefones dos alunos, e o sistema passou a emitir mensagens via e-mails e mensagens via aplicativo de mensagens (whatsapp), o Professor/Tutor continuou o seu papel, já o agente, realizava o envio das mensagens dentro e fora do ambiente virtual de aprendizagem. Esse mecanismo serviu para verificar que as mensagens de alertas não funcionariam se enviadas apenas via AVA, pois alguns alunos não entravam no sistema com tanta frequência, e até alguns alunos estavam evadidos. Tais ações para obter a atenção e atrair os alunos para a sala de aula virtual e fazê-los reingressar nas atividades do curso foram satisfatórias, houve o retorno de 2 dos 4 alunos dados como evadidos.

5. Conclusões

Uma das principais funções em cursos na modalidade EaD por parte do Professor/Tutor é emitir constantes feedbacks aos alunos, entretanto essa função ocupa demasiado tempo e sobrecarga de trabalho. Utilizar de um sistema de monitoramento por meio de alertas automatizados pode facilitar a prática docente e auxiliar os tutores em suas tomadas de decisões, além de manter os alunos sempre informados de suas responsabilidades junto ao curso. Afinal o trabalho de envio de feedbacks comumente realizado, nem sempre é suficiente para monitorar todos os discentes. O envio automatizado de alertas vem contribuir e retirar do professor a preocupação de sua 'onipresença' no ambiente, colaborando ainda mais para a sensação de presencialidade dos atores envolvidos em EaD.

Desta forma, este trabalho propôs o sistema automatizado de alertas, o Alert-LV, que alinhado ao modelo de avaliação quali-quantitativo *Learning Vectors*, faz envio automatizado de feedbacks aos alunos. O diferencial do Alert-LV é valer-se do conceito de avaliação formativa, processual e de acompanhamento constante facilitada pela

métrica pedagógica Fator β do Modelo LV, tal fator ajuda a categorizar o desempenho dos alunos, assim as mensagens de alerta e feedbacks foram intensificados. O Alert-LV pode contribuir também na redução da evasão ao auxiliar no resgate dos alunos ausentes da sala de aula virtual. Buscar alternativas de comunicação externas ao ambiente Moodle foi o que diferenciou este trabalho dos demais.

Como trabalhos futuros propõem-se inserir no Modelo LV outros mecanismos de inteligência artificial que aprimorem a avaliação a ser realizada pelo tutor. Ou ainda utilizar de linguagens de marcação como AIML para transformar e tornar mais dinâmicas as notas LV. Utilizando destes sistemas os próprios emoticons dos LVs seriam transformados em chatterbots capazes de "conversarem" com os alunos, os LVs passariam a deter de "inteligência" e interatividade.

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) - Brasil - Código de Financiamento 001.

Referências

Almeida, A. de (2013). *Estratégias para o Monitoramento de Ações de Tutoria na Educação a Distância*. Dissertação, Curso de Engenharia da Informação, Universidade Federal do ABC, Santo André, pp. 102.

Jade, J. *Agent Development Framework*. (2016). Disponível em: < http://jade.tilab.com > Acessado em: 01 nov. 2018.

Kampff, A. J. C. (2009). *Mineração de Dados Educacionais Para geração de alertas em Ambientes Virtuais de Aprendizagem como apoio à pratica docente*. Tese, Porto Alegre.

Leite, E. A. M., Machado A., Sales G. L., Zednik H., & Maia S. M. (2015). Distance Education Evaluation: An Analysis of the β Factor from LV Model Subjectivity. In: Rocha A., Correia A., Costanzo S., Reis L. (eds) *New Contributions in Information Systems and Technologies. Advances in Intelligent Systems and Computing*, vol 353. Springer, Cham. Disponível em: https://goo.gl/jiAxdK Acessado em: 21 Jun 2018.

Losada, M., & Heaphy, E. (2004). The role of positivity and connectivity in the performance of business teams: A nonlinear dynamics model. *American Behavioral Scientist,* vol. 47, nº 6, 740-765.

Mercado, L. P. L. (2007). Dificuldades na educação a distância online. In: *Congresso internacional de educação a distância*. Curitiba.

Moore, M. (1993). *Theoretical principles of distance education*. New York: Routledge, p. 20-35.

Moore, M., & Kearsley, G. (2007). *Educação a Distância*. São Paulo: Cengage Learning, pp. 398.

Musa, D. L. (2001). *Um sistema de Alertas Inteligentes para Ambientes de Ensino na Internet*. Rio Grande do Sul, Porto Alegre, pp.123.

Sales, G. L. (2010). *Learning Vectors (LV): um modelo de avaliação da aprendizagem em EaD online aplicando métricas não-lineares*. Tese Doutorado. Departamento de Engenharia de Teleinformática. Universidade Federal do Ceará. 239f. Disponível em: < https://goo.gl/2ULNJU > Acessado em: 08 Maio 2018.

Sales, G. L., Barroso, G. C., & Soares, J. M. (2012). Learning Vectors (LV): A Procedural Assessment Model with Non-Linear Measurement of Learning in Distance Education Online. *Brazilian Journal of Computers in Education*. Available at: <http://www.br-ie.org/pub/index.php/rbie/article/view/1368>. Date accessed: 20 jun. 2018.

Sales, G. L., Albuquerque, M. C. N., Leite, E. A. M., & Paz, R. P. (2018). Dez Anos do Modelo Learning Vectors na Avaliação Formativa no AVA Moodle: de Emoticons a Gifs Animados In: *Tecnologias da Educação: passado, presente, futuro. Coleção História da Educação*.1 ed.Fortaleza: UFC, v.1, p. 63-78.

Silva, F. A. G. da (2018). *Alert-LV: Um sistema de monitoramento e apoio à tutoria para suporte à avaliação em ambientes virtuais de aprendizagem*. Dissertação (Mestrado). Programa de Pós-Graduação em Ciência da Computação (PPGCC) do Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE). 97f.

Silveira, R. A. (2001). *Modelagem Orientada a Agentes Aplicada a Ambientes Inteligentes Distribuídos de Ensino: JADE - Java Agent Framework Learnings Environments*. Porto Alegre, pp. 133.

Tori, R. (2010). *Educação sem distância: as tecnologias interativas na redução de distâncias em ensino e aprendizagem*. São Paulo: Editora Senac.

Valente, J. A. (1999). *Diferentes abordagens de educção a Distância*. São Paulo, pp. 77.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 02/09/2018
Aceitação/Acceptance: 23/11/2018

Avaliação Heurística e Teste de Usabilidade para Software de *Design* de Interiores

Noemi Pereira Scherer Cilumbriello[1], Valéria Farinazzo Martins[2], Maria Amélia Eliseo[2], André Luiz Satoshi Kawamoto[1]

noemischerer13@gmail.com, mariaamelia.eliseo@mackenzie.br, valeria.farinazzo@mackenzie.br, kawamoto@utfpr.edu.br

[1] Federal University of Technology – Paraná, Via Rosalina Maria dos Santos, 1233, 87301-899, Campo Mourão – PR, Brazil

[2] Mackenzie Presbyterian University, Rua da Consolação 930, 01302-907, São Paulo-SP, Brazil

Pages: 90–101

Resumo: Um *designer* de interiores é o profissional responsável por criar e modificar qualquer espaço interno por meio do planejamento e execução de um projeto. Ele utiliza como auxílio para seu trabalho softwares capazes de criar e mostrar um ambiente por meio de projeções bidimensionais ou tridimensionais. Muitos desses programas podem apresentar problemas em suas interfaces, dificultando ou impossibilitando seu uso pelos usuários. Assim, surge a necessidade de realizar uma avaliação capaz de identificar e classificar esses problemas. Esse trabalho visa aplicar duas avaliações de usabilidade: heurística e um teste de usabilidade. A primeira utiliza as heurísticas propostas por Nielsen, originalmente específicas para aplicações *Web*, que foram adaptadas para o contexto de *design* de interiores, enquanto a segunda utiliza *System Usability Scale (*SUS). Para comprovar sua importância, essas avaliações foram aplicadas a um software real. Os resultados obtidos permitem classificar a usabilidade da interface do programa avaliado.

Palavras-chave: *design* de interiores*;* usabilidade; avaliação heurística; teste de usabilidade; *system usability scale*.

Heuristic evaluation and usability test for interior design software

Abstract: An interior designer is the professional responsible for creating and modifying any internal space through the planning and execution of a project. He uses software to create and display an environment through two-dimensional or three-dimensional projections. Many of these programs may have problems in their interfaces, making it difficult or impossible for users to use. Thus, it is necessary arises to carry out an evaluation capable of identifying and classifying these problems. The present work applies two usability evaluations, a heuristic and a usability test. The first one is based on the heuristics proposed by Nielsen, originally specific to Web applications, that were adapted in this work to the context

of interior design, whereas the second uses the System Usability Scale (SUS). To confirm the importance of these evaluations, they were applied to real software. The results allowed to classify the interface software of the evaluated program.

***Keywords**:* interior design; usability; heuristic evaluation; usability test; system usability

1. Introdução

Um *designer* de interiores é o profissional responsável por cuidar de espaços internos, tanto residenciais como corporativos, seguindo leis ergonômicas, acústicas, térmicas e luminotécnicas. Ele deve verificar as necessidades dos seus clientes e as atribuir ao projeto, delinear completamente o ambiente, mas não deve alterar nenhuma estrutura física (Vela et al., 2016).

A evolução tecnológica tem possibilitado a criação de diversos softwares de projeção 2D e 3D que podem ser utilizados por *designers* de interiores, com diferentes funcionalidades entre si (Montenegro, 1987). Sua utilização torna o trabalho de um *designer* mais fácil e prático, reduz o tempo de criação, revisão e alteração do projeto, garante mais economia, boa precisão e melhora do fluxo das informações. Entretanto, estudos indicam que as interfaces desses programas são projetadas de maneira incorreta, tornando a execução de uma tarefa mais difícil e de maior tempo (Medeiros, Cybis, 2000).

A usabilidade de um programa é um fator relacionado à percepção dos usuários a respeito de um sistema durante sua interação. Assim, ela pode ser usada para avaliar a qualidade de um sistema. Entre os instrumentos existentes para tal, podemos citar a avaliação heurística e o teste de usabilidade, que ajudam a determinar quais recursos do programa não foram projetados de maneira efetiva aos usuários.

O presente trabalho apresenta os resultados de uma avaliação heurística e de um teste de usabilidade aplicados a um programa de *design* de interiores. A avaliação foi adaptada do conjunto de heurísticas originalmente proposto por Nielsen (Nielsen, Mack, 1994) para o contexto da aplicação, enquanto o teste de usabilidade é baseado no SUS (*System Usability Scale*) (Brooke et al., 1996).

Os dois instrumentos de avaliação são usados por diferentes tipos de avaliadores: a avaliação heurística deve ser feita por especialistas da área de Interação Humano-Computador (IHC), e o teste de usabilidade por usuários finais. A aplicação desses dois tipos de avaliação deve ser capaz de indicar diferentes tipos de erros e, ao mesmo tempo, indiretamente, aferir a qualidade de um sistema.

O presente artigo está organizado da seguinte forma: na Seção 2 é apresentado o conceito de Interação Humano-Computador (IHC) e dois métodos de avaliação de usabilidade; na Seção 3 são definidas as heurísticas adaptadas para o contexto de *design* de interiores; na Seção 4 são apresentados os processos adotados na avaliação heurística e no teste de usabilidade; a Seção 5 apresenta os resultados obtidos e, por fim, na Seção 6, são apresentadas as conclusões e trabalhos futuros.

2. Referencial Teórico

2.1. Interação Humano-Computador

A Interação Homem-Computador, ou Interface Humano-Computador, (IHC) é a área da Ciência da Computação que se preocupa com o *design*, avaliação e desenvolvimento de interfaces para que usuários finais possam interagir com sistemas computacionais de uma forma eficiente e intuitiva (Oliveira, 2011).

Um dos maiores desafios da IHC é acompanhar a evolução tecnológica e atender, ao mesmo tempo, a todos os grupos de usuários, uma vez que nem todos os usuários conseguem dominar as funcionalidades de um sistema. Dessa forma os desenvolvedores dos softwares possuem o desafio de criar novas interfaces, mas que não deixem de lado a objetividade e a clareza (Oliveira, 2011).

Um conceito intimamente relacionado à IHC é o de usabilidade. A usabilidade é uma medida que permite o uso de determinados produtos por usuários para eles alcançarem objetivos específicos com eficácia, eficiência e satisfação (Alves et al., 2015, Preece et al., 2005, Gloria, 2015).

Quando um ou mais usuários encontram dificuldades em realizar qualquer tarefa, pode-se afirmar que a interface de um sistema apresenta problemas de usabilidade, o que pode causar perda de dados, diminuição da produtividade e até mesmo a desistência do uso do software (Winckler, Pimenta, 2002).

Para garantir que a usabilidade esteja presente em um sistema são realizadas avaliações de usabilidade. Essas técnicas podem ser classificadas como métodos indiretos e métodos diretos.

Nos **métodos indiretos** os avaliadores analisam a usabilidade do sistema aplicando técnicas de coleta de dados sem a necessidade de envolver usuários finais, cujo principal objetivo é encontrar problemas de usabilidade (Gloria, 2015).

A avaliação heurística é um exemplo de método indireto, a qual permite encontrar diferentes problemas em interfaces de sistemas. Consiste de um conjunto de regras gerais (heurísticas) que descrevem propriedades comuns em interfaces derivadas do conhecimento de aspectos psicológicos, computacionais e sociológicos (Gloria, 2015).

Os dez princípios fundamentais de usabilidade desenvolvidos por Nielsen fornecem uma fundamentação para realizar uma avaliação de produtos de software. Originalmente voltadas para aplicações *Web*, essas dez heurísticas podem ser utilizadas em outros tipos de aplicações e combinadas com outras análises, desde que essas realizem avaliações em produtos com o objetivo de determinar a usabilidade da interface do sistema (Preece et al., 2005). As dez heurísticas definidas por Nielsen são: *Visibilidade do status do sistema*; *Compatibilidade do sistema com o mundo real*; *Controle do usuário e liberdade*; *Consistência e padrões*; *Ajuda os usuários a reconhecer, diagnosticar e recuperar de erros; Prevenção de erros; Reconhecimento em vez de memorização; Flexibilidade e eficiência de uso; Estética e design minimalista;* e *Ajuda e documentação* (Nielsen, Mack, 1994).

Os **métodos diretos**, por sua vez, utilizam ferramentas e instruções em simulações e as aplicam ao próprio usuário do sistema, com o objetivo de observar seu comportamento e

ouvir sua opinião. As interações com usuários acontecem por meio de questionários ou sobre a experiência de uso com o sistema (Machado et al., 2014).

O teste de usabilidade é um método direto que tem como principal objetivo avaliar a qualidade de interação entre o usuário e um sistema, determinando e medindo impactos da interação e identificar problemas na interface responsáveis por gerar desconfortos aos usuários (Machado et al., 2014).

Nesse teste, o ambiente de realização deve ser real ou próximo do real e os avaliadores devem criar um roteiro o qual os participantes da pesquisa devem seguir. Durante a execução desse roteiro, os avaliadores deverão acompanhar e monitorar todas as decisões realizadas pelos participantes. Além disso, é preciso organizar materiais, selecionar usuários e definir um modelo de questionário e análise dos resultados. Nessa proposta, adotou-se a escala *System Usability Scale* (SUS).

A SUS consiste de um questionário composto por dez questões, com cinco opções de respostas cada uma, que variam desde "Discordo totalmente" a "Concordo totalmente" (Brooke et al., 1996). Essas questões são apresentadas na Tabela 1.

O resultado do SUS é obtido pela soma da contribuição individual de cada questão, definida como *score*. Para as questões ímpares o *score* é dado pela subtração de 1 (um) à resposta do usuário, já para as pares é dado pela subtração da resposta do usuário por 5 (cinco). Após obter o *score* de cada questão, esses são somados e multiplicados por 2,5. Assim, o resultado final será um índice de satisfação do utilizador, variando de 0 a 100 (Brooke et al., 1996).

1.	**Eu gostaria de utilizar este sistema frequentemente**
2.	**Eu achei o sistema desnecessariamente complexo**
3.	**Eu achei o sistema fácil de utilizar**
4.	**Eu acho que precisaria do apoio de um suporte técnico para ser possível utilizar este sistema**
5.	**Eu achei que diversas funções deste sistema foram bem integradas**
6.	**Eu achei que houve muita inconsistência neste sistema**
7.	**Eu imaginaria que a maioria das pessoas aprenderiam a utilizar esse sistema rapidamente**
8.	**Eu achei o sistema muito pesado para uso**
9.	**Eu me senti muito confiante utilizando esse sistema**
10.	**Eu precisei aprender uma série de coisas antes que eu pudesse continuar a utilizar esse sistema.**

Tabela 1 – Questões do SUS

De acordo com Boucinha e Tarouco, cada grupo de questões do SUS permite mapear os componentes de qualidade indicados por Nielsen: a facilidade de aprendizagem (questões 3, 4, 7 e 10); a eficiência (questões 5, 6 e 8); a facilidade de memorização (questão 2); a minimização dos erros (questão 6); e a satisfação (questões 1, 4 e 9) (Boucinha; Tarouco, 2013).

3. Adaptação do conjunto de Heurísticas de Nielsen para o contexto de softwares de *design* de interiores

Adaptar o trabalho de Nielsen para a área de *design* de interiores, nesse projeto, consiste em, baseado na definição de cada heurística, apresentar exemplos de funcionalidades as quais programas dessa área devem conter. Essas adaptações são descritas a seguir.

Visibilidade do status do sistema: o programa deve fornecer *feedback* ao usuário quando realizada alguma ação. Para ações que são divididas em etapas o usuário deve ser informado constantemente em qual etapa a ação se encontra. Por exemplo, durante a importação de um objeto, o sistema deve apresentar uma barra indicando o progresso do andamento da operação.

Compatibilidade do sistema com o mundo real: A linguagem do sistema deve ser clara, de forma que o usuário esteja familiarizado com ela. Para facilitar a compreensão, é possível utilizar imagens que correspondem à descrição. Um programa de *design* de interiores deve possuir diversas ferramentas necessárias para seu uso e, para facilitar a localização dessas ferramentas, podem existir imagens metafóricas correspondentes às suas descrições, como o símbolo de parte de um cubo, como uma analogia que corresponde à criação de paredes.

Controle do usuário e liberdade: Os usuários precisam da sensação de controle sob todas as ações do sistema. Para garantir isso, erros no programa devem ser evitados, como o fechamento da interface gráfica durante o seu uso, ou propagandas que são abertas na execução de um recurso do sistema. Em qualquer sistema o usuário deve conseguir cancelar a ação e a mesma ser cancelada, como por exemplo, ao importar um novo modelo para o programa, deve ser possível cancelar essa opção e retomar ao estado anterior.

Consistência e padrões: O programa deve possuir a linguagem, janelas, mensagens de erros, botões, hierarquia de cores e imagens padronizadas. Em um programa de *design* de interiores as mensagens de erros, avisos e confirmações de uma ação devem ser padronizadas, ou seja, devem apresentar os mesmos ícones e linguagem utilizada.

Prevenção de erros: Os erros são desagradáveis e também as principais causas de frustração com o programa, por esse motivo o software deve evitar que eles aconteçam. As informações devem ser bem organizadas e as ferramentas que o programa apresenta devem ser bem descritas para não haver erros ao selecioná-la, não obtendo ações esperadas. Quando não é possível realizar uma ação em um programa, ela deve ser desabilitada para que o usuário não corra o risco de selecioná-la e obter nenhuma resposta. Por exemplo, filtros nas janelas para impedir a importação de arquivos em formato errado ou não reconhecido pelo sistema.

Reconhecimento em vez de memorização: As informações ou opções disponíveis no programa devem ser intuitivas, de forma que o usuário não necessite lembrar o seu significado. Para satisfazer essa heurística é necessário informar ao usuário em qual janela ou qual ação está sendo realizada. Um exemplo mais comum, é que a maioria dos programas deve oferecer, é o *menu*. Nele, deve ser possível identificar a localização e as ações possíveis, como os diversos atalhos demonstrados no menu ferramentas.

Flexibilidade e eficiência de uso: O programa deve ser criado de forma que atenda as necessidades de diferentes perfis de usuários, desde os iniciantes aos experientes. O

sistema deve fornecer interações que atendem a todos os tipos de usuários. Por exemplo, usuários mais experientes usam o atalho do teclado, já os usuários novatos navegam pelos menus.

Estética e design minimalista: As principais informações e opções do programa devem estar presentes ao iniciá-lo. Elas devem ser destacadas, ou seja, devem ser exibidas de forma que o usuário não precise procurar por elas (fácil visualização). Por exemplo, o programa deve exibir uma barra de tarefa contendo as suas principais funções (criação de paredes, pisos, seleção e ajuda), uma área que permite a visualização ou criação do projeto e dos objetos nele contido.

Ajuda aos usuários no reconhecimento, diagnóstico e correção de erros: O programa deve fornecer uma forma para que os usuários solucionem possíveis erros encontrados no mesmo, garantindo o reconhecimento da gravidade do problema. Por exemplo, ao encontrar uma mensagem informando determinado erro, a mesma deve explicar o motivo ou conter uma solução.

Ajuda e documentação: O programa deve fornecer ajuda para tirar dúvida dos usuários. Um exemplo específico para um programa na área de *design* de interiores é a documentação apresentada ao usuário quando se seleciona a opção "AJUDA" no menu, que explica como o usuário deve proceder quando deseja acionar alguma funcionalidade específica.

4. Aplicação das Avaliações

As duas avaliações propostas foram aplicadas a um software conhecido como Sweet Home 3D, um programa de *design* de interiores *Open Source*[1] disponível tanto para uso *online* quanto para instalação no computador do usuário, o qual permite criar um ambiente em um plano 2D e simultaneamente realizar visualização em 3D.

4.1. Aplicação da Avaliação heurística

Para realizar a avaliação heurística, foram convidados especialistas da área de Interação Humano-Computador, inclusive com formação em Arquitetura em Urbanismo.

Inicialmente, cada avaliador navegou na interface do programa até encontrar um problema de usabilidade. Feito isso, o problema foi classificado de acordo com as heurísticas propostas. Por exemplo, se o avaliador não conseguisse localizar uma ferramenta de ajuda no programa, deveria categorizar esse problema "Heurística 10: Ajuda e documentação". Cada avaliador realizou esse passo três vezes, no mínimo.

Todos os problemas encontrados foram anotados em uma tabela e, depois, foi atribuído um grau de severidade, variando de 0 a 4, de acordo com a opinião do avaliador: 0 (zero) para um problema sem importância; 1 (um) para problema cosmético, ou seja, não requer correção, a menos que se tenha tempo extra; 2 (dois) para aqueles considerados simples; 3 para os graves; e 4 para os catastróficos. Também foram anotados o contexto e a possível causa, informações que facilitam a análise da falha e a identificação de meios para corrigi-la.

[1] http://www.sweethome3d.com/pt/features.jsp

4.2. Aplicação do Teste de Usabilidade

Para o teste de usabilidade não foi necessário adaptar a escala SUS, uma vez que suas questões são abrangentes o suficiente para detectar problemas de usabilidade na interface de sistemas.

O teste foi aplicado ao final de um minicurso para capacitar usuários familiarizados com programas da área de *design* de interiores. Para assegurar usuários com esse perfil, foram convidados professores e alunos de um curso de graduação na área de Arquitetura e Urbanismo. No total, trinta e dois alunos participaram da avaliação.

O minicurso contou com um material, disponibilizado aos alunos em formato digital, que continha exemplos de como utilizar as ferramentas disponíveis e os processos básicos para a criação de um projeto no programa, tais como: a instalação do software em diferentes sistemas operacionais; como importar plantas-baixas, desenhar e editar paredes, adicionar portas, janelas e móveis; importar objetos 3D e desenhar cômodos.

Para a realização do minicurso e aplicação do teste, uma equipe, devidamente treinada, ficou responsável por auxiliar e observar os usuários. A tarefa escolhida foi importar e projetar os ambientes de uma casa de 19.50 m² com dois cômodos, um banheiro e uma cozinha anexa ao quarto.

Os alunos responderam às dez questões do SUS por meio de um formulário online. Além dessas questões, uma questão discursiva não obrigatória, na qual o avaliador poderia descrever oportunidades de melhoria do sistema, foi adicionada.

5. Resultados e Análises dos Resultados

5.1. Resultados da avaliação heurística

Como resultados da avaliação heurística, onze problemas na interface do Sweet Home 3D foram encontrados. A distribuição e a classificação desses problemas são discutidas a seguir.

Três problemas foram classificados na heurística de “Compatibilidade do sistema com o mundo real”, sendo dois considerados simples e um grave. O primeiro problema simples especifica que a nomenclatura adotada pelo software não corresponde à realidade ao se alterar a cor de uma parede. O usuário escolhe uma cor do lado “esquerdo” e outra cor do lado “direito”. Entretanto, dependendo da posição não existem tais lados, mas sim um lado “interno” e um lado “externo”. O segundo problema simples ocorre quando usuário clica com o botão direito sobre um objeto qualquer para ativar o módulo de edição das características do mesmo. Nesse momento, são mostradas apenas as funcionalidades gerais em vez de informações sobre o objeto clicado. Por fim, o problema grave considerado grave diz respeito ao nome do comando “construir cômodo”, que pode confundir o usuário, pois na verdade ele permite inserir piso no ambiente (tanto no pavimento térreo quanto em outros).

No que se refere à heurística “Controle do usuário e liberdade”, foi encontrado um problema considerado grave na função de desenhar paredes: o usuário não consegue desenhar paredes no contexto 3D. Não se sabe, entretanto, se essa função realmente está implementada no sistema.

Um problema considerado simples foi apontado na heurística "Consistências e padrões": mesmo após a aplicação dos *patches* para correção do tamanho da fonte, algumas telas de diálogo permaneceram com fonte bem diferente. A causa identificada foi de que a correção do tamanho da fonte proposto na documentação não foi abrangente a todas as telas de interface do sistema.

Para a heurística "Reconhecimento em vez de memorização", os avaliadores detectaram três problemas, um grave e dois simples. O problema grave está relacionado à definição de dimensões precisas ao desenhar paredes ou cômodos. Não existe no menu ou caixa de ferramenta relacionada uma opção (caixa de texto) para o usuário inserir os valores exatos que ele deseja. Os dois problemas considerados simples dizem respeito, respectivamente, a não encontrar a possibilidade de inserir portas com aberturas em ambos os sentidos (abrir para a esquerda, e abrir para a direita), e a exportar o projeto para o formato SVG, o qual deveria ser opção do menu "arquivo", mas se encontra no menu "Plano". Em tempo, todas as opções de exportar deveriam estar no menu "Arquivo", uma vez que esse é o padrão em outros softwares.

Um problema grave, classificado como "Flexibilidade e eficiência de uso", está relacionado à inserção de paredes com uma dimensão específica, não havendo alternativas para inserir essas dimensões como um campo de entrada ou um comando específico no menu.

Para heurística "Estética e design minimalista" um problema foi considerado catastrófico: ao iniciar o aplicativo, a interface aparece com a lista de itens do menu à direita sobrepostos, bem como todos os textos da interface. Esses textos permanecem ilegíveis mesmo após maximizar a janela da aplicação.

Finalmente, um problema grave, relacionado ao problema relatado de textos sobrepostos, foi classificado na heurística "Ajuda e documentação". De acordo com o avaliador, uma solução reportada no fórum dos desenvolvedores indicava inserir um arquivo de configuração. Essa solução, entretanto, não foi adequada, uma vez que os textos da interface continuaram sobrepostos (agora com fontes maiores). A solução final foi encontrada numa lista de discussão de usuários, que seria um local inadequado.

Para as heurísticas "Visibilidade do status do sistema", "Prevenção de erros" e "Ajuda aos usuários no reconhecimento, diagnóstico e correção de erros", nenhum dos avaliadores encontrou problema. Porém, isso não garante que o software avaliado não apresente problemas de interface relacionados a essas heurísticas, uma vez que diferentes especialistas poderiam apontar problemas que se enquadram nessas classificações.

Apesar de não ter sido considerado um erro, foi sugerido que o software disponibilize vídeos explicativos, uma vez que, atualmente, a ajuda é composta apenas por textos e muitos usuário desistem de ler e realizam qualquer tarefa em forma de "tentativa-e-erro". Essa sugestão está relacionada à heurística "Ajuda e documentação".

O grau de severidade facilita o estabelecimento de prioridades, determinando a ordem das correções dos problemas. Dessa forma, o pior problema encontrado, classificado como catastrófico, diz respeito à interface, que apresenta textos sobrepostos, impossibilitando a utilização do programa.

5.2. Resultado do teste de usabilidade

5.2.1. Análise dos resultados por avaliador

Utilizando a fórmula proposta pela SUS, foi calculada a pontuação para cada avaliador um dos 32 avaliadores (Figura 1).

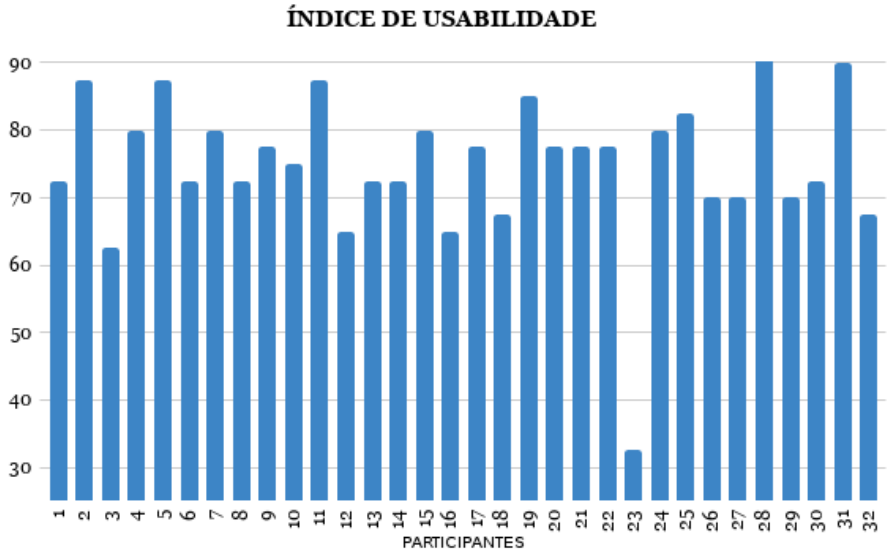


Figura 1 – Resultado da SUS por Avaliador

De acordo com Sauro (2009), após analisar resultados do SUS para mais de 5.000 usuários em 500 avaliações diferentes, pontuações médias abaixo de 68 indicam que a usabilidade pode estar comprometida, sendo aconselhável realizar mudanças na interface desse sistema.

É possível observar que, do total de 32, a grande maioria (28 usuários) pontuou o sistema com resultados muito próximos ou superiores a 68. Dos usuários restantes, 3 pontuaram abaixo (usuário 3 com 62.5 pontos e os usuários 12 e 16 com 65 pontos), e apenas um usuário pontuou com um valor muito inferior (usuário 23 com 32.5 pontos).

A fim de investigar melhor esses resultados, os usuários que menos pontuaram o sistema foram separados em dois grupos, o primeiro contendo os usuários 3, 12 e 16, e o segundo com o usuário 23. Nessa investigação, verificou-se a existência de resposta para a questão discursiva, além de quais questões foram respondidas de forma negativa.

Para o primeiro grupo, observou-se que a maioria dos usuários ficou satisfeita com as funcionalidades do programa, em especial o usuário 16, que deixou uma opinião bastante positiva quanto ao mesmo. Entretanto, foi apontada a necessidade de apoio técnico para o uso do sistema, bem como melhorias que poderiam ser introduzidas.

O usuário 23, por sua vez, não respondeu a questão discursiva, que poderia contribuir para determinar qual foi a principal causa de frustração com o programa. Porém, levando em consideração as suas respostas, foi possível concluir que esse usuário não se sentiu confiante em usar o software, pois suas respostas indicam que ele considerou o sistema inconsistente, além de complexo e "pesado" para uso, tendo a necessidade de apoio técnico. Além disso, todas as questões pares do questionário (que possuem maior peso para a avaliação negativa do sistema) foram mal avaliadas justificando uma pontuação tão inferior quando comparada com as outras.

Apesar disso, a avaliação geral do sistema é considerada boa. A média das 32 avaliações foi de 75 pontos, indicando que a usabilidade do programa não está comprometida.

5.2.2. Análise dos resultados de cada questão do SUS

A pontuação obtida em cada questão foi ponderada para a escala de 0 a 100. A Tabela 2 mostra esses resultados.

Questão	1	2	3	4	5	6	7	8	9	10
Pontuação	75.78	85.94	78.13	68.75	71.09	65.63	78.91	79.69	73.44	72.66

Tabela 2 – Resultados das Questões do SUS

Tenório et al. (2010), por sua vez, associam cada componente de qualidade de Nielsen à média de um conjunto de questões específicas do SUS. A **facilidade de aprendizagem do sistema** está associada às questões 3, 4, 7 e 10. A média dos resultados dessas questões no teste aplicado foi de 74.61, uma pontuação considerada neutra. Apesar dos usuários terem conseguido realizar a tarefa proposta, eles consideram que o sistema poderia ser mais intuitivo, facilitando o seu uso.

A **eficiência do sistema** é indicada pelas questões 5, 6 e 8. A média dos resultados dessas questões foi de 72.41, uma pontuação nem boa nem ruim. Dessa forma, é possível concluir que o sistema poderia apresentar um melhor nível de produtividade e desempenho.

A **facilidade de memorização** está associada à questão 2 e a sua pontuação foi de 85.94, um resultado considerado muito bom. Assim, é possível afirmar que os usuários foram capazes de realizar suas tarefas no sistema sem a necessidade de reaprender como interagir com o mesmo.

As **inconsistências do sistema** ou **minimação dos erros** são identificadas por meio da questão 6, cuja pontuação foi de 65.625, um valor considerado ruim. Assim, é possível afirmar que o sistema não apresenta soluções simples ou rápidas para eventuais erros. Alguns testes mais elaborados podem ser feitos no programa com o objetivo de encontrar outras irregularidades.

A **Satisfação dos usuários com o sistema** está representada nas questões 1, 4 e 9. A média dos resultados dessas questões foi de 72.66 pontos. Em geral, a satisfação refere-se ao nível de conforto que o usuário sente ao utilizar o sistema, de forma que ele alcance seus objetivos ao navegar no mesmo. Dessa forma, é possível concluir que

os usuários poderiam estar mais confortáveis com o sistema se algumas mudanças na interface fossem realizadas.

Identificar oportunidade de melhoria do sistema: Embora a usabilidade do Sweet Home 3D tenha sido classificada como boa em quase todas as questões, a observação das respostas dadas às questões discursivas possibilitou analisar a existência de insatisfações e dificuldades em relação ao programa, como: a interface com um visual muito antigo e bordas quadradas; dificuldade em movimentar a planta da casa; necessidade de mais opções de visualização do projeto; o mal funcionamento da ferramenta zoom; necessidade de aumentar a sensibilidade do teclado para diminuir a dependência do mouse para ações com o sistema, ou seja, mais teclas de atalhos; e facilitar a inserção de telhados ao projeto.

6. Conclusões e Trabalhos Futuros

Este trabalho propõe uma adaptação do conjunto de heurísticas de Nielsen e uma abordagem para aplicação de um teste de usabilidade, baseado na escala SUS, para o contexto de programas de *design* de interiores.

Para validar essa proposta, essas avaliações foram aplicadas a um programa conhecido como Sweet Home 3D. Os resultados dessas avaliações determinaram diferentes tipos de erros. O problema mais grave apontado na avaliação heurística diz respeito a "Estética e design minimalista". Entretanto, esse problema não foi percebido pelos outros avaliadores, tampouco durante o teste de usabilidade. Possivelmente o contexto no qual esse problema foi detectado era bastante específico, por isso não foi replicado pelos demais avaliadores.

Por outro lado, alguns problemas de usabilidade apontados pelo teste não foram percebidos pelos especialistas, como por exemplo o mal funcionamento da ferramenta de *zoom*. Esse fato evidencia a importância de aplicar diferentes avaliações, determinando o maior número de falhas para melhor classificar a interface do sistema.

A principal contribuição desse trabalho é a adaptação das heurísticas para o contexto de *design* de interiores. O uso da escala SUS mostrou-se adequado para classificar a usabilidade de sistemas, como aconteceu com o Sweet Home 3D, bem como para identificar componentes de qualidade que necessitam de mudança e, ainda avaliar a opinião e classificação do sistema por cada usuário.

Como trabalhos futuros, propõe-se aplicar o conjunto de heurísticas adaptado nesse trabalho para outros softwares da área, a fim de validar sua aplicabilidade e utilidade nesse contexto. Também, é possível aplicar o teste de usabilidade usando outros questionários, a fim de observar se os resultados são condizentes com a qualidade aferida pelo SUS nesse trabalho.

Referências

ALVES, João Marcus; SAVARIS, Alexandre; WANGENHEIM, Christiane Gresse von; WANGENHEIM, Aldo von. Criando e validando um checklist de avaliação heurística a para sistemas de laboratórios clínicos. *Computer on the Beach*, p. 1–10, 2015.

BOUCINHA, Rafael Marimon; TAROUCO, Liane Margarida Rockenbach. Avaliação de ambiente virtual de aprendizagem com o uso do sus-system usability scale. *RENOTE*, v. 11, n. 3, 2013.

BROOKE, John et al. Sus-a quick and dirty usability scale. *Usability evaluation in industry*, London–, v. 189, n. 194, p. 4–7, 1996.

GLORIA, Heloisa de Souza. Avaliação de um conjunto de heurísticas de usabilidade para aplicativos de smartphones na área da saúde por meio de testes de usabilidade. p. 1–75, 2015.

MACHADO, Lais; FERREIRA, Evelise Pereira; VERGARA, Lizandra Garcia Lupi. Métodos de avaliação de usabilidade: Características e aplicações. 30. *CONEPRO-SUL, Joinville. 30. CONEPROSUL*, 2014.

MEDEIROS, Marco Aurélio; CYBIS, Walter de Abreu. Método de avaliação de usabilidade de software a partir da satisfação de usuários e da aplicação de quesitos da norma isso 9241. In: *IHC'2000–III Workshop sobre Fatores Humanos em Sistemas de Computação*. [S.l.: s.n.], 2000. p. 93–101.

MONTENEGRO, Gildo. *A invenção do projeto*. [S.l.]: Edgard Blucher, 1987. ISBN 9788521200079.

NIELSEN, Jakob; MACK, Robert L. *Usability Inspection Methods Computer*. [S.l.]: John Wiley & Sons, 1994. ISBN 0471018775.

OLIVEIRA, Igor Aguiar. Interface de usuário: A interação homem-computador através dos tempos. Olhar Científico, v. 1, n. 2, p. 178–184, 2011.

PREECE, Jennifer; ROGERS, Yvone; SHARP, Helen. *Design de Interação*. [S.l.]: Bookman, 2005. ISBN 9788536304946.

SAURO, Jeff. *Measuring Usability with the System Usability Scale (SUS)*. 2009. https://measuringu.com/sus/. Acessado em 08-04-2018.

TENÓRIO, Josceli Maria; COHRS, Frederico Molina; SDEPANIAN, Vera Lúcia; PISA, Ivan Torres; MARIN, Heimar de Fátima. Desenvolvimento e avaliação de um protocolo eletrônico para atendimento e monitoramento do paciente com doença celíaca. Revista de Informática Teórica e Aplicada, v. 17, n. 2, p. 210–220, 2010.

VELA, João Carlos; NASCIMENTO, Beatriz Andrielly de Souza; FEDER, Franciane; ISTVANDIC, Mikaella; JAHN, Ana Claudia. Design de interiores: Entre o bacharelado e o tecnológico: uma reflexão. In: *Congresso brasileiro de pesquisa e desenvolvimento em design*. [s.n.], 2016. v. 2, n. 9, p. 1–12. Disponível em:<http://pdf.blucher.com.br.s3-sa-east-1.amazonaws.com/designproceedings/ped2016/0237.pdf>.

WINCKLER, Marco; PIMENTA, Marcelo Soares. Avaliação de usabilidade de sites web. *Escola de Informática da SBC SUL (ERI 2002) ed. Porto Alegre: Sociedade Brasileira de Computação (SBC)*, v. 1, p. 85–137, 2002.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 15/10/2018
Aceitação/Acceptance: 03/12/2018

OpenChatBotUNL: Propuesta de plataforma de ejecución de Agentes Conversacionales

Jose-Luis Granda[1], Luis Chamba-Eras[1], Milton Labanda-Jaramillo[1], Edison Coronel-Romero[1], Rene Guaman-Quinche [2], Carlos Maldonado-Ortega[2]

{jose.l.granda, lachamba, miltonlab, edisoncor, rguaman, carlos.v.maldonado}@unl.edu.ec

[1] Grupo de Investigación en Tecnologías de la Información y Comunicación (GITIC), Carrera de Ingeniería en Sistemas, Facultad de Energía, Universidad Nacional de Loja, Av. Pío Jaramillo Alvarado y Reinaldo Espinosa, EC110110, Loja, Ecuador.

[2] Carrera de Ingeniería en Sistemas, Facultad de Energía, Universidad Nacional de Loja, Av. Pío Jaramillo Alvarado y Reinaldo Espinosa, EC110110, Loja, Ecuador.

Pages: 102–112

Resumen: En este trabajo se propone una arquitectura de ejecución de Agentes Conversacionales (ChatBots) basada en herramientas Open Source, denominado OpenChatBotUNL. La propuesta integra en una arquitectura algunos paquetes de software como el protocolo XMPP, Openfire, martinbigio/chatterbot, Python ChatterBot, como soporte para la infraestructura de comunicación, gestión, monitoreo e inteligencia artificial. OpenChatBotUNL se plantea como una alternativa a las plataformas de implementación y ejecución (SDKs), disponibles bajo la modalidad de Software como Servicio, representando una oportunidad hacia la innovación en la línea de investigación de Procesamiento de Lenguaje Natural en las universidades en Ecuador.

Palabras-clave: Procesamiento de Lenguaje Natural; Inteligencia Artificial; Agente Conversacional; ChatBot; Open Source

OpenChatBotUNL: Proposal for the execution platform of Conversational Agents

Abstract: In this work we propose an architecture of execution of Conversational Agents (ChatBots) based on Open Source tools, called OpenChatBotUNL. The proposal integrates various software package on an architecture, such as the XMPP protocol, Openfire, martinbigio/chatterbot, Python ChatterBot, as support for the communication, management, monitoring and artificial intelligence infrastructure. OpenChatBotUNL is proposed as an alternative to implementation and execution platforms (SDKs), available under the Software as a Service modality, representing an opportunity for innovation in the Natural Language Processing research line at universities in Ecuador.

Keywords*: N*atural Language Processing; Artificial Intelligence; Conversational Agent; ChatBot; Open Source

1. Introducción

Para crear un ChatBot, actualmente es necesario disponer de una cuenta en alguna red social y habilitar la aplicación correspondiente. Si se requiere algo más avanzado, basta un par de horas, en algunos de los SDK (Software Development Kit) disponibles en las plataformas de Amazon, Google, IBM, Facebook, Slack, Twitter, Whatsapp, WeChat (William Vorhies, 2017).

Fabricantes de SDK como IBM o Google, ofrecen versiones gratuitas, basadas en la nube, destinadas a reducidas cargas de trabajo. Una vez superados ciertos límites, se debe migrar a versiones empresariales (IBM, 2017). Adicional a la limitante de carga, los SDK ocultan la complejidad del software para materializar un agente conversacional (ChatBot). Así, los SDKs representan una caja negra, desde el punto de vista del investigador que pretenda agregar valor o innovar en la línea del procesamiento del lenguaje natural (PLN).

En países, en vías de desarrollo, como el Ecuador, la inversión en innovación, la adquisición, absorción, modificación y creación de conocimiento tecnológico y no tecnológico, son actividades indispensables para el desarrollo económico (Schwartz & Guaipatín, 2014), en esta misma línea, el impulso a la ciencia abierta trae consigo prácticas de innovación que pueden detonar en el desarrollo de los países, especialmente si se destierran los mitos que por diversos motivos se vienen asociando al movimiento del conocimiento en abierto (Ramírez-Montoya & García-Peñalvo, 2018).

En la industria del software, la filosofía de software Open Source, presenta una oportunidad para adquirir, absorber, modificar y construir tecnología. En este trabajo, se propone una plataforma de ejecución de agentes conversacionales, *OpenChatBotUNL* construida mediante la integración de software Open Source. La plataforma tiene por objetivo transparentar el acceso a la complejidad, de la caja negra, de los SDK de los grandes fabricantes de software, al mismo tiempo de empoderar al investigador académico en estas tecnologías con miras a futuras innovaciones.

La organización del trabajo es la siguiente: la sección 2, hace una revisión de literatura respecto a los agentes conversacionales y sus ambientes de ejecución. Así también, se describe software Open Source de soporte a funcionalidades de mensajería instantánea como candidatos a integrar la plataforma propuesta. La sección 3, plantea el diseño de arquitectura de la plataforma *OpenChatBootUNL* y ejemplos de algunas funcionalidades esperadas. Finalmente, en la sección 4, se mencionan las conclusiones y posibles líneas futuras.

2. Revisión de literatura

Un agente conversacional, a simple vista, se puede percibir como un conjunto de interfaces de usuario que nos permite enviar y/o recibir un mensaje, a través de un mecanismo de entrada/salida de cómputo, con cierta “inteligencia”.

Mas ¿qué hay debajo de las interfaces? ¿Es tan sencillo, implementar un agente conversacional, conforme a la promesa de los gigantes del software? Para responder estas preguntas y plantear una alternativa a la ejecución de los agentes conversacionales,

es importante compendiar el estado actual de los conceptos y desarrollos de software en esta línea.

2.1. Agentes conversacionales (ChatBots).

Los agentes conversacionales (ChatBots), son tecnologías que se usan para conseguir una experiencia de usuario más amigable (Kerlyl, Hall, & Bull, 2006). Desde la acuñación del término Inteligencia Artificial (IA), en los años 50, por Jhon Macarthy; y el planteamiento del test para etiquetar como "inteligencia" a una máquina, propuesto por Alan Turing (Russell & Norvig, 2004), una de las primeras aplicaciones, en la línea de conversaciones inteligentes, capaz de entablar una conversación básica, fue ELIZA en 1965 (Russell & Norvig, 1995). Posteriormente ALICE, extendió la base de conocimiento de ELIZA e introdujo el Lenguaje de Marcas de IA, AIML (AI Markup Language). Actualmente, extensiones de este lenguaje soportan consideraciones de contexto, ramificaciones condicionales y aprendizaje supervisado para producir nuevas respuestas (Kerlyl et al., 2006).

Otros ejemplos de agentes conversacionales son Microsoft's Clippy de MS Office97; algunos más recientes y populares como Alexa, Cortana, Siri, and Google Assistant (William Vorhies, 2017). Estos últimos emplean técnicas de aprendizaje automático, para monitorear y cambiar dinámicamente los patrones (pregunta, respuesta) y así mejorar la experiencia del usuario.

En (William Vorhies, 2017) se hace referencia al estudio de 2017 de la empresa Gartner, donde se concluye "que todavía estamos en una curva ascendente para los ChatBots. Sin embargo, todos los principales proveedores de redes o redes sociales, incluidos Amazon, Google, IBM, Facebook, Microsoft, Slack, Twitter, Whatsapp, WeChat y una gran cantidad de independientes, ofrecen SDK fáciles de usar para desarrollarlos. El campo está evolucionando rápidamente tanto en términos de tecnología como de adopción". También se mencionan los campos de aplicación, en el contexto empresarial: servicio al cliente, ventas/mercadotecnia, procesamiento de pedidos, redes sociales, pagos y contratación. En el ámbito educativo: tutoría, preguntas y respuestas, prácticas de conversación en estudio de idiomas, agentes pedagógicos y acompañamiento de aprendizaje, diálogos para fortalecer las habilidades cognitivas y de reflexión, entre otros. En general, estos agentes son capaces de ofrecer apoyo personalizado para cada individuo, reconocer y construir sobre las fortalezas, los intereses y habilidades de las personas con el fin de fomentar la participación y aprendizajes independientes (Kerry, Ellis, & Bull, 2009).

2.2. Software Open Source para Mensajería Instantánea.

La mensajería instantánea representa un conjunto de interfaces para interactuar con los agentes conversacionales. Aplicaciones como WhatsApp, Facebook Messenger, Hangouts son parte de la vida cotidiana de los "conectados"[1]. Usar estas aplicaciones resulta trivial, gracias al fuerte soporte de infraestructura software, como: protocolos de comunicación, servidores de colaboración en tiempo real, panel de administración,

[1] Los usuarios de Internet y dispositivos móviles

motores de PLN, entre otros. A continuación, la descripción de algunas opciones de software de soporte, disponible en modalidad Open Source:

2.3. XMPP

Son las siglas para el protocolo extensible de mensajería y presencia. El protocolo está basado en XML y permite el intercambio casi en tiempo real de datos estructurados pero extensibles entre dos o más entidades de red (Saint-Andre, 2011).

El protocolo central de XMPP soporta: instalación y desmontaje de flujos XML, cifrado de canal, autenticación, manejo de errores y comunicación primitivas para mensajería, disponibilidad de red ("presencia"), interacciones solicitud-respuesta (Saint-Andre, 2011). Mayor detalle de la especificación de XMPP, se desarrolla en (Xuefu & Ming, 2012).

2.4. Openfire

Es un servidor de colaboración en tiempo real (RTC), con licencia de Open Source Apache License, que utiliza como protocolo de mensajería XMPP (Realtime, Ignite, 2008).

Openfire representa una alternativa al dominio de las grandes compañías de software en el área de Mensajería Instantánea Empresarial (MIE). En (Xuefu & Ming, 2012) se hace notar la diferencia entre Mensajería Instantánea (MI) y la Mensajería instantánea Empresarial, estos últimos implican mayores requerimientos en cuanto a funcionalidades extras como envió/recepción de archivos, vídeo conferencia, entre otros. En la presente propuesta nos limitaremos a la MI como interface del agente conversacional.

2.5. Aplicación martinbigio/chatterbot[2]

Es un agente inteligente simple capaz de chatear con los clientes de un centro de llamadas y cuyo comportamiento se puede monitorear y personalizar a través de la interfaz Web. El servicio de mensajería instantánea esta implementada sobre OpenFire.

Para administrar el sistema, se cuenta con una aplicación Web como panel de control. En ella se pueden administrar los usuarios del sistema, monitorear las conversaciones de los agentes de chat (chatterbots) en tiempo real, consultar los registros y estadísticas del sistema y configurar la semántica del lenguaje que deben interpretar los chatterbots (Bigio, 2012/2018).

2.6. Python ChatterBot

Es una biblioteca de Python que facilita el generar respuestas automáticas a la entrada de un usuario. Python ChatterBot utiliza una selección de algoritmos de aprendizaje automático para producir diferentes tipos de respuestas. Permite crear agentes conversacionales (ChatBots) y automatizar las conversaciones con los usuarios.

Una instancia no entrenada de agente conversacional, comienza sin conocimiento de cómo comunicarse. Cada vez que un usuario ingresa una declaración, la biblioteca guarda el texto ingresado y el texto al que la declaración respondió. A medida que ChatterBot

[2] Se usa el nombre de usuario GitHub y el repositorio de la aplicación, para diferenciarla de Python ChatterBot también descrito y usado posteriormente

recibe más información, la cantidad de respuestas que puede responder y la precisión de cada respuesta en relación con la declaración de entrada aumentan[3]. El diagrama de flujo de procesos se detalla en la figura 1.

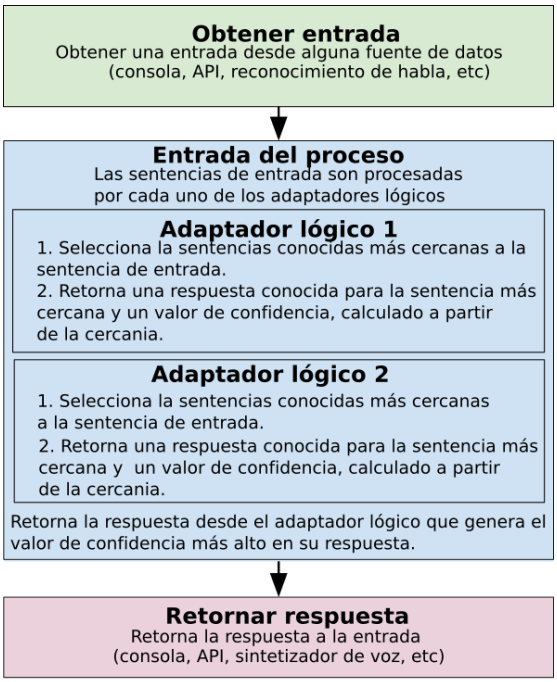


Figura 1 – Diagrama de flujo de procesos Python ChatterBot

2.7. Plataformas de ejecución de agentes conversacionales

Como se mencionó, los principales proveedores de redes o redes sociales y servicios de software, proveen plataformas sobre las cuales se puede implementar agentes conversacionales. El denominador común de estas plataformas es la modalidad de negocio, software como servicio o, en este contexto, ChatBot como servicio. Cabe recalcar que en su mayoría disponen de una versión de libre acceso, limitada en la cantidad de mensajes en las conversaciones.

En (Rahman, Al Mamun, & Islam, 2017) se hace una clasificación de las plataformas antes mencionadas, así: chats sin programación, chats orientados a conversaciones y plataformas de chat proveídas por los "gigantes"[4].

[3] About ChatterBot — ChatterBot 0.8.7 documentation. (s. f.). Recuperado 26 de julio de 2018, de https://chatterbot.readthedocs.io/en/stable/index.html
[4] Se refiere a los gigantes de la industria del software, como: Google, Facebook, IBM, entre otros

3. Especificación de la arquitectura

3.1. Arquitectura

OpenChatBotUNL plataforma de ejecución de agentes conversacionales, se concibe como el resultado de la integración de herramientas de software disponibles bajo modalidad Open Source. La plataforma se basa en la arquitectura para MIE propuesta en (Xuefu & Ming, 2012). Esta arquitectura integra servicios basados en XMPP, clientes para protocolo XMPP e integración con otros servicios XMPP. La figura 2 muestra la propuesta de la arquitectura *OpenChatBotUNL*.

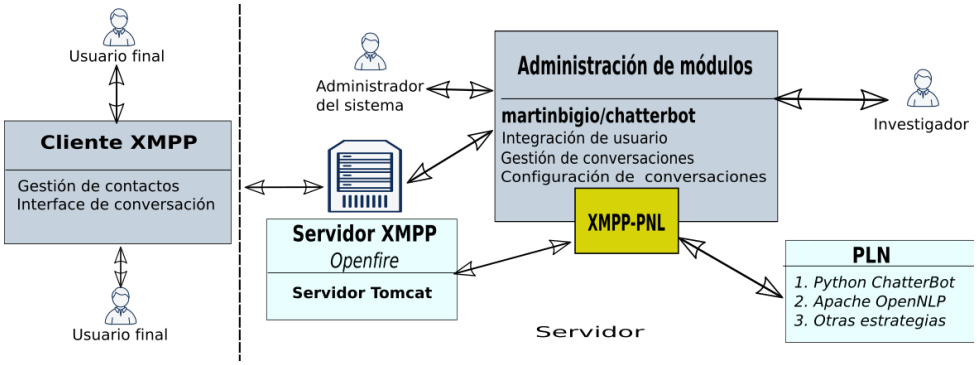


Figura 2 – Arquitectura de *OpenChatBotUNL*.

El componente ***XMPP-PLN*** permitirá la conversación entre XMPP y algún modelo de PLN, en la primera implementación de la plataforma se propone *Python Chatterbot* como motor de PLN, adicionalmente *martinbigio/chatterbot* implementa PLN con la librería Java *Apache OpenNLP*[5], así se amplían las posibilidades de la plataforma. El componente XMPP-PLN ha de implementarse empleando tecnología *"enchufar-usar"* (plugin), así será posible usar otras tecnologías como R, Octave u otras.

Se consideran 3 tipos de usuarios: el usuario final quien usa el chat, el administrador quien gestiona las conversaciones, la información y datos resultantes; y finalmente el investigador quien crea el o los agentes conversacionales y analiza los resultados. La identificación de estos roles de usuario, también determina las herramientas a proveer a cada uno de ellos. El usuario final, en producción, puede elegir entre la gran variedad de clientes XMPP[6] (Ej.: pidgin, converse, entre otros) y adicionalmente, la plataforma debe implementar un cliente basado en Web para efectos de pruebas. El administrador cumple un rol de orquestador y sus herramientas son de naturaleza más técnica a fin de dar soporte tanto a los usuarios y/o investigadores. El investigador por su parte debe poder desplegar sus agentes conversacionales y/o ajustarlos en tiempo de ejecución, debe poder acceder a datos estadísticos de ejecución y al contenido de las conversaciones, a fin de poder hacer nuevas experimentaciones sobre sus modelos de agentes conversacionales u otros en la línea de PLN.

[5] Apache OpenNLP. Recuperado 08/11/2018. https://opennlp.apache.org/
[6] XMPP Clientes. Recuperado 08/11/2018. https://xmpp.org/software/clients.html

3.2. Vistas previas

En las figuras 3, 4 y 5, se observa los primeros prototipos basados en la plataforma *OpenChatBotUNL*.

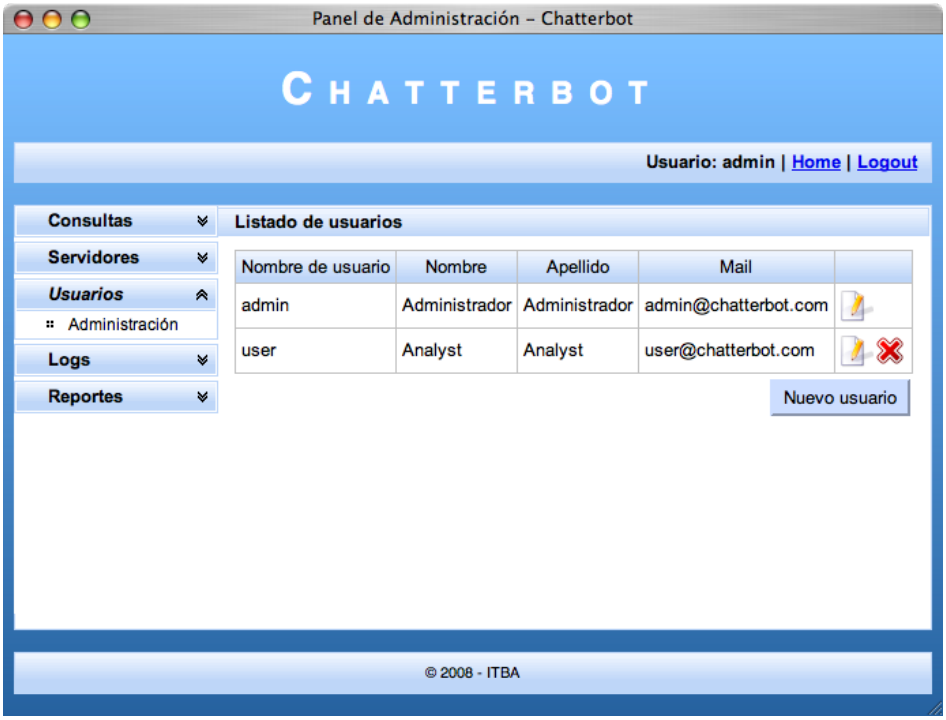


Figura 3 – Aplicación *martinbigio/chatterbot: Panel de Administración*

4. Conclusiones y líneas futuras

En los trabajos (Kerlyl et al., 2006), (Kerry et al., 2009; García et al., 2009; Lima, Algredo-Badillo, Moreno, Bautista, & Rosales, s. f.; Xu, Liu, Guo, Sinha, & Akkiraju, 2017; Acedo & Ramírez, 2017; Cambizaca & Francisco, 2018) se describen los detalles de implementación de agentes conversacionales en diferentes ámbitos de aplicación. Un factor común es la diversidad de plataformas de ejecución.

Disponer de una plataforma de ejecución de Agentes Conversacionales (ChatBots) implementada sobre software Open Source, proveída desde la academia, representa una oportunidad para propender a la innovación transparentando la complejidad de los componentes de la plataforma de ejecución. Este conocimiento puede ser un motor para la innovación, la resolución de problemas y el planteamiento creativo de posibilidades (Ramírez-Montoya & García-Peñalvo, 2018).

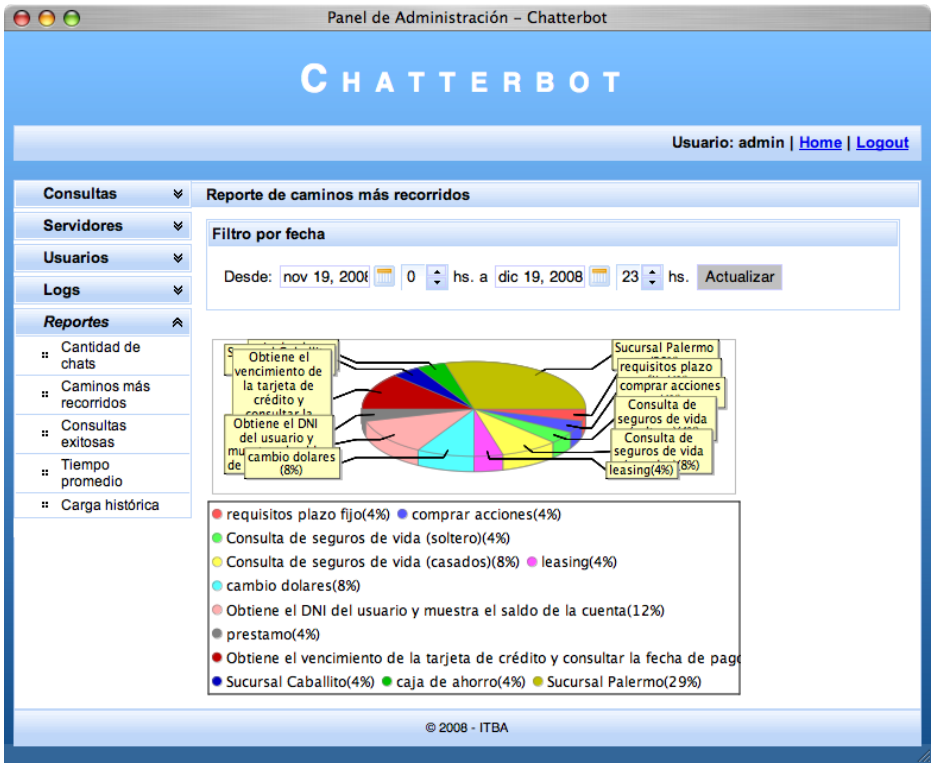


Figura 4– Aplicación *martinbigio/chatterbot: Reportes y estadística*

En concepto la plataforma permitirá realizar experimentación sobre modelos de conversaciones a través de ChatBots, pese a que la plataforma no está integrada aún, el software *martinbigio/chatterbot* puede usarse de inmediato como versión beta, de estas interacciones se obtendrán datos e ideas para evolucionarla.

Un debate que empieza a tomar fuerza, es la conveniencia o no de entregar datos (en el contexto de IA: inteligencia) a los grandes fabricantes de software y sus herramientas de IA. La plataforma propuesta, permitiría evaluar si es posible o no alejarse de los grandes fabricantes.

No es descartable que, como fruto de la investigación, la plataforma pueda ser llevada a entornos de producción empresarial.

Agradecimientos

Este trabajo forma parte de las actividades académicas y de investigación del Grupo de Investigación GITIC, agradecemos al profesorado, autoridades y alumnado de la Carrera de Ingeniería en Sistemas de la Universidad Nacional de Loja (UNL).

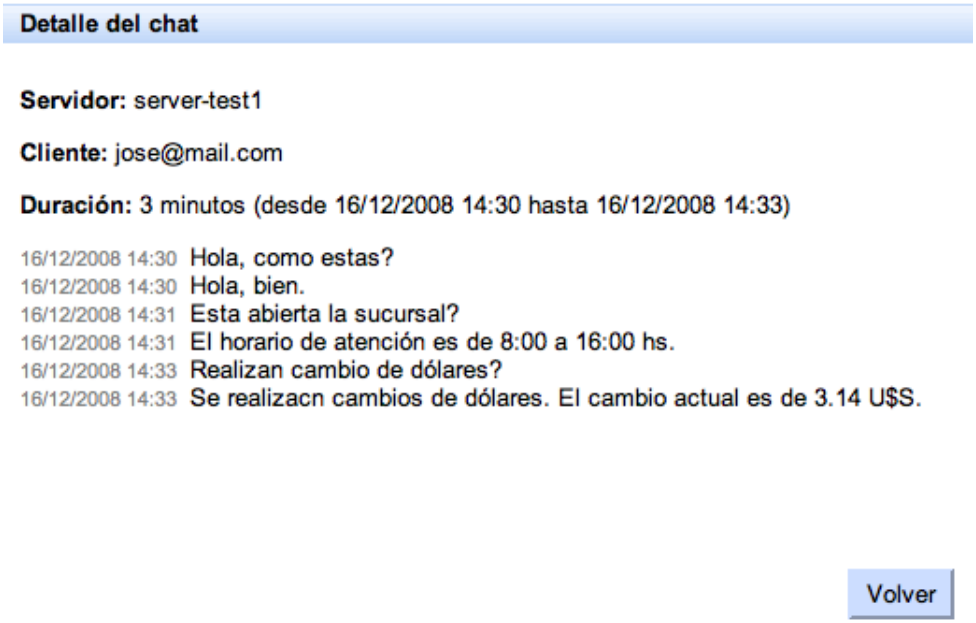


Figura 5 – Aplicación *martinbigio/chatterbot: Una conversación de chat*

Contribuciones de autoría

Jose-Luis Granda y Luis Chamba-Eras y concibieron la propuesta, la arquitectura, las vistas. Así mismo, Milton Labanda-Jaramillo, Edison Coronel-Romero, Rene Guaman-Quinche y Carlos Maldonado-Ortega, revisaron la redacción del texto y aportaron comentarios en la revisión de literatura sobre el objeto de estudio, la especificación de la arquitectura y otros elementos para la mejora el artículo y aprobaron la versión final para su publicación.

Referencias

Acedo, E. G., & Ramírez, P. S. (2017). BotMentor: Bot de ayuda al estudiante en la plataforma Telegram.

Bigio, M. (2018). *chatterbot: A simple intelligent agent for a call center, written in Java (final project at itba.edu.ar)*. Java. Recuperado de https://github.com/martinbigio/chatterbot (Original work published 2012)

Cambizaca, T., & Francisco, A. (2018). Desarrollo de un chatbot que ayude a responder a preguntas frecuentes referentes a becas en la Universidad Técnica Particular de Loja. Recuperado de http://dspace.utpl.edu.ec/handle/20.500.11962/21874

Extensible Messaging and Presence Protocol. (2018). En *Wikipedia, la enciclopedia libre*. Recuperado de https://es.wikipedia.org/w/index.php?title=Extensible_Messaging_and_Presence_Protocol&oldid=108402227

García, E. G., López, R. J., Moreno, J. J. M., Abad, A. S., Blasco, B. C., & Pol, A. P. (2009). La metodología del Data Mining. Una aplicación al consumo de alcohol en adolescentes. *Adicciones*, *21*(1), 65–80.

Realtime, Ignite. (2008). Ignite Realtime: Openfire Server. Recuperado 25 de julio de 2018, de https://www.igniterealtime.org/projects/openfire/index.jsp

Kerlyl, A., Hall, P., & Bull, S. (2006). Bringing Chatbots into education: Towards Natural Language Negotiation of Open Learner Models. En *Applications and Innovations in Intelligent Systems XIV* (pp. 179-192). Springer, London. https://doi.org/10.1007/978-1-84628-666-7_14

Kerry, A., Ellis, R., & Bull, S. (2009). Conversational agents in E-Learning. En *Applications and innovations in intelligent systems XVI* (pp. 169–182). Springer.

Lima, F. T., Algredo-Badillo, I., Moreno, N. T., Bautista, D. P., & Rosales, L. A. M. (s. f.). Recorrido Virtual y Agente Conversacional Usando Photosynth y Python para un Sitio Web Publicitario.

Open Source Initiative OSI - Jabber Open Source License:Licensing | Open Source Initiative. (s. f.). Recuperado 25 de julio de 2018, de https://opensource.org/licenses/jabberpl.php

Rahman, A. M., Al Mamun, A., & Islam, A. (2017). Programming challenges of chatbot: Current and future prospective. En *Humanitarian Technology Conference (R10-HTC), 2017 IEEE Region 10* (pp. 75–78). IEEE.

Ramírez-Montoya, M. S., & García-Peñalvo, F. J. (2018). Co-creación e innovación abierta: Revisión sistemática de literatura. Recuperado de https://repositorio.grial.eu/handle/grial/1048

Russell, S. J., & Norvig, P. (1995). *Artificial intelligence: a modern approach*. Malaysia; Pearson Education Limited,.

Russell, S. J., & Norvig, P. (2004). *Inteligencia Artificial: un enfoque moderno*.

Saint-Andre, P. (2011). *Extensible messaging and presence protocol (XMPP): Core*.

Wallace, R. S. (2009). The anatomy of ALICE. En *Parsing the Turing Test* (pp. 181–210). Springer.

William Vorhies. (2017, noviembre 7). Beginners Guide to Chatbots. Recuperado 10 de julio de 2018, de https://www.datasciencecentral.com/profiles/blogs/beginners-guide-to-chatbots

XMPP | The XMPP Blog. (s. f.). Recuperado 24 de julio de 2018, de https://xmpp.org/2015/03/no-its-not-the-end-of-xmpp-for-google-talk/

Xu, A., Liu, Z., Guo, Y., Sinha, V., & Akkiraju, R. (2017). A New Chatbot for Customer Service on Social Media. En *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 3506–3510). New York, NY, USA: ACM. https://doi.org/10.1145/3025453.3025496

Xuefu, B., & Ming, Y. (2012). Design and implementation of web Instant Message System based on XMPP. En *2012 IEEE International Conference on Computer Science and Automation Engineering* (pp. 83-88). https://doi.org/10.1109/ICSESS.2012.6269411

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 09/09/2018
Aceitação/Acceptance: 19/11/2018

Análisis comparativo de predicción dentro de bases de datos de cáncer: una aplicación de aprendizaje automático.

Gabriel MauricioMartínez-Toro[1], Dewar Rico-Bautista[2], Efrén Romero-Riaño[3]

gmartinez714@unab.edu.co, dwricob@ufpso.edu.co,efren.romero@saber.uis.edu.co

[1]Universidad Autónoma de Bucaramanga, Cra 34#41-51, 680003, Bucaramanga, Colombia.

[2]Departamento Sistemas e Informática. Universidad Francisco de Paula Santander Ocaña, Sede Algodonal Vía Acolsure, 546551, Ocaña, Colombia.

[3]Escuela de Estudios Industriales y Empresariales. Universidad Industrial de Santander, 680002, Bucaramanga, Colombia.

Pages: 113–122

Resumen: El cáncer no es solo una enfermedad, es un conjunto de enfermedades de gran impacto en la salud pública. En ese sentido, los esfuerzos por consolidar métodos de análisis basados en grandes datos que contribuyan a su predicción, constituye un área de especial interés para los científicos y analistas de datos. El objetivo de este artículo, es comparar el rendimiento de la predicción delos métodos: i) Logisticregression, ii) K NearestNeighbor, iii) K-means, iv) Random Forest, v) Support Vector Machine, vi) Análisis de Discriminantes Lineales, vii) Gaussian Naive Bayes viii) MultilayerPerceptron dentro de una base de datos de cáncer.Para el caso de los modelos de aprendizaje no supervisado, se evidencia la relevancia de los centroides para el caso del algoritmo k means, así como las asignaciones de la tasa de aprendizaje y los parámetros para el caso de MultilayerPerceptron. Para el caso de los modelos de aprendizaje supervisado, el de mejor desempeño es el de SVM.

Palabras-clave: big data, aprendizaje de máquina, predicción de cáncer.

Comparative analysis of prediction within cancer databases: a machine learning application.

Abstract:Cancer is not only a disease, it is a set of diseases with a great impact on public health. In that sense, efforts to consolidate methods of analysis based on large data that contribute to its prediction, is an area of special interest for scientists and data analysts. The objective of this paper is to compare the performance of method prediction: i) Logistic regression, ii) K Nearest Neighbor, iii) K-means, iv) Random Forest, v) Support Vector Machine, vi) Linear Discriminant Analysis, vii) Gaussian Naive Bayes viii) Multilayer Perceptron, within a cancer database. In the case of unsupervised learning models, the relevance of the centroids for the k means algorithm is evident, as well as the learning rate assignments and parameters for

the Multilayer Perceptron case. In the case of supervised learning models, SVM performs best.

***Keywords**:* big data, machine learning, cancer prediction

1. Introducción

El cáncer no es solo una enfermedad, es un conjunto de enfermedades. Siendo el cáncer de mama el segundo cáncer más común en todo el mundo después del cáncer de pulmón, es la causa más frecuente de muerte por cáncer en mujeres. Si se diagnostica a una edad temprana las posibilidades de supervivencia son mayores.

Los datos utilizados, corresponden a una muestra obtenida de 11 tumores, relacionados con cáncer de seno. El cáncer es la principal causa de muerte en países desarrollados y la segunda en países en desarrollo. Dentro de las categorías de cáncer, el cáncer de seno es el de mayor frecuencia de diagnóstico y la primera causa de muerte en mujeres(Jemal, Bray, &Ferlay, 1999). La importancia de los métodos y modelos de aprendizaje supervisado y no supervisado, para identificar patrones y características en la clasificación de cáncer, se evidencia en el avance en el desempeño de las clasificaciones basadas en los mismos. El objetivo de este paper es mostrar la comparación del desempeño de clasificación de cáncer de nueve modelos, enmarcados dentro de aprendizaje supervisado (Machine Learning) y no supervisado (Deep Learning).

El reconocimiento de características de "serumproteomics", es ampliamente utilizado para la detección de cáncer de ovario (Baggerly et al., 2004), de seno (Pablo Guillen & Guillen, 2017), entre otros. Múltiples estudios, han utilizado selección de genes para la clasificación de cáncer usando: support vector machines(Guyon, Isabelle, w, Eston& Stephen Barnhill, 2002), randomforest (Statnikov, Wang, &Aliferis, 2008), redes bayesianas (Griffis, Allendorfer, &Szaflarski, 2016; Velikova, Lucas, Samulski, &Karssemeijer, 2013), Deep Learning (Chen, T., Chefd'hotel, 2014) y comparando el rendimiento entre pares (Dreiseitl& Ohno-Machado, 2002; guyon, isabelle, w, eston&stephenbarnhill, 2002) de métodos, sin embargo no se evidencia una comparación exhaustiva de este tipo, usando nueve modelos.

Machine Learning fue definido por Arthur Samuel como el campo de estudio que da a los computadores, la habilidad de aprender sin ser explícitamente programados (Martens, 1959). Deep Learning es originado de la investigación de redes neurales artificiales(Pablo Guillen & Guillen, 2017), (Chen, 1966). El aprendizaje profundo, descubre una estructura compleja en grandes conjuntos de datos al usar el algoritmo de retro propagación, para indicar cómo una máquina debe cambiar sus parámetros internos que se utilizan para calcular la representación en cada capa, a partir de la representación en la capa anterior(Lecun, Bengio, & Hinton, 2015). Se generaron matriz de confusión, la cual se puede utilizar para medir el rendimiento de un algoritmo de aprendizaje de máquina, generalmente uno de aprendizaje supervisado. Cada fila de la matriz de confusión representa las instancias de una clase real y cada columna representa las instancias de una clase pronosticada. Dentro de las tareas de clasificación, la precisión de una clase, es el número de verdaderos positivos. En ese sentido, se presenta la interpretación de la matriz de confusión del ejercicio uno, la cual muestra tres grupos de datos, por fuera de la diagonal principal de la matriz de confusión. De manera adicional se identifica, dos clases predichas con valores altos de presión, entre siete y ocho.

Este artículo se encuentra dividido principalmente en (i)Introducción, (ii) Materiales y métodos, (iii) resultados y (iv) por último las conclusiones.

2. Materiales y métodos

2.1. Aprendizaje automático.

Existen diversos marcos de trabajo conceptuales para aprendizaje automático. Aquí se resume algunos aspectos asociados a las arquitecturas usadas. A continuación, se describen brevemente los algoritmos de aprendizaje de máquina:

La regresión logística es una técnica de clasificación bien establecida que se utiliza ampliamente en estudios epidemiológicos. Por lo general, es que se utiliza como referencia, en comparación con otras técnicas de analizando datos médicos. La regresión logística también se conoce como regresión logit, clasificación de máxima entropía (MaxEnt) o el clasificador log-lineal (Dreiseitl& Ohno-Machado, 2002).

El análisis discriminante lineal es un clasificador lineal. Estos clasificadores son atractivos porque tienen soluciones de forma cerrada, que se pueden calcular fácilmente. Se puede usar para realizar la reducción de dimensionalidad supervisada, proyectando los datos de entrada a un subespacio lineal que consiste en las direcciones que maximizan la separación entre clases(Dudoit, Fridlyand, &Speed, 2002).

K-nearestneighbor (aprendizaje basado en la instancia) es un clásico Algoritmo de aprendizaje basado en instancias en el que se clasifica un nuevo caso basado en la clase conocida del vecino más cercano, por medio de una mayoría de muestras. El principio detrás de los métodos de “K NearestNeighborn” es encontrar un número predefinido de muestras de entrenamiento más cercanas en distancia al nuevo punto, y predecir la etiqueta de estos. El número de muestras puede ser una constante definida por el usuario (aprendizaje k-vecino más cercano) o variar según la densidad local de puntos (aprendizaje vecino basado en radio). Este tipo de algoritmo también se llama aprendizaje perezoso porque no hay ningún paso de construcción de modelos y toda la informática procedimiento (es decir, la búsqueda del vecino más cercano) se realiza directamente durante la predicción(Dreiseitl& Ohno-Machado, 2002).

Los métodos de Naive Bayes, son un conjunto de algoritmos de aprendizaje supervisado basados en la aplicación del teorema de Bayes con la suposición de independencia, entre cada par de características. El algoritmo funciona bien con tipos de datos heterogéneos y también con valores perdidos, debido al tratamiento independiente de cada variable predictiva para la construcción del modelo (Griffis et al., 2016).

Las máquinas de soporte vectorial (SVM), son un conjunto de métodos de aprendizaje supervisados utilizados para la detección de clasificación, regresión entre otras. Las principales ventajas de SVM son: i) efectivos en espacios de alta dimensión, ii) siguen siendo efectivo en casos donde el número de dimensiones es mayor que el número de muestras. y ii) son eficientes desde el punto de vista de la memoria(Chang & Lin, 2011).

El perceptrón/backpropagation multicapa, es el principal material artificial de la red neuronal. Cuando no hay ninguna capa oculta en la red, este algoritmo es equivalente a la regresión logística, pero puede resolver problemas más difíciles con una red más compleja

arquitecturas. El precio de usar arquitecturas complejas es que produce modelos que son más difíciles de interpretar. Además, puede ser computacionalmente caro.

Randomforest es un algoritmo de aprendizaje por máquina del método "familia de algoritmos "ensemble", que crea múltiples modelos (llamados estudiantes débiles) y los combina para tomar una decisión, para aumentar la precisión de la predicción. La idea principal de esto es construir un "bosque" de "árboles" de decisión aleatoria y utilizar para clasificar un nuevo caso. Cada árbol se genera usando un árbol aleatorio de las variables predictoras del candidato y un subconjunto de variables aleatorias. subconjunto de datos, generados mediante bootstrap. Este algoritmo también se puede utilizar para estimar la relevancia variable.

2.2. Aprendizaje profundo.

El algoritmo de K-Means agrupa los datos tratando de separar las muestras en n grupos de igual varianza, minimizando un criterio conocido como inercia o suma de cuadrados dentro del grupo. Este algoritmo requiere que se especifique la cantidad de clústeres. Se adapta bien a una gran cantidad de muestras y se ha utilizado en una amplia gama de áreas de aplicación en muchos campos diferentes(Tsai, 2014).

Los arboles de decisión (RF), son un método de aprendizaje supervisado, no paramétrico utilizado para la clasificación y la regresión. El objetivo es crear un modelo que prediga el valor de una variable objetivo mediante el aprendizaje de reglas simples de decisión inferidas a partir de las características de los datos (Kurzynski, Krysmann, Trajdos, &Wolczowski, 2016).

2.3. Datos y análisis.

El presente paper se basa en el análisis de la base de datos de once tumores de cáncer de seno,procedentes de los datos de diagnóstico de la universidad de Wisconsin (BCW). El proceso de preparación para el procesamiento incluye ajuste de los datos dentro de las plataformas Anaconda Spyder (aprendizaje supervisado) y H2O (aprendizaje no supervisado).

Las medidas de comparación del desempeño de los pronósticos de cada modelo, incluyen descriptores como la media y la desviación estándar de los resultados de predicción dentro de los modelos de aprendizaje supervisado. De manera adicional se establece una medida de la eficiencia de cada uno de los modelos.

3. Resultados

Este apartado se presenta en dos componentes: i) el análisis de desempeño de modelos de aprendizaje supervisado con base en cinco modelos y ii) desempeño de modelos no supervisado.

3.1. Aprendizaje supervisado

Tomando el documento "CancerClassificationBasedonMicroarray Gene Expression Data Using Deep Learning"(P Guillen &Ebalunode, 2016), el objetivo es realizar matriz

de confusión para resumir el rendimiento de un algoritmo de clasificación, a la data suministrada para tal efecto, e identificar como salida de este proceso.(Olivera et al., 2017).

Los cinco modelos aplicados a dicha data son: LogisticRegression, models.append((‘LDA’, LinearDiscriminantAnalysis, models.append ‘KNN’, KNeighborsClassifier, models. append ‘NB’, GaussianNB y el models.append. (Berg, Hjelmervik, Henrik, Stender, & S, 2016)

En referencia al aprendizaje supervisado, se muestra es resultado de la implementación de cinco modelos: Regresión Logística (LR), Análisis de Discrimantes Lineales (LDA), K-Vecinos más Cercanos (KNN), Gaussian Naive Bayes (NB), Máquinas de Soporte Vectorial (SVM). Estos modelos son optimizados con base a la estrategia de uso del 80% de los datos para entrenamiento y el 20% de los datos para validación. Los modelos son evaluados mediante la técnica de evaluación “accuracy” por medio de la media y la desviación estándar. Se presenta la figura 1, diagrama de caja y bigote, con el fin de visualizar el desempeño de los modelos.

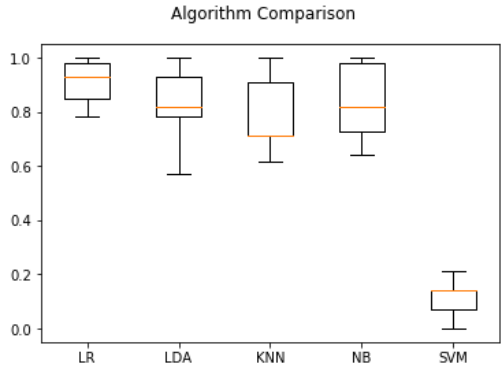


Figura 1 – Comparación del desempeño de cinco modelos de aprendizaje supervisado.
Fuente: elaboración propia a partir de Anaconda Spyder

La tabla 1, muestra la comparación de resultados de desempeño de los cinco modelos supervisados. Se identifica que el mejor desempeño se obtiene con base en regresión logística, el cual presenta una media de 0,9 y desviación estándar de 0,079. El desempeño evaluado por el modelo muestra una eficiencia del 100%, resultado que ratifica la alta relevancia y pertinencia del modelo para el análisis de los datos suministrados.

Modelo	Media	Desviación estandard
LR	0.906044	0.079336
LDA	0.834066	0,116214
KNN	0,782967	0,127118
NB	0,841209	0,127810
SVM	0,931800	0,066529

Fuente: elaboración propia a partir de Anaconda Spyde

Tabla 1 – Comparación desempeño modelos supervisados.

Calcular una matriz de confusión puede darle una mejor idea de lo que su modelo de clasificación está haciendo bien y qué tipos de errores está cometiendo. La definición de la matriz de confusión es un paso fundamental para la comprensión y la sustentación de la validez de los análisis basados en máquina de aprendizaje. De forma específica, las matrices de confusión permiten identificar parámetros clave tales como la precisión de los pronósticos. En ese sentido se construye el código para la lectura de las variables y la construcción de la matriz, ver Fig 2.

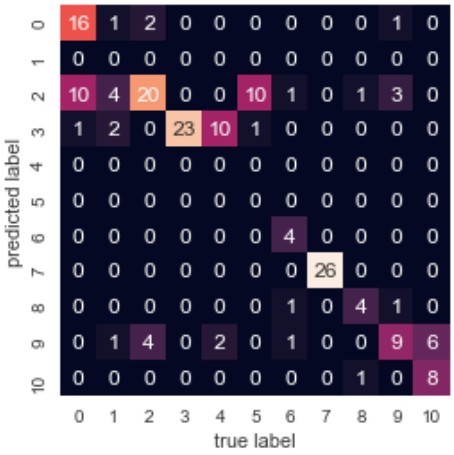


Figura2 – Matriz de confusión.
Fuente: elaboración propia a partir de Anaconda Spyder

3.2. Aprendizaje No-supervisado

Como medida para establecer una relación entre el desempeño de los modelos de aprendizaje supervisado y no supervisado, se realiza la comparación de algoritmos a través del gráfico de caja y bigote que se muestra en la figura 3.

Esta comparación se implementa con base en las condiciones estándar de cada uno de los métodos no supervisados, incluidos los parámetros. Con base a esta condición estándar, se obtiene precisiones de bajas con base al método de k means (0.14) Respecto a los otros algoritmos se obtiene: Random Forest (0.69), Neural Multilayer (0.6011) y DecisionTree (0.658). Con el fin de mejorar el desempeño de la clasificación de cada uno de los modelos, se implementa la estrategia de trabajo individual de cada algoritmo, con base a la identificación de parámetros claves para cada modelo en particular.

Para el caso de la estrategia de optimización de la clasificación de tumores, basado en el algoritmo K means, se incluyen parámetros adicionales a los parámetros que se incluyen por defecto. Los cuatro parámetros incluidos y sus valores son: n_clusters=10,random_state=170,n_jobs=10,tol=0.001. Se direcciona k=10, debido a que existe la información que existen 10 clase dentro de la data. El parámetro n_jobs se prueba desde n= 4, obteniendo precisión del 50 % con tolerancia: 0.001. Finalmente, la precisión estimada es de setenta y un porciento (0.7126436781609196). Esta estrategia genera como resultado un mejoramiento significativo del rendimiento desde 0.14 hasta 0.71.

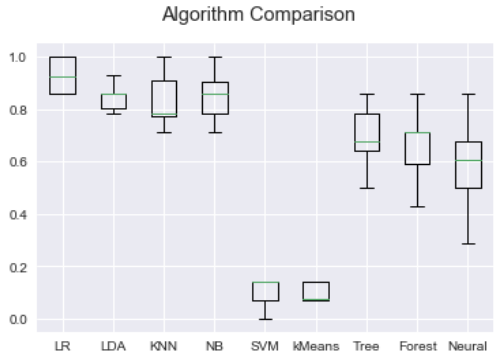


Figura 3 – Comparación del desempeño de cinco modelos de aprendizaje no supervisado.
Fuente: elaboración propia a partir de Anaconda Spyde

Para el caso de Random Forest, resulta determinante la elección de la semilla. Se contrasta con valores de 4 y 170, evidenciando cambios del orden del 50% de mejoramiento de accuracy. El valor estimado con seed= 170, fue de 0,6347. Esta estrategia de mejoramiento no genera mejor resultado que la original, donde el primer valor obtenido de 0.69 es mayor que el valor de 0.63.

Para el caso de Neural MultilayerPerceptron, se implementa un arreglo basado en una partición de la data de 80% para entrenamiento, de 20% de prueba y una semilla de 42. Se implementan dos capas con 50 neuronas cada una. Los parámetros incluyen alfa=0,0001 y una tasa de aprendizaje constante. Como resultado se obtiene una precisión de 94,28. De manera adicional se obtiene un recall= 0,91 y precisión=0,95.

El análisis comparativo de los algoritmos, muestra mejor resultado de desempeño dentro de los métodos no supervisados. El mejor rendimiento dentro de los algoritmos de aprendizaje supervisado es SVM con un valor de 0,931800; y el mejor rendimiento para el caso de aprendizaje no supervisado es Neural MP, con un valor de accuracy de 0.9428.

4. Conclusiones

Se aplicaron cinco métodos de reconocimiento apropiados en clase LogisticRegression, models.append((‘LDA’, LinearDiscriminantAnalysis, models.append ‘KNN’, KNeighborsClassifier, models.append ‘NB’, GaussianNB y el models.append, para el análisis de la data identifica once clases, resultado que demuestra la estructura y las clases de cáncer de los datos de suministrado.

El modelo que muestra mejores resultados, medidos con base a la media y desviación estándar, es LR, el cual presenta media igual 0.906044 y desviación estándar de 0.079336. El desempeño evaluado por el modelo muestra una eficiencia del 100%, resultado que ratifica la alta relevancia y pertinencia del modelo para el análisis de los datos suministrados.

Dentro de las tareas de clasificación, la precisión de una clase, es el número de verdaderos positivos. En ese sentido, se presenta la interpretación de la matriz de confusión del ejercicio uno, la cual muestra tres grupos de datos, por fuera de la diagonal principal de la matriz de confusión. De manera adicional se identifica, dos clases predichas con valores altos de presión, entre siete y ocho.

La tabla 2 resume el rendimiento de los modelos de aprendizaje no supervisado y supervisado, para la clasificación de cáncer. Se identifica la sensibilidad de los diferentes modelos para la clasificación de cáncer, relacionada con diversos parámetros propios de los métodos, así como la sensibilidad a procedimientos ocultos como las fórmulas de estimación de distancias entre puntos, entre otros.

Modelo	Media
KMEANS	0,712643
RF	0,6347
MP	0,9428
SVM	0.9318
LR	0.906044
LDA	0.834066
KNN	0,782967
NB	0.8412

Fuente: elaboración propia a partir de Anaconda Spyde

Tabla 2 – Comparación desempeño modelos no supervisados y supervisados.

El estudio plasmado en el artículo se llevó a cabo para comprobar la precisión del rendimiento de algoritmos de agrupación de k-means sobre el cáncer de mama, usando la base de datos de diagnóstico de Wisconsin (BCW). Se estudiaron los siguientes parámetros (Dubey, Gupta, & Jain, 2016):

- Exactitud de la agrupación en la separación de lo benigno y lo maligno tumores.
- El efecto de las medidas de centroide, distancia y división en k-significa.
- El papel de las iteraciones y la época en la búsqueda del centroide.
- El impacto de los atributos del conjunto de datos BCW en la aplicación de algoritmo de agrupación de medios K.

Se estima el mejor desempeño para el algoritmo de Neural MP. Para el caso de los modelos de aprendizaje no supervisado, se evidencia la relevancia de los centroides para el caso del algoritmo k means, así como las asignaciones de la tasa de aprendizaje y los parámetros para el caso de MultilayerPerceptron. Para el caso de los modelos de aprendizaje supervisado, el de mejor desempeño es el de SVM. Se identifica el carácter determinante del valor de los parámetros de aprendizaje para el caso de support vector machine, SVM y una baja sensibilidad de este al tamaño de la semilla seleccionado. Se evidencia la dificultad de trabajar este tipo de datos con el modelo de K means, así como la facilidad de mejoramiento del desempeño, basados en aprendizaje no supervisado, MultilayerPerceptron.

Referencias

Baggerly, K. A., Edmonson, S. R., Morris, J. S., Coombes, K. R., Liotta, L. A., Petricoin, E. F., … Conrads, T. P. (2004). High-resolution serum proteomic patterns for ovarian cancer detection (multiple letters). Endocrine-Related Cancer, 11(4), 583–587. https://doi.org/10.1677/erc.1.00868

Berg, H., Hjelmervik, K. T., Henrik, D., Stender, S., & S, T. S. (2016). A Comparison of Different Machine Learning Algorithms for Automatic Classification of Sonar Targets. Oceans 2016 MtsIeee, 1–8.

Chang, C.-C., & Lin, C.-J. (2011). Libsvm. ACM Transactions on Intelligent Systems and Technology, 2(3), 1–27. https://doi.org/10.1145/1961189.1961199

Chen, T., Chefd'hotel, C. (2014). Deep learning based automatic immune cell detection for immunohistochemistry images. In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics (p. 17–24.).

Chen, C. H. (1966). A note on sequential decision approach to pattern recognition and machine learning. Information and Control, 9(6), 549–562. https://doi.org/10.1016/S0019-9958(66)80015-3

Dreiseitl, S., & Ohno-Machado, L. (2002). Logistic regression and artificial neural network classification models: A methodology review. Journal of Biomedical Informatics, 35(5–6), 352–359. https://doi.org/10.1016/S1532-0464(03)00034-0

Dubey, A. K., Gupta, U., & Jain, S. (2016). Analysis of k-means clustering approach on the breast cancer Wisconsin dataset. International Journal of Computer Assisted Radiology and Surgery, 11(11), 2033–2047. https://doi.org/10.1007/s11548-016-1437-9

Dudoit, S., Fridlyand, J., & Speed, T. P. (2002). Comparison of discrimination methods for the classification of tumors using gene expression data. Journal of the American Statistical Association, 97(457), 77–86. https://doi.org/10.1198/016214502753479248

Griffis, J. C., Allendorfer, J. B., &Szaflarski, J. P. (2016). Voxel-based Gaussian naïve Bayes classification of ischemic stroke lesions in individual T1-weighted MRI scans. Journal of Neuroscience Methods, 257, 97–108. https://doi.org/10.1016/j.jneumeth.2015.09.019

Guillen, P., &Ebalunode, J. (2016). Cancer Classification Based on Microarray Gene Expression Data Using Deep Learning. In 2016 International Conference on Computational Science and Computational Intelligence (CSCI) (pp. 1403–1405). https://doi.org/10.1109/CSCI.2016.0270

Guillen, P., & Guillen, P. (2017). Cancer Classification Based on Microarray Gene Expression Data Using Deep Learning Cancer Classification Based on Microarray Gene Expression Data Using Deep Learning, (February), 1403–1405. https://doi.org/10.1109/CSCI.2016.269

Guyon, Isabelle, W, Eston, J., & Stephen Barnhill. (2002). Gene Selection for Cancer Classification using Support Vector Machines. Machine Learning, 46, 389–422. https://doi.org/10.1155/2012/586246

Jemal, A., Bray, F., &Ferlay, J. (1999). Global Cancer Statistics: 2011. CA Cancer J Clin, 49(2), 1,33-64. https://doi.org/10.3322/caac.20107.Available

Kurzynski, M., Krysmann, M., Trajdos, P., &Wolczowski, A. (2016). Multiclassifiersystemwithhybridlearningappliedto the control of bioprosthetic hand. Computers in Biology and Medicine, 69, 286–297. https://doi.org/10.1016/j.compbiomed.2015.04.023

Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444. https://doi.org/10.1038/nature14539

Martens, H. H. (1959). Two notes on machine "Learning." Information and Control, 2(4), 364–379. https://doi.org/10.1016/S0019-9958(59)80014-0

Olivera, A. R., Roesler, V., Iochpe, C., Schmidt, M. I., Vigo, Á., Barreto, S. M., & Duncan, B. B. (2017). Comparisonof machine-learningalgorithmstobuild a predictivemodelfordetectingundiagnosed diabetes - ELSA-Brasil: accuracystudy. Sao Paulo Medical Journal, 135(3), 234–246. https://doi.org/10.1590/1516-3180.2016.0309010217

Statnikov, A., Wang, L., &Aliferis, C. F. (2008). A comprehensive comparison of random forests and support vector machines for microarray-based cancer classification. BMC Bioinformatics, 9, 1–10. https://doi.org/10.1186/1471-2105-9-319

Tsai, C. F. (2014). Combining cluster analysis with classifier ensembles to predict financial distress. ACM Transactions on Intelligent Systems and Technology, 16(1), 46–58. https://doi.org/10.1016/j.inffus.2011.12.001

Velikova, M., Lucas, P. J. F., Samulski, M., &Karssemeijer, N. (2013). On the interplay of machine learning and background knowledge in image interpretation by Bayesian networks. Artificial Intelligence in Medicine, 57(1), 73–86. https://doi.org/10.1016/j.artmed.2012.12.004

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 22/10/2018
Aceitação/Acceptance: 28/12/2018

Madurez de la Gestión del Conocimiento en la Ingeniería de Requisitos

Edgar Serna M.[1], Alexei Serna A.[2]

eserna@eserna.com, alexeiserna@gmail.com

[1,2] Universidad Autónoma Latinoamericana;[2] Instituto Antioqueño de Investigación Medellín, Antioquia.

Pages: 123–141

Resumen: Desde finales del siglo XX se ha publicado una serie de propuestas para Gestionar el Conocimiento en la Ingeniería de Requisitos que, aunque proceden del concepto de gestión tradicional basado en el mando y el control, en el futuro deberán dar lugar a un nuevo concepto de gestión que se ocupe más de desarrollar, apoyar, conectar, aprovechar y empoderar al equipo como trabajadores del conocimiento. En este artículo se presentan los resultados de una investigación para determinar el nivel de madurez de la Gestión del Conocimiento en la Ingeniería de Requisitos. Se realizó una revisión de la literatura acerca de las propuestas para gestionar este conocimiento y de las valoraciones que publican los investigadores luego de experimentarlas. Se seleccionaron siete propuestas, se encontraron ocho características valoradas y seis niveles de valoración, y se estructuró un modelo de madurez de cuatro niveles: 0. Infantil, 1. Adolescente, 2. Adulto, y 3. Veterano. La conclusión es que, hasta el momento, el nivel de madurez de KM en RE es Adolescente.

Palabras-clave: Gestión del conocimiento, Información, modelo de madurez, conocimiento tácito, conocimiento explícito.

Maturity of Knowledge Management in Requirements Engineering

Abstract: Since the end of the XX century, a series of proposals to Knowledge Management in Requirements Engineering have been published, although they come from the concept of traditional management based on command and control, in the future they should give rise to a new concept of management that is more responsible for developing, supporting, connecting, harnessing, and empowering the team as knowledge workers. In this article is presented the results of a research to determine maturity level of Knowledge Management in Requirements Engineering. A review of the literature on the proposals to manage this knowledge and the evaluations published by the researchers after the experiment was carried out. Seven proposals were selected, were found eight rated characteristics and six levels of assessment, and a four-level maturity model was structured: 0. Children, 1. Teenager, 2. Adult, and 3. Stager. The conclusion is that until now the maturity level of KM in RE is Teenager.

Keywords*:* Knowledge Management, Information, tacit knowledge, explicit knowledge, maturity model.

1. Introducción

En toda organización se crean documentos, se digitan datos, se envía información y, de muchas maneras, se documentan los procesos diarios de trabajo y producción. Por otro lado, los empleados intercambian ideas diariamente, para informar o clarificar dudas o discutir temas formales o informales sobre sus funciones. Todo esto es un proceso permanente en el que se refinan los datos y la información hasta convertirla en conocimiento (Davenport, 2010; Rus & Lindvall, 2002). El problema con ese conocimiento es que, mayoritariamente, reside en los individuos y, esporádicamente, en pequeños grupos. De esta manera se puede catalogar como explícito o tácito de acuerdo con la forma como se comparte: si se deja en documentos o se publica por algún medio entonces se considera explícito, porque es posible utilizarlo y aplicarlo en las actividades donde se requiera. Pero, si solamente se encuentra en la mente de los empleados, ya sea porque se ha acumulado por la experiencia o porque simplemente no se desea compartir, entonces es tácito, porque está oculto y no se publica ni se comunica.

En este sentido, el objetivo de KM es lograr un uso satisfactorio y productivo para el modelo organizacional o el objetivo del proceso (Rus & Lindvall, 2002). Por eso es que, para la Ingeniería de Requisitos es tan importante gestionar adecuadamente el conocimiento de los actores involucrados en esta fase del ciclo de vida del software. A este respecto, es conveniente tener en cuenta que toda esta gestión parte de la cadena datos-información-conocimiento relacionada con el problema que se desea resolver con el producto, y que se rige por las fases del conocimiento organizacional: 1) creación, cuando se comparten, formal o informalmente, datos o información entre los actores (a partir de dimensiones internas o externas al problema); 2) captura, que se logra cuando los actores lo documentan o almacenan en el cuerpo de conocimiento del problema; 3) transformación, cuando el equipo de trabajo lo organiza, mapea y convierte de forma transdisciplinar en una herramienta de solución al problema; 4) capitalización, cuando el equipo lo almacena en la Enciclopedia de Conocimiento y se convierte en activo de valor para resolver el problema; y 5) aprovechamiento, el equipo utiliza el conocimiento capitalizado para construir el documento de la especificación de requisitos.

Aunque en la Ingeniería del Software (SE) cada problema es único y cada modelo-solución tiene una metodología específica, tienen en común una serie de procesos complejos orientados a encontrarle una solución. Además, encontrar una solución software consiste de aplicar un enfoque sistemático, disciplinado y cuantificable para al desarrollo, operación y mantenimiento del producto. Por lo que los integrantes del equipo que lleva a cabo la Ingeniería de Requisitos (RE) se consideran trabajadores del conocimiento (Ibarra & Valle, 2017), y deben ser capaces de crear, capturar, transformar y capitalizar el conocimiento, explícito y tácito, que poseen las partes interesadas acerca del problema. De otra manera no podrán elicitar adecuadamente los requisitos, que el producto-solución debe satisfacer con calidad, fiabilidad y seguridad (Camacho, Sánchez & Galvis, 2013). Pero este es un tema que todavía necesita investigación porque, prácticamente, ninguno de los modelos de desarrollo de software tiene explícitamente incorporada una propuesta para gestionar el conocimiento en la Ingeniería de Requisitos (Natali & Falbo, 2002).

En la práctica, la Ingeniería de Requisitos se desarrolla en un ambiente de conocimiento intensivo, por lo que el equipo de trabajo debe gestionarlo para aprovecharlo eficientemente en pro de encontrar una posible solución al problema (Natali & Falbo,

2002; Ward & Aurum, 2004). En la Figura 1 se presenta la proximidad entre las fases del conocimiento y las etapas de RE, donde el objetivo es encontrar el conocimiento explícito y transformar el táctico, para generar un documento de especificación de requisitos sólido y fiable. En esta fase del ciclo de vida el conocimiento es diverso y creciente, por lo que los equipos tienen problemas para identificarlo, ubicarlo y encontrar la mejor manera de utilizarlo (Rus & Lindvall, 2002). Esto se debe a que primero hay que caracterizarlo y determinar las dimensiones de origen: en los procesos de negocio, la toma de decisiones, lo empresarial, lo declarativo, lo procedimental y la heurística, y luego darle el nivel de formalidad o informalidad necesario (Vasconcelos et al., 2009).

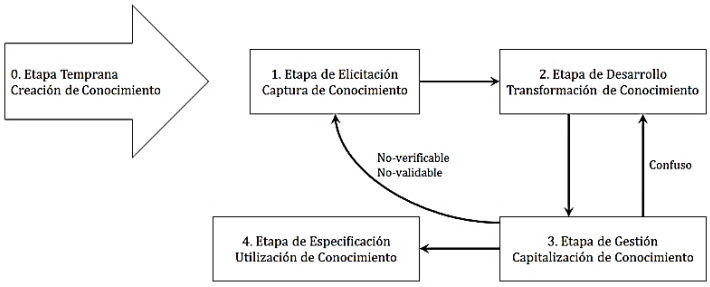


Figura 1 – Proximidad entre las fases de KM y las etapas de RE

Pero materializar esta proximidad entre las fases del conocimiento y las etapas de RE no es una tarea trivial, porque quien lo posee debe *traducirlo* a un lenguaje natural y utilizar canales de comunicación verbales y no-verbales para transferirlo a los demás integrantes del equipo, para que juntos lo decodifiquen de acuerdo con sus modelos mentales (Achive, 2018). Otro factor que incrementa la complejidad de esta actividad es el hecho de que el equipo y las partes interesadas son profesionales de diversas disciplinas (transdisciplinares), con experiencias, perspectivas, intereses y expectativas diferentes, que tienen datos-información-conocimiento sobre el problema tanto tácitos como explícitos y, en la mayoría de ocasiones, producto de su familiaridad, relación y función en el mismo. Esto genera lo que se conoce como *simetría de la ignorancia*, es decir, una brecha de conocimiento entre lo que creen saber del problema y lo que realmente saben del mismo, y esto dificulta el proceso de comunicación (Achive, 2018).

Diversos autores han presentado propuestas para gestionar el conocimiento en la Ingeniería de Requisitos, sin embargo, aplicarlas aisladamente, tal como se hace actualmente, no es suficiente (Shadbolt & Smart, 2015). Si el objetivo es logar una gestión eficiente primero hay que analizarlas y determinar su nivel de madurez y eficiencia (Rocha, 2011), para encontrar lo mejor de cada una y seleccionar la que mejor se adapte a las generalidades de RE. La idea es que sean altamente efectivas, que minimicen el tiempo entre las fases de KM y que permitan la conversión de lo tácito en explícito de manera ágil. Por lo tanto, este trabajo se sustenta en la hipótesis de que las propuestas para gestionar el conocimiento en la Ingeniería de Requisitos no tienen un adecuado nivel de madurez para atender la complejidad de los problemas actuales, pero que, al integrarlas, se podrían utilizar como punto de partida para desarrollar un modelo de KM para RE.

2. Necesidad de mejorar el nivel de madurez de KM en RE

En términos generales, el objetivo de KM es mejorar la manera como se comparte el conocimiento en las organizaciones y gestionar sus flujos. Nonaka y Takeuchi (1995) propusieron el modelo SECI: Socializar, Externalizar, Combinar e Interiorizar, para convertir y aprovechar el conocimiento en las organizaciones, porque el conocimiento se crea a partir de las interacciones sociales. *Socialización*: se transfiere entre individuos mediante el intercambio de modelos mentales y habilidades técnicas; *Externalización*: el tácito se convierte en explícito mediante el desarrollo de modelos, protocolos y guías; *Combinación*: la recombinación o reconfiguración de los cuerpos de conocimiento existentes crea nuevo conocimiento; *Internalización:* se alcanza aprendizaje mediante procesos repetitivos.

El interés por gestionar el conocimiento adecuadamente en RE ha llevado a diversos autores a emprender investigaciones para encontrar la mejor manera de hacerlo. En sus propuestas tratan la cuestión del conocimiento en las diversas etapas de RE, tales como la elicitación (Friedrich & Van Der Poll, 2007), el desarrollo (Medeni et al., 2011) o la especificación (Stone & Sawyer, 2006), a la vez que desarrollan propuestas metodológicas (Friedrich & Van Der Poll, 2007), herramientas específicas (Stoiber & Glinz, 2009) o técnicas (Mohamed, 2010), y Gacitua et al. (2009) investigan el papel del conocimiento en RE y proponen cómo mitigar y gestionar sus efectos; sin embargo, ninguno enfrenta el problema desde una perspectiva de KM.

Por otro lado, algunos autores han adoptado el modelo SECI para trabajarlo en la Ingeniería de Requisitos, tales como Pilat y Kaindl (2011), quienes proponen una perspectiva de KM para crear conocimiento; Wan et al. (2010) plantean una propuesta para convertir el conocimiento en las actividades de la elicitación de requisitos, con el objetivo de minimizar la simetría de la ignorancia entre los integrantes del equipo de trabajo. Por su parte, Vázquez et al. (2012) realizan una clasificación de las técnicas de elicitación y proponen cómo seleccionarlas de acuerdo con el modelo SECI, su objetivo es reducir las dificultades de una adecuada selección. Todos los autores mencionados hasta el momento comparten una perspectiva de KM, pero ninguno formula una propuesta, metodología, estrategia o proceso para determinar el nivel de madurez de las implementaciones de KM en RE.

Por otro lado, pero centrados en la Ingeniería del Software (SE), Olmos y Ordas (2013) describen un proceso basado en conocimiento para la elicitación de requisitos. Para ellos, todavía persisten muchos problemas en la industria del software, porque utilizan técnicas informales y no hay algoritmos definidos que puedan ayudar a la solución. Esta falta de madurez en las técnicas aplicadas crea ambigüedades e inconsistencias, por lo que lo más conveniente es gestionar el conocimiento de otra manera. Khalid y sus colegas (2015) proponen un modelo de conocimiento para ayudar a resolver los problemas de eficiencia de las propuestas de KM en Global Software Development (GSD), utilizando un repositorio central de conocimientos como servidor de KM en la nube. La propuesta de Ward y Aurum (2004) se orienta a desarrollar un modelo de conocimiento que se ha aplicado en algunas organizaciones de software para modelar el conocimiento. Los autores describen la inmadurez de otras propuestas para gestionar el conocimiento en SE.

Por su parte, Kess y Haapasalo (2002) argumentan que, esencialmente, SE es un proceso en el que KM se trabaja marginalmente, por lo que estructuran una propuesta

para integrarlo. Dingsøyr y sus colegas (2001) describen los problemas de la industria por falta de propuestas adecuadas para la gestión del conocimiento, mientras que Aurum et al. (2004) reafirman la necesidad de buenas propuestas para hacerlo en SE. Proponen la solución como un reto para los investigadores y detallan algunos factores clave que pueden tener en cuenta para lograrlo. En este sentido, Li (2007) aporta un procedimiento para compartir y crear conocimiento en estas organizaciones; Wongthongtham y Chang (2007) describen las ontologías como modelos para gestionar el conocimiento; Georgiades et al. (2005) definen una solución basada en la semántica del lenguaje natural; y Cuena y Molina's (2000) identifican como una falacia las fases de SE tradicional en las que se implementa KM.

Otro asunto que no ha permitido la maduración de las propuestas para gestionar el conocimiento en la Ingeniería de Requisitos, es que no se encuentra unanimidad en el lenguaje utilizado para formular las terminologías y estrategias de KM (Pilat & Kaindl, 2011). Estos autores afirman que el problema central de las propuestas es que en RE los actores no se comunican adecuadamente, debido a que tienen culturas y profesiones diferentes. White (2010) afirma que el equipo de trabajo debe ser capaz de comprender las acciones humanas y las normas culturales de cliente en relación con el problema. Debido a que el proceso RE es básicamente trabajo en equipo, debe estar bien definido, utilizar prototipos, un vocabulario común para la comunicación, un ambiente de apoyo y abordar la gestión del conocimiento en forma de anotaciones y símbolos para resolverlos por separado (Rouse, Cannon & Salas, 1992). Para estos autores aquí es donde radica la falta de eficiencia de las propuestas para hacerlo. Además, deben tener en cuenta que el conocimiento, como los requisitos, no proviene de una fuente específica, sino que se origina en diversas dimensiones (Lambe, (2007), tales como el dominio, el contexto, el medio ambiente, la experiencia, el grupo y la cultura. Además, se encontraron algunos esfuerzos relacionados con la madurez de KM en RE: algunos centrados en herramientas TIC (Uenalan et al., 2008); otros sugieren metodologías para medirla en partes específicas de KM (Supakkul et al., 2009) o en nichos específicos de RE (Schmid, 2009). Mientras que Gacitua y sus colegas (Gacitua et al., 2009) intentan hacerlo de forma más amplia y discuten indirectamente cómo afecta el nivel de madurez de las propuestas actuales a la especificación de requisitos.

3. Metodología

Metodología es el estudio de los métodos, el conjunto de técnicas y los procesos utilizados para la investigación y el desarrollo del trabajo científico. Debido a que la naturaleza de esta investigación se utilizó el Diseño de Investigación Científica (DSR), que es fundamentalmente una metodología para resolver problemas con el objetivo de crear innovaciones que definan ideas, prácticas, capacidades técnicas o productos, a los que se puede mejorar su eficiencia y efectividad mediante el análisis, el diseño, la implementación, la gestión y el uso de las experiencias publicadas (Hevner et al., 2004).

Además de las actividades en las diferentes etapas de la Ingeniería de Requisitos, DSR tiene su ciclo de investigación, por lo que es una metodología ampliamente utilizada en la realización de investigaciones relacionadas con el desarrollo de software. Esta metodología se estructura alrededor de varias directrices y, para esta investigación, se toman las siguientes: 1) relevancia del problema, 2) contribuciones al cuerpo del

conocimiento, 3) rigor de la investigación, 4) diseño como proceso de búsqueda, y 5) comunicación de resultados (Hevner, 2007). Además, la base del conocimiento pueden ser teorías, métodos, experiencias, especialistas, procesos y artefactos, pero el rigor de la misma depende de una adecuada selección, aplicación y evaluación de ese conocimiento para analizarlo y obtener los resultados previstos.

Una actividad que pretenda ser eficiente y objetiva, como es el caso de las involucradas en la Ingeniería del Software, se convierte en un tema relevante para un equipo de desarrollo de software, porque debe dedicar hasta un 60% del tiempo del ciclo de vida a tratar de entender su entorno, a través de una adecuada gestión del conocimiento en la Ingeniería de requisitos. Por eso es que este problema es relevante para la comunidad. DSR es una metodología inherentemente iterativa, por lo que el diseño como proceso de búsqueda se basa en una revisión de la literatura, a la que se añade la experiencia personal de los autores y de los especialistas consultados. Por otro lado, es esencialmente un proceso de búsqueda para descubrir el estado de la cuestión y proponer una solución al problema.

Para esta investigación se adapta la propuesta DSR de Hevner (2007) de tres ciclos (Figura 2): En el *ciclo de diseño* se diseñan, analizan y evalúan los procesos de búsqueda que se aplicarán; en el *ciclo de relevancia* se itera la búsqueda en las bases de datos hasta encontrar los trabajos relevantes para responder a las preguntas de investigación; y en *el ciclo de rigor* se construye el cuerpo de conocimiento con base en la validación de la valoración a las propuestas analizadas.

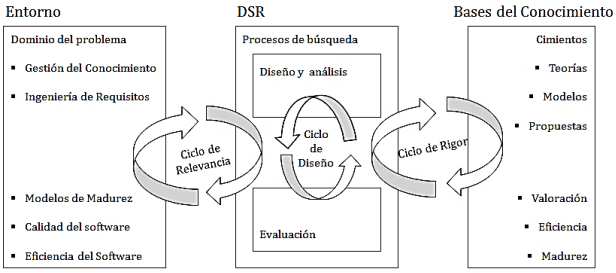


Figura 2 – Metodología adaptada de Hevner (2007)

4. Resultados

Las preguntas en las que se basa esta investigación son:

1. ¿Qué Propuestas para Gestionar el Conocimiento en la Ingeniería de Requisitos (PGCIR) se han publicado?
2. De acuerdo con la evaluación a estas propuestas, ¿Cuál es el nivel de madurez de la Gestión del Conocimiento en la Ingeniería de Requisitos?

4.1. Pregunta 1: PGCIR encontradas

Luego de aplicar el *ciclo de diseño* de la metodología y llevar a cabo una búsqueda de las características del entorno de investigación, se encontraron 24 PGCIR. En el *ciclo de relevancia* se realiza una evaluación concreta, en la que tuvo en cuenta la estructura,

Propuesta	Descripción	Observaciones
ISO STEP (ISO, 1994)	Se puede utilizar para intercambiar datos e información en varias industrias, entre ellas la del software, especialmente en la Ingeniería de Requisitos.	Proporciona información específica del dominio y los protocolos ofrecen conocimiento específico (Heimannsfeld & Müller, 2000). Un riguroso intento por identificar y demostrar una estrategia de razonamiento en RE (Scott & Cook, 2003). Los requisitos se expresan en lenguaje natural, con los problemas de lo tácito en él (Ratchev et al., 2005). Se adapta en el tiempo para clarificar y actualizar los requisitos. No resuelve el problema de la ambigüedad acerca de qué es un requisito y los actores no logran una comprensión del problema (Johnson, 2006).
MTKISDP (Hickey & Davis, 2003)	Se basa en los principios de comprender, conocer y elicitar. Da prioridad al conocimiento crítico necesario y define la base subyacente para su implementación en lenguaje formal.	Presenta un modelo de elicitación centrado en la selección de la técnica, deja la gestión del conocimiento en segundo plano (Bendjenna et al., 2008). No logra capturar el conocimiento relevante para cada dominio porque KM debe abarcar mucho más que identificar y elicitar (Vásquez et al., 2014). Permite representar los usuarios y utiliza grupos y comunidades; no tiene en cuenta dimensiones ni transdisciplina de requisitos (Tuunanen, 2003). Otros autores han publicado análisis indirectos a esta propuesta (Aranda et al., 2005; Aranda et al., 2008; Huang et al., 2005; Mason & Pauleen, 2003; Barkha & Sumaira, 2010), no logra una valoración sobresaliente para las características como proponen aquí.
RMKMSE (Andrade et al., 2006)	Se compone de un esquema de formalización para representar, capturar y transmitir el conocimiento relevante. Se centra en el mantenimiento del software.	Es una experiencia de uso en KM no-exitosa (Aung & Nyunt, 2013). Se orienta esencialmente a la captura del conocimiento en RE; usa la memoria corporativa como principal fuente; no aprovecha eficientemente a las partes interesadas (Carreteiro et al., 2016). Se focaliza en la participación en el conocimiento y la reutilización de las lecciones aprendidas; deja de lado aspectos importantes de RE (Rodríguez et al., 2004; Talib et al., 2010). No aprovecha suficientemente el conocimiento de los actores involucrados; se centra básicamente en la organización (Anquetil et al., 2007).
KBRE (Pavanasam et al., 2010)	Es un modelo integrado de sistemas expertos y modelado específico de dominio. Uno define las reglas y propiedades del sistema y el otro requiere una herramienta metacase.	Es un método centrado en cómo modelar y utilizar eficazmente el conocimiento del dominio, pero se olvida del tácito que poseen las partes interesadas; satisface algunas características de madurez, pero deja de lado las más necesarias (Valero et al., 2013).
KMPRE (Pilat & Kaindl, 2011)	Se basa en el conocimiento compartido acerca del problema y su dominio. Su objetivo es superar el problema del intercambio de conocimiento en RE, enfatizando en la importancia de elicitar a partir del mismo.	Es un enfoque en el que predomina la comunicación entre las partes; facilita la transferencia y el intercambio de conocimiento (Venkatesh & Kumar, 2013). Es incompleta y costosa y requiere tiempo; no logra la transformación ni la transferencia de conocimiento en la dimensión que se requiere (Olmos & Rodas, 2014). Se basa en el modelo SECI (Nonaka & Takeuchi, 1995) y ofrece ideas y técnicas básicas para comprender y facilitar el proceso de transferencia de conocimiento y transformación, pero nada más (Olmos & Rodas, 2016).
KMFSE-SECI (Chikh, 2011)	Utiliza ontologías para describir la especificación y representar el contenido de los requisitos; se sustenta en el modelo SECI.	Necesita ir más allá del explícito, porque en RE el tácito es muy importante (Mohamed et al., 2015). Es algo flexible y adaptable, pero es difícil de aplicar a otras funcionalidades (Hanafiah at al., 2015).
KMoS-RE (Olmos & Rodas, 2014)	Se orienta a la transformación y transferencia de conocimiento; incorpora métodos para identificar, capturar, indexar y hacer explícito el tácito.	Tiene en cuenta el modelo de Jackson (Jackson, 1995) y se basa en SECI (Nonaka & Takeuchi, 1995). Es un modelo de conocimiento sobre una estrategia para RE; no tiene en cuenta la complejidad y las dimensiones de los requisitos, el dominio informal estructurado, ni las ambigüedades (Humayoun & Qazi, 2015).

Tabla 1 – PGCIR encontradas en la revisión de la literatura

la relevancia del autor en este campo y el seguimiento que ha recibido cada propuesta con base en las citaciones recibidas. Posteriormente, en el *ciclo de rigor* se extractaron las validaciones y valoraciones para cada propuesta, para construir las bases de conocimiento que orientan la selección de la muestra. Al final de este proceso se seleccionaron las *siete* propuestas que se describen en la Tabla 1.

4.2. Pregunta 2: Madurez de la Gestión del Conocimiento en RE

Para determinar esta madurez se estructuró el siguiente procedimiento: 1) identificar y definir las características valoradas por los autores (Tabla 2), 2) seleccionar los niveles de valoración aplicados (Tabla 3), 3) resumir los niveles de valoración a cada característica con base en los aportes de los autores y sus análisis (Tabla 4), 4) construir un modelo de madurez para analizar las valoraciones, y 5) encontrar el nivel de madurez de la Gestión del Conocimiento en RE (Tabla 5).

1. *Características de valoración*. En el contexto de la investigación y a partir de la estructura y definición de sus etapas e interrelaciones, se construyó un marco de evaluación al nivel de madurez de las PCGIR con base en las descripciones encontradas en la revisión de la literatura. El objetivo fue estructurar un marco de referencia y análisis para identificar su eficiencia y eficacia, teniendo en cuenta los diferentes aspectos que evaluaron los autores en la experimentación de las propuestas. En esta investigación se encontró que los investigadores tienen en cuenta que la propuesta debe ser sistémica, además de responder a las características propias de los sistemas actuales, como se resumen en la Tabla 2.

Características	Descripción
Flexibilidad (F)	Se puede implementar en todo tipo de problemas, sin importar tamaño, grado de complejidad o nivel de exigencia.
Adaptabilidad (A)	Se puede adecuar en función del contexto y tipo de organización, el equipo de trabajo, la infraestructura, los objetivos estratégicos y la base de conocimiento, entre otros.
Multidimensionalidad (Mu)	Permite la selección de diversas estrategias para gestionar el conocimiento proveniente de las dimensiones involucradas en el problema.
Transdisciplina (T)	Permite gestionar la cadena datos-información-conocimiento proveniente de una amplia variedad de disciplinas.
Evolutividad (E)	Facilita el desarrollo de KM en la medida que las etapas de RE evolucionan hacia el documento de especificación y que la comunicación crecen en alcance y complejidad.
Capacidad de realimentación (R)	Se regula a sí mismo y cada resultado de un sub-proceso incide en el todo, reintegrándolo y modificándolo, permitiendo que el modelo tenga capacidad de control y autocorrección.
Medible (Mi)	Ofrece métricas para los resultados, la evolución y el impacto de la gestión del conocimiento en el logro de la especificación.
Prospectiva (P)	Permite identificar oportunidades de mejora y de realizar modificaciones o cambios oportunos en la gestión y líneas de acción adoptadas por el equipo.

Tabla 2 – Características de valoración

2. *Niveles de valoración*. En la Tabla 3 se describen los niveles de valoración tomados de los mismos análisis que realizan los autores a las PGCIR.

	Detalle	Descripción
0	Inexistente	No la permite (*Nunca*)
1	Inicial	La permite parcialmente, es adecuada y corresponde a lo esperado (*Casi nunca*)
2	Repetible	Está homologada, pero no está estandarizada ni tiene gobernabilidad (*Rara vez*)
3	Definida	Está homologada, no está estandarizada, pero tiene gobernabilidad (*A veces*)
4	Administrada	Está completamente homologada en la funcionalidad del modelo, está estandarizada para las distintas etapas y cuenta con gobernabilidad (*Casi siempre*)
5	Optimizada	Opera completamente y se puede refinar hasta niveles de mejores prácticas, se basa en los resultados de mejoras y diseños continuos (*Siempre*)

Tabla 3 – Niveles de valoración

3. *Valoración de las características*. En la Tabla 4 se presenta la valoración a las características de las PGCIR, resultado del análisis hecho por los autores consultados en la experimentación de las mismas.

Propuestas	Características valoradas							
	F	A	Mu	T	E	R	Mi	P
ISO STEP - AP233	3	3	0	0	2	2	3	0
MTKISDP	1	1	0	0	3	2	3	0
RMKMSE	2	2	0	0	2	2	2	0
KBRE	2	1	0	0	0	3	2	1
KMPRE	1	2	0	0	0	2	2	1
KMFSE-SECI	3	3	0	1	1	2	1	1
KMoS-RE	3	3	1	0	1	3	3	1

Tabla 4 – Valoración de las características de las PGCIR

4. *Modelo de madurez aplicado*. Conocer el nivel de madurez de KM en RE será posible en la medida que las propuestas para hacerlo demuestren eficacia y eficiencia, y esto se logra ubicando el proceso dentro de un nivel determinado. Aunque la mayoría de investigadores están familiarizados con los niveles de madurez utilizados por CMM, en esta investigación se propone determinar esta madurez aplicando una perspectiva con niveles comparativos al proceso en los seres humanos, es decir: Infantil, Adolescente, Adulto y Veterano.

- *Nivel 0: Infantil*: cuando necesita mucho cuidado, atención y *afecto*; sus características son:
  - La gestión del conocimiento se realiza solamente al finalizar la fase, porque no está en sintonía con el ciclo de vida del desarrollo de software.
  - Los problemas de comunicación hacen que la gestión falle y que el equipo

no llegue a un acuerdo dialogado, porque la falta de un lenguaje común invalida cualquier intento de conclusión.
    - Se generan re-procesos porque, en cualquier caso, hay que modificar o actualizar la cadena datos-información-conocimiento que se intenta gestionar.
    - El equipo de trabajo invierte más tiempo en tratar de concluir que en comprender la Transdisciplina y Multidimensionalidad de la cadena que desea gestionar.
    - La organización que presenta el problema impone la visión de una disciplina desde la que se debe realizar la gestión del conocimiento.
    - Debido a que hay que *cuidar*, *mimar* y prestarle mucha *atención*, la gestión del conocimiento no se puede dejar sin vigilancia por mucho tiempo, por lo que se puede afirmar que el nivel de madurez de este proceso desalienta, en lugar de alentar al equipo a seguir adelante.

- *Nivel 1: Adolescente*: cuando se puede aplicar y seguir sin mucho detalle durante un tiempo razonable, pero todavía se desconfía de su *responsabilidad*; sus características son:
    - Un error en la gestión ocasiona problemas en las otras etapas de la propuesta.
    - Es importante conocer el error, pero no se necesitará mucho esfuerzo para demostrarlo.
    - Se desperdicia mucho tiempo intentando analizar la causa de cada problema a la vez que se deja de ejecutar varias etapas de la propuesta.
    - Para encontrar los problemas, el equipo debe solucionar las inconsistencias y volver a correr las etapas recorridas, un ciclo que se tiene que repetir muchas veces.
    - En este nivel, y al igual que con los *humanos adolescentes*, la gestión del conocimiento es sorprendentemente útil, pero se comporta de manera irresponsable y puede causar más daño que beneficio a RE.

- *Nivel 2: Adulto*: cuando es digna de confianza y se puede aplicar sin mucha *vigilancia* durante algunas etapas de RE, pero todavía se desvía de las *responsabilidades*; sus características son:
    - Al final de cada etapa el equipo logra analizar buena parte de la cadena datos-información-conocimiento.
    - Aunque el equipo todavía no llega a conclusiones dialogadas, los aportes disciplinares se tienen en cuenta y se analiza cómo integrarlos.
    - La gestión del conocimiento es algo más que reunir datos-información-conocimiento.
    - El equipo se puede concentrar en encontrar y corregir los problemas de comunicación, mientras se progresa en las demás etapas de la propuesta sin *intervención*.
    - En este nivel, la gestión del conocimiento no requiere *vigilancia* extrema y, aunque el proceso es *responsable* y se puede *confiar* en él, todavía no es auto-dirigido porque su madurez no ha llegado a un nivel de independencia que le permita al equipo dejar que la propuesta guie autónomamente.

- *Nivel 3: Veterano*: cuando ha evolucionado y crecido a través de los niveles anteriores y es posible dejar que siga su curso sin *vigilancia permanente*; el conocimiento es totalmente gestionable; es autónoma; se analiza, discute y concluye la cadena datos-información-conocimiento con aportes desde todas las disciplinas y dimensiones necesarias. Sus características son:
  - La gestión del conocimiento es eficiente y eficaz, y el equipo desarrolla todo el proceso de forma transdisciplinar y multidimensional.
  - El equipo reutiliza la cadena datos-información-conocimiento en sus conclusiones, porque la entiende y dialoga permanentemente.
  - Al final de RE se tiene conocimiento validado y maduro, y una serie de reportes que denotan la calidad y fiabilidad de la gestión realizada.
  - La gestión del conocimiento se estructura a partir de una propuesta planificada que involucra la transdisciplina y multidimensionalidad de la cadena datos-información-conocimiento.
  - Se puede aplicar procedimientos de planeación estratégica en las etapas.
  - La gestión del conocimiento es un proceso *autónomo* en cuanto a reproducción y selección de la cadena de entrada, y a la validación, discusión y exposición de resultados.
  - La organización cuenta con un modelo de gestión del conocimiento reutilizable y fácil de proyectar a cada solución analizada.
  - Como en el caso de los humanos, en este nivel KM aporta experiencia y madurez, con el objetivo de que el producto de esta fase del ciclo de vida sea de utilidad para las demás.

5. *Madurez de la Gestión del Conocimiento en la Ingeniería de Requisitos*. Con base en los resultados presentados en la Tabla 4 y en las opiniones, reflexiones y críticas que hacen los investigadores a las propuestas, en la Tabla 5 se presenta el nivel de madurez de la Gestión del Conocimiento en la Ingeniería de Requisitos, calculado a partir de los resultados de la evaluación a las características que hacen los autores consultados, los niveles de valoración definidos y el modelo de madurez aplicado.

Características	**Valoración promedio**	**Nivel de madurez**			
		Infantil	**Adolescente**	**Adulto**	**Veterano**
Flexibilidad (F)	Definida		X		
Adaptabilidad (A)	Definida		X		
Multidimensionalidad (Mu)	Inicial	X			
Transdisciplina (T)	Inicial	X			
Evolutividad (E)	Inicial	X			
Capacidad de realimentación (R)	Repetible		X		
Medible (Mi)	Repetible		X		
Prospectiva (P)	Inicial	X			
Valoración final			X		

Tabla 5 – Madurez de la Gestión del Conocimiento en la Ingeniería de Requisitos

5. Análisis de resultados

Al analizar la aplicación de las diferentes propuestas en la industria, el equipo de investigadores se encontró con evaluaciones que despreciaban cada una cada de ellas, pero al consultar con los equipos de RE se pudo constatar que las empresas hacen prevalecer los costos y el tiempo en el ciclo de desarrollo, por lo que ellos están presionados a no seguir las fases estructuradas sino a utilizar las que puedan necesitar. Debido a esta realidad, en la que pocas empresas tienen la paciencia y los recursos para aplicar el proceso completo de las propuestas, los autores tomaron la decisión de analizar únicamente los comentarios y evaluaciones que otros investigadores han publicado, luego de experimentar con las PGCIR.

Por otro lado, el problema se complica porque los clientes se han vuelto cada vez más exigentes y los sistemas software cada vez más complejos. El aseguramiento de la calidad se está convirtiendo en una condición permanente para la sobrevivencia de las empresas y, actualmente, esta una realidad en la industria del software. A pesar de que en ciertas esferas se afirma que la calidad del software ha mejorado en los últimos años, todavía está muy lejos del escenario ideal y de la madurez esperada, esencialmente porque la Ingeniería de Requisitos todavía presenta falencias al especificar las necesidades que debe satisfacer el producto. Para colaborar a llenar este vacío, muchos investigadores han estructurado y publicado diferentes propuestas para gestionar el conocimiento en esta fase, pero, de acuerdo con los resultados analizados en esta investigación, todavía son limitadas y muchas no se centran directamente en KM.

Aunque la Ingeniería de Requisitos es una fase esencial para desarrollar un producto que satisfaga al cliente y una parte integral de todo el ciclo de vida del desarrollo de software, la gestión del conocimiento en ella requiere un nivel de madurez que le haga eficiente, efectiva y flexible. Estas características se hacen cada vez más complicadas de lograr, por lo que continuamente se incrementa la demanda por una RE madura que asegure que la especificación de requisitos ocurra sobre una base regular de KM. Además, que facilite la construcción de escenarios de comprensión del problema cada vez más reales, que requieran menos esfuerzo y que les permita a los equipos de desarrollo obtener y gestionar el conocimiento tácito y explícito sobre el sistema en desarrollo. Pero, de acuerdo con los resultados de esta investigación, los autores que han experimentado las PGCIR afirman que todavía no se puede demostrar que lo logran.

De acuerdo con los resultados obtenidos en esta investigación algunas empresas utilizan las PGCIR como una especie de *bala de plata*, que podría resolver problemas en la calidad del producto, ayudar a elicitar y especificar todos los requisitos y ahorrar tiempo y esfuerzo. Este no es el caso, porque la implementación implica una curva de aprendizaje progresivo hasta el equipo alcance las habilidades necesarias. Eso significa que, en realidad, en los primeros proyectos la propuesta seleccionada puede incrementar el esfuerzo y el costo de RE. También es importante destacar que no se encontró una propuesta genérica que se pudiera aplicar en los diversos contextos y problemas en los que se desarrolla el software. Esto significa que la mayoría requiere alguna adaptación previa y la reinducción del equipo en cada caso, porque no hay algo así como una propuesta de *talla única*.

La creencia de que una propuesta determinada les ayudará a reducir costos y tiempos de entrega, hace que las empresas se convenzan de poder cumplir con los plazos.

Pero actualmente es poco probable que alguna de ellas ayude en este sentido, porque los resultados demuestran que no reducen inmediatamente el esfuerzo y el tiempo, debido a que requieren capacitación y el desarrollo de habilidades para utilizar de forma dinámica. Además, la complejidad del software no da cabida a una reutilización inmediata de ninguna herramienta de KM, y las empresas piensan en términos de fiabilidad y reutilización en lugar de simplemente en la eficacia de la gestión. Otro asunto que la industria no comprende a cabalidad es que, actualmente, no todo el conocimiento en RE se puede descubrir, comprender y gestionar, muchos usuarios y clientes *esconden* conceptos clave, lo que representa un problema porque es poco probable que una PGCIR por sí sola pueda ser capaz de encontrar el cien por ciento de ese conocimiento.

En términos generales y de acuerdo con los resultados en esta investigación, se puede concluir que el nivel de madurez de la Gestión del Conocimiento en la Ingeniería de Requisitos, como área de investigación, desarrollo y aplicación, es *Adolescente*. Una conclusión a la que llegan los autores luego de analizar e interpretar los datos y valoraciones que los investigadores publicaron, luego de experimentar con las propuestas seleccionadas en la revisión de la literatura. Mejorar este nivel es una estrategia necesaria y, para muchas empresas, la única opción para optimizar los estándares de calidad del software, pero en esta investigación se encontró que todavía se encuentra en el proceso de maduración que requiere cualquier aplicación compleja en el mundo actual. Por otro lado, en la práctica las características evaluadas en cada una de las propuestas se encuentran en diferentes niveles de madurez, porque en cada etapa de RE surge un conocimiento único que le permite al equipo pasar a la gestión en la etapa siguiente. Pero un problema para lograrlo es que actualmente en la industria esta gestión se observa como un proceso más del desarrollo de software, y las propuestas para hacerlo se estructuran, al parecer, sin una planeación real de las necesidades de la industria. Esto genera algunos desafíos que debe afrontar una PGCIR para alcanzar un nivel de madurez *Veterano*:

- El tiempo que se invierte para aplicarla no debe extender las actividades de RE.
- Los cambios permanentes en la formulación y los escenarios de los requisitos le exigen adaptación continua.
- Debe gestionar el conocimiento que generan tanto los requisitos explícitos como los tácitos.
- La industria invierte dinero, tiempo y esfuerzo en el desarrollo de software, por lo tanto, la propuesta debe permitir alcanzar resultados rápidamente.
- Las personas con las habilidades para gestionar conocimiento eficientemente son escasas, por lo que la propuesta debe ser fácil de comprender y aplicar.
- Un desafío importante para que una PGCIR se utilice y mejore su nivel de madurez como área de investigación y de desarrollo es el costo de aplicación. La industria opta por no aplicar herramientas que incrementen el tiempo o el valor del desarrollo de software.

6. Conclusiones

Es importante observar cómo en las últimas décadas se ha incrementado el interés de los investigadores por divulgar Propuestas para Gestionar el Conocimiento en la Ingeniería de Requisitos, en parte, motivados por aportar al mejoramiento y la escalabilidad del

proceso de desarrollo de software. Al mismo tiempo, también llama la atención que, aunque se conozcan y compartan estas propuestas, a la industria parece que le faltara algo: no las aplican rigurosamente. Una de las principales cosas que recalcan sus autores es la importancia de no saltarse las fases, porque para que sea eficiente y eficaz para el equipo de trabajo, la PGCIR seleccionada de debe aplicar en su totalidad. De acuerdo con el modelo de madurez que se propone en este trabajo, las personas crecen a través de cada etapa y a partir de lo que lograron en la anterior, es decir, no es posible dejar de ser niño para pasar automáticamente a ser adulto, aunque se intente desesperadamente, porque si se salta la adolescencia esa persona está destinada a fracasar en la vida. Lo mismo sucede con las PGCIR.

El cambio de paradigma y el aprovechamiento de la Gestión del Conocimiento en todo tipo de actividades está creciendo rápidamente, en parte porque las organizaciones se han dado cuenta de su potencialidad y porque los investigadores han innovado y mejorado sus propuestas. Esto ha permitido que se pueda utilizar en áreas como la Ingeniería de Requisitos para buscar una mejor especificación. Los conceptos y procesos fundamentales que se analizan en esta investigación reflejan un cambio creciente en el pensamiento y la práctica de KM en el desarrollo de software. Este cambio de paradigma, aunque procede del concepto de gestión tradicional, basado en el mando y el control, debe dar lugar a un nuevo concepto de gestión que se ocupe más de desarrollar, apoyar, conectar, aprovechar y empoderar al equipo de RE como trabajadores del conocimiento.

En el entorno de la Ingeniería del Software, la Gestión del Conocimiento les interesa a muchos profesionales e investigadores, y muchas organizaciones la prueban como un potencial repositorio de soluciones relacionadas con problemas en el desarrollo de software. Los desarrollos recientes en TI permiten compartir el conocimiento, documentado o no, independientemente del tiempo y el espacio. En este sentido, se puede prever que en el futuro se podrá capturar desde y difundir en diversos formatos y dimensiones, permitiéndole al equipo de RE elicitar los requisitos a cualquier escala y complejidad. Pero esto depende en gran medida de la eficiencia y efectividad de las PGCIR y del mejoramiento en el nivel de madurez de esta área de investigación y desarrollo.

Teniendo en cuenta que uno de los factores clave de éxito de un proyecto software reside en el hecho de que su propósito responde al cumplimiento de un objetivo estratégico de la organización, la Gestión del Conocimiento en la Ingeniería de Requisitos debe responder a sinergias de valor agregado y a dinámicas de intercambio en la cadena datos-información-conocimiento, lo que supone un importante reto socio-cultural-administrativo para el equipo de trabajo. Por eso es que una propuesta para gestionar el conocimiento que se comunica, analiza y difunde en esta fase, debe ser dinámica y adaptativa al problema y al propósito de la organización, a la vez que evolutiva conforme se asimila o ajusta el conocimiento tácito y explicito que se elicita.

Referencias

Achive, T. (2018). Coexistence of the natural and the artificial in the mind: Engineering the non-possible. *Actas de Ingeniería*, 4(1), 35–42.

Andrade, J. et al. (2006). A Reference Model for Knowledge Management in Software Engineering. *Engineering Letters*, 13(2), 14–21.

Anquetil, N. et al. (2007). Software maintenance seen as a knowledge management issue. *Information and Software Technology*, 49(5), 515–529.

Aranda, G. et al. (2005). Choosing groupware tools and elicitation techniques according to stakeholders' features. In: *International Conference on Enterprise Information Systems*. Miami, USA.

Aranda, G. et al. (2008). Strategies to minimize problems in global requirements elicitation. In: *Special issue of best papers presented at workshop doctoral colloquium with one paper selected from CLEI Electronic Journal,* 11(1), Paper 3.

Aung, Z. & Nyunt, K. (2013). Constructive Knowledge Management Model and Information Retrieval Methods for Software Engineering. In: Buragga, K. & Zaman, N. (Eds.), *Software Development Techniques for Constructive Information Systems Design* (pp. 377-393). IGI Global: USA.

Aurum, A., Parkin, P. & Cox, K. (2004). Knowledge Management in Software Engineering Education. In: *IEEE Intern. Conference on Advanced Learning Technologies*. Joensuu, Finland.

Barkha, J. & Sumaira, S. (2010). Process support for requirements engineering activities in global software development: A literature based evaluation. In: *International Conference on Computational Intelligence and Software Engineering*. Wuhan, China.

Bendjenna, H, Zarour, N. & Charrel, P. (2008). Enhancing Elicitation Technique Selection Process in a Cooperative Distributed Environment. *Lectures Notes on Computer Science*, 5025, 23–36.

Camacho, J., Sánchez, J. & Galvis, E. (2013). Understanding the Process of Knowledge Transfer in Software Engineering: A Systematic Literature Review. *Int. Jour. of Soft. Comp. and Soft. Eng.*, 3(3), 219–229.

Carreteiro, P. et al. (2016). A knowledge management approach for software engineering projects development. In: Rocha, A. et al. (Eds.), *New Advances in Infor. Systems and Tech.* (pp. 59-68). Springer: Switzerland.

Chikh, A. (2011). A Knowledge Management Framework in Software Requirements Engineering based on the SECI Model. *Journal of Software Engineering and Applications*, 4(12), 718–728.

Cuena, J. & Molina's, M. (2000). The Role of Knowledge Modeling Techniques in Software Development: A General Approach Based on Knowledge Management Tools. *International Journal of Human-Computer Studies,* 52(3), 385–421.

Davenport, T. (2010). *Process Management for Knowledge Work. Handbook on Business Process Management*. Springer: USA.

Dingsøyr, T., Brede, N. & Nytro, O. (2001). Augmenting Experience Reports with Lightweight Postmortem Reviews. In: *3rd International Conference on Product Focused Software Process Improvement*. Kaiserslautern, Germany.

Friedrich, W. & Van Der Poll, J. (2007). Towards a methodology to elicit tacit domain knowledge from users. *Interdisciplinary Journal of Information, Knowledge, and Management 2*, 178–193.

Gacitua, R. et al. (2009). Making tacit requirements explicit. In: *Second International Workshop on Managing Requirements Knowledge*. Atlanta, USA.

Gacitua, R. et al. (2009). Making Tacit Requirements Explicit. In: *Second International Workshop on Managing Requirements Knowledge*. Washington, USA.

Georgiades, M., Andreou, A. & Pattichis, C. (2005). A Requirement Engineering Methodology Based on Natural Language Syntax and Semantics. In: *13th IEEE International Conference on Requirements Engineering*. Paris, France.

Hanafiah, M. at al. (2015). Towards Developing Collaborative Experience Based Factory Model for Software Development Process in Cloud Computing Environment. *Intern. Review on Comp. and Soft.*, 10(3), 340–350.

Heimannsfeld, K. & Müller, D. (2000). Requirements Engineering Knowledge Management based on STEP AP233. *IMW – Institutsmitteilung*, 25, 73–78.

Hevner, A. (2007). A Three Cycle View of Design Science Research. *Scandinavian Journal of Information Systems*, 19(2), 87–92.

Hevner, A. et al. (2004). Design Science in Information Systems Research. *Mis Quartely*, 28(1), 75–105.

Hickey, A. & Davis, A. (2003). Requirements Elicitation and Elicitation Technique Selection: A Model for Two Knowledge-Intensive Software Development Processes. In: *36th Hawaii International Conference on Systems Sciences*. Hawaii, USA.

Huang, Y. et al. (2005). Development of an expert system for tackling the public's perception to climate-change impacts on petroleum industry. *Expert Systems with Applications*, 29(4), 817–829.

Humayoun, M. & Qazi, A. (2015). Towards Knowledge Management in RE Practices to Support Software Development. *Jou. of Software Eng. and Applications*, 8, 407–418.

Ibarra, Y. & Valle, R. (2017). Knowledge management - A way to contribute to risk decision making in video game development. *Revista Antioqueña de las Ciencias Computacionales y la Ingeniería de Software (RACCIS)*, 7(1), 68–73.

ISO (1994). ISO 10303-1:1994. *Industrial automation systems and integration — Product data representation and exchange* — Part 1: Overview and fundamental principles. Intern. Organization for Standardization.

Jackson, M. (1995). The world and the machine. In: *17th international conference on software engineering*. Seattle, USA.

Johnson, J. (2006). How does AP233 support a Systems Engineering process (e.g. ANSI/EIA-632)? An update. In: *16th Annual International Symposium INCOSE*. Swindon, UK.

Kess, P. & Haapasalo, H. (2002). Knowledge Creation through a Project Review Process in Software Production. *International Journal of Production Economics,* 80(1), 49–55.

Khalid, S., Shehryar, T. & Arshad, S. (2015). The Role of Knowledge Management in Global Software Engineering. In: *International Conference on Industrial Engineering and Operations Management*. Dubai, Dubai.

Lambe, P. (2007). *Organizing Knowledge and Organizational Effectiveness*. Chandos: Oxford.

Li, J. (2007). Sharing Knowledge and Creating Knowledge in Organizations: The Modeling, Implementation, Discussion and Recommendations of Web-Log Based Knowledge Management. In: *International Conference on Service Systems and Service Management*. Chengdu, China.

Mason, D. & Pauleen, D. (2003). Perceptions of knowledge management: A qualitative analysis. *Journal Knowledge Management,* 7(4), 38–48.

Medeni, T. et al. (2011). Tacit knowledge extraction for software requirements specification: A proposal of research methodology design and execution for knowledge visualization. In: *55th Annual Meeting of the ISSS*. Hull, UK.

Mohamed, A. (2010). Facilitating tacit-knowledge acquisition within requirements engineering. In: *10th International Conference on Applied Computer Science*. Iwate, Japan.

Mohamed, B., Basir, G. & Salam, R. (2015). Tacit requirements elicitation framework. *ARPN Journal of Engineering and Applied Sciences*, 10(2), 572-578

Natali, A. & Falbo, R. (2002). Knowledge Management in Software Engineering Environments. In: *XVI Brazilian Symposium on Software Engineering*. Vitoria, Brazil.

Nonaka, I. & Takeuchi, H. (1995). The knowledge-creating company: How Japanese companies foster creativity and innovation for competitive advantage. *Harvard Business Review,* 69(6), 96–104.

Olmos, K. & Rodas, J. (2014). KMoS-RE: knowledge management on a strategy to requirements engineering. *Requirements Engineering*, 19(4), 421–440.

Olmos, K. & Rodas, J. (2014). KMoS-RE: Knowledge Management on a Strategy to Requirements Engineering. *Requirements Engineering*, 19, 421–440.

Olmos, K. & Rodas, J. (2016). A Strategy of Requirements Engineering for Informally Structured Domains. *International Journal of Combinatorial Optimization Problems and Informatics*, 7(2), 49–56.

Olmos, K., Ordas, J. (2013). A Strategy to Requirements Engineering Based on Knowledge Management. *Requirements Engineering*, 19, 421–440.

Pavanasam, V. et al. (2010). Knowledge Based Requirement Engineering Framework for Emergency Management System. In: *Second Intern. Conf. on Computer Engineering and Applications*. Bangalore, India.

Pilat, L. & Kaindl, H. (2011). A Knowledge Management Perspective of Requirement Engineering. In: *5th Intern. Conference on Research Challenges in Information Science*. Gosier, Guadalupe.

Pilat, L. & Kaindl, H. (2011). A Knowledge Management Perspective of Requirements Engineering. In *Fifth International Conference on Research Challenges in Information Science*. Gosier, Guadeloupe.

Pilat, L., Kaindl, H. (2011). A knowledge management perspective of requirements engineering. In: *Fifth Intern. Conference on Research Challenges in Information Science*. Gosier, Guadeloupe.

Ratchev, S. et al. (2005). Knowledge-enriched Requirement Specification for One-of-a-kind Complex Systems. *Concurrent Engineering*, 13(3), 171–183.

Rocha, Á. (2011). Evolution of information systems and technologies maturity in healthcare. *International Journal of Healthcare Information Systems and Informatics*, 6(2), 28–36.

Rodríguez, O. et al. (2004). Applying Agents to Knowledge Management in Software Maintenance Organizations. In: *Workshop on Agent-Mediated Knowledge Management*. Valencia, Spain.

Rouse, W., Cannon, J. & Salas, E. (1992). The Role of Mental Models in Team Performance in Complex Systems. *IEEE Transactions on Systems, Man, and Cybernetics*, 22(6), 1296–1308.

Rus, I. & Lindvall, M. (2002). Knowledge Management in Software Engineering. *IEEE Software*, 19(3), 26–38.

Schmid, K. (2009). Reasoning on Requirements Knowledge to Support Creativity. In: *Second International Workshop on Managing Requirements Knowledge*. Washington, USA.

Scott, W. & Cook, S. (2003). An Architecture for an Intelligent Requirements Elicitation and Assessment Assistant. In: *13th Annual International Symposium INCOSE*. Crystal City, USA.

Shadbolt, N. & Smart, P. (2015). Knowledge elicitation. In: Wilson, J. & Sharples, S. (Eds.), *Evaluation of human work* (pp. 163-196). CRC Press: USA.

Stoiber, R. & Glinz, M. (2009). Modeling and managing tacit product line requirements knowledge. In: *Second Intern. Workshop on Managing Requirements Knowledge*. Atlanta, USA.

Stone, A. & Sawyer, P. (2006). Identifying tacit knowledge-based requirements. *IEEE Proceedings-Software*, 153(6), 211–218.

Supakkul, S. et al. (2009). Capturing, Organizing, and Reusing Knowledge of NFRs: An NFR Pattern Approach. In: *Second International Workshop on Managing Requirements Knowledge*. Washington, USA.

Talib, A. et al. (2010). MASK-SM: Multi-Agent System Based Knowledge Management System to Support Knowledge Sharing of Software Maintenance Knowledge Environment. *Computer and Information Science*, 3(2), 52–78.

Tuunanen, T. (2003). A New Perspective on Requirements Elicitation Methods. *Journal of Information Technology Theory and Application,* 5 (3), 45–62.

Uenalan, O. et al. (2008). Using Enhanced Wiki-based Solutions for Managing Requirements. In: *First Inter. Workshop on Managing Requirements Knowledge*. Barcelona, Spain.

Valero, R., Boza, A. & Vicens, E. (2013). Decision Support Systems for Incident Management. In: *7th Intern. Conf. on Industrial Engineering and Industrial Management*. Valladolid, Spain.

Vasconcelos, J. et al. (2009). A Knowledge-Engine Architecture for a Competence Management Information System. In: *Proceedings UK Academy for Information Systems Conference* (Paper 14). Oxford, UK.

Vásquez, D. et al. (2012). Combining software engineering elicitation technique with the knowledge management lifecycle. *International Journal of Knowledge Society Research*, 3(1), 1–13.

Vásquez, D. et al. (2014). Knowledge management acquisition improvement by using software engineering elicitation techniques. *Computers in Human Behavior*, 30, 721–730.

Venkatesh, K. & Kumar, P. (2013). An Efficient Risk Analysis based Risk Priority in Requirement Engineering using Modified Goal Risk Model. *International Journal of Computer Applications*, 73(14), 15–25.

Wan, J. et al. (2010). Research on knowledge creation in software requirement development. *Journal of Software Engineering and Applications*, 3(5), 487-494.

Ward, J. & Aurum, A. (2004). Knowledge Management in Software Engineering - Describing the Process. In: *Australian Software Engineering Conference*. Melbourne, Australia.

White, S. (2010). Application of Cognitive Theories and Knowledge Management to Requirements Engineering. In: *4th Annual IEEE of Systems Conference*. San Diego, USA.

Wongthongtham, P. & Chang, E. (2007). Ontology Instantiations for Software Engineering Knowledge Management. In: *IEEE International Symposium on Industrial Electronics*. Vigo, Spain.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 15/09/2018
Aceitação/Acceptance: 17/11/2018

Estrategias de gamificación con fines de mejora de procesos software en la gestión de proyectos

Liliana Machuca-Villegas[1,2], Gloria Piedad Gasca-Hurtado[2]

liliana.machuca@correounivalle.edu.co, gpgasca@udem.edu.co

[1] Universidad del Valle, Calle 13 # 100-00, 760032, Cali, Valle del Cauca, Colombia.

[2] Universidad de Medellín, Carrera 87 No. 30-65, 50026, Medellín, Colombia.

Pages: 142–155

Resumen: En el contexto de la Ingeniería de Software, en el que se presentan dificultades relacionadas con la gestión de proyectos de software, la gamificación se constituye como una herramienta de apoyo para la mejora de sus procesos. En este sentido, a través de estrategias de gamificación, se busca aplicar nuevos enfoques de trabajo que ayuden a solventar parte de estas dificultades. Por tal razón, este artículo presenta un estudio de estrategias de gamificación con fines de mejora en los procesos de gestión de proyectos de software. Las estrategias se han clasificado de acuerdo con las áreas de gestión de proyectos de software, el impacto esperado y los elementos de gamificación.

Palabras-clave: Estrategia de gamificación; mejora de procesos de software; gestión de proyectos software.

Gamification strategies for the purpose of improving software processes in project management

Abstract: In the context of Software Engineering, where there are difficulties related to the software project management, gamification is constituted as a support tool for the improvement of its processes. In this way, through gamification strategies, we seek to apply new work approaches that help to solve part of these difficulties. For this reason, this paper presents a study of gamification strategies for improving software project management processes. The strategies have been classified according to the areas of software project management, the expected impact and the gamification elements.

Keywords*:* Gamification strategy; Software Process Improvement; Software Project Management.

1. Introducción

Hoy en día, la gamificación está siendo utilizada en diferentes contextos de trabajo, para hacer el desarrollo de las actividades más atractivas y divertidas (Pedreira, García,

Brisaboa, & Piattini, 2015) y promover habilidades como: la motivación, el compromiso, el rendimiento y la colaboración en los participantes. La intención de la gamificación es usar los elementos de juego para influir en un cambio en el comportamiento de un usuario (Engedal, 2015).

En el contexto de la Ingeniería de Software, en el que se presentan dificultades relacionadas con la gestión de proyectos de software (Standish Group, 2015), (Dutra, Prikladnicki, & França, 2015), (Lombriser, Dalpiaz, Lucassen, & Brinkkemper, 2016), la gamificación se constituye como una herramienta de apoyo para la mejora de procesos.

El proceso de desarrollo de software es considerado como una actividad centrada en las personas (Pirzadeh, 2010),(Yilmaz, 2013). Por lo tanto, la gamificación puede ser utilizada para la implementación de estrategias que susciten el desarrollo de factores sociales y humanos. Dichos factores intervienen en la gestión de los proyectos y de esta manera es posible contribuir con el éxito y calidad de los proyectos.

Algunos estudios exploratorios de la gamificación en la gestión de proyectos de software (Machuca-Villegas & Gasca-Hurtado, 2018a), (Machuca-Villegas & Gasca-Hurtado, 2018b) han reportado resultados que relacionan las áreas gestión de proyectos de software con intervención de la gamificación como estrategia de mejora. Las áreas reportadas en estos estudios han sido identificadas a partir del Project Management Body of Knowledge (PMBOK) (Project Management Institute, 2013). Algunas de las más destacadas son: la gestión de la integración, la gestión de los recursos, la gestión del alcance, la gestión de la calidad y la gestión del riesgo.

Con el propósito de hacer una completa exploración de la gamificación y su influencia en la mejora de los procesos de software en el contexto de la gestión de proyectos, se amplía una revisión exploratoria (Machuca-Villegas & Gasca-Hurtado, 2018b) para obtener resultados relacionados con estrategias de gamificación implementadas en las áreas de gestión de proyectos. El estudio identifica aquellas estrategias de gamificación que aportan a disminuir las causas de fracaso de los proyectos de desarrollo de software relacionadas con los procesos de gestión.

Estas estrategias han sido obtenidas del análisis de 34 estudios primarios obtenidos de la revisión sistemática de literatura (RSL) descrita en (Machuca-Villegas & Gasca-Hurtado, 2018b). Las estrategias se han clasificado de acuerdo con las áreas de gestión de proyectos de software, el impacto esperado por estas estrategias y los elementos de gamificación. Los tipos de estrategia de gamificación que se han identificado se categorizan en: herramientas de software, marcos de trabajo, modelos, métodos, juegos serios, prototipos, entre otros, siendo la categoría de herramientas de software la más utilizada.

El resto del artículo está estructurado de la siguiente manera: en la sección II se presentan algunos trabajos relacionados con el tema; la sección III describe la metodología para la identificación y clasificación de las estrategias; en la sección IV se incluyen los resultados; en la sección V se presenta la discusión de resultados; y finalmente, en la sección VI se plantean las conclusiones y el trabajo futuro.

2. Trabajo relacionado

La gamificación busca incorporar mecánicas y elementos de juego en contextos de no juego, con el fin de mejorar el compromiso, la motivación y el rendimiento de los usuarios, haciendo más atractivas sus tareas (Pedreira et al., 2015). Es un área en la que se han ido desarrollando estudios que promueven su aplicación en contextos de Ingeniería de Software. Se han identificado estudios relacionados con revisiones sistemáticas de literatura (RSL) que permiten evidenciar una creciente evolución en el uso de la gamificación en la mejora de los procesos de Ingeniería de Software.

La revisión sistemática realizada por (Pedreira et al., 2015) reúne elementos importantes que permiten ver que la gamificación es un campo de investigación preliminar. La mayoría de los estudios que encontraron han sido publicados en conferencias o workshops, enfocados al desarrollo de software. También, se encontraron vacíos en el apoyo metodológico para la aplicación de la gamificación en el contexto de los procesos de software. De igual manera, (Hernández, Muñoz, Mejia, & Peña, 2016) presentan una revisión sistemática de la aplicación de la gamificación en equipos de trabajo en la Ingeniería de Software. Sus conclusiones permiten ver que la gamificación en la industria de desarrollo de software enfocada en la creación de equipos de trabajo es incipiente. Se ve reflejado en que la mayoría de los trabajos han sido realizados en un ambiente académico.

Se han desarrollado diferentes trabajos en los que se aprecia la utilidad de la gamificación relacionados con la mejora de procesos software, tales como: (Herranz, Colomo-palacios, Seco, & Sánchez-Gordón, 2016) proponen un framework gamificado para la mejora de procesos de software, orientado en las necesidades de la organización. (Steffens & Marczak, 2016) proponen un framework gamificado como un motivador para la colaboración en equipos de desarrollo de software.

Otros trabajos se enfocan en la ingeniería de requisitos, como el caso de (Lombriser et al., 2016) con una propuesta para mejorar el compromiso y rendimiento de los stakeholders. Otros buscan mejorar el intercambio de la información en código, aplicando técnicas de gamificación como una forma de motivar a los desarrolladores para aplicar las convenciones de codificación (Prause & Jarke, 2015b). (Yilmaz & O'Connor, 2016) presentan un enfoque de gamificación integrado a la metodología híbrida Scrumban para guiar la mejora de procesos de software.

Los trabajos mencionados contribuyen con sus resultados para continuar buscando soluciones nuevas e innovadoras en la mejora de los procesos de software, en la gestión proyectos software y usando la gamificación como estrategia para mejorar. Este interés por el uso de la gamificación en la mejora de procesos y la gestión de proyectos de software favorece el desarrollo de habilidades de tipo colaborativo, de integración y liderazgo en un equipo de trabajo. Sin embargo, aún quedan líneas de trabajo futuro por desarrollar relacionadas con la gestión de los proyectos y en su enfoque en la industria de software.

3. Materiales y métodos

A través de la RSL presentada en (Machuca-Villegas & Gasca-Hurtado, 2018b) se han obtenido 49 estudios base. Esta base permite hacer la extracción y análisis de datos con

el fin de identificar y clasificar las estrategias de gamificación utilizadas. Se establecieron los siguientes criterios de selección para facilitar el proceso:

1. Excluir estudios de tipo filosófico. Estos trabajos se orientan a revisiones de literatura y estudios reflexivos que carecen de una estrategia de gamificación en sí mismos.
2. Incluir los estudios que den respuesta a las preguntas de evaluación indicada en la RSL:
3. PE1: ¿El estudio se enfoca en la aplicación de la gamificación en un área de gestión de proyectos de software?
4. PE2: ¿El estudio refleja un impacto de la gamificación en aspectos como la comunicación, compromiso, colaboración, motivación, entre otros?
5. Incluir aquellos estudios que tengan una contribución o aporte concreto y claridad en la presentación de la estrategia de gamificación.

De la aplicación de estos criterios se obtuvieron 34 estudios que constituyen la muestra de este trabajo. A partir de estos estudios se llevó a cabo el análisis de revisión. El proceso de selección y análisis de los estudios se describe en la Figura 1.

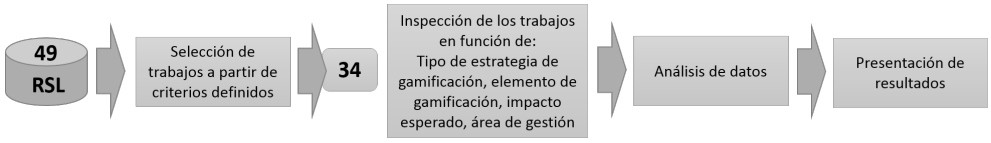


Figura 1 – Proceso de identificación y análisis de estrategias de gamificación

En el contexto de este trabajo, se considera como una estrategia de gamificación a la forma en que se llevó a cabo la implementación de la gamificación como mejora de algún proceso de gestión de proyectos de software. Para efecto de una clasificación de las estrategias identificadas, se tipificaron de la siguiente forma: a) herramienta de software; b) juegos; c) modelos; d) marcos de trabajo; e) métodos; f) prototipo y g) otros.

El análisis de los datos y la presentación de los resultados se realizó en función de los siguientes parámetros: a) clasificación de las estrategias de gamificación; b) relación impacto esperado por las estrategias y elementos de gamificación y c) métodos empleados para medir el impacto esperado.

4. Resultados

La clasificación de los estudios primarios en función de los tipos de estrategia de gamificación y las áreas de gestión de proyecto se presenta en la Tabla 1. Según (Machuca-Villegas & Gasca-Hurtado, 2018a) y (Machuca-Villegas & Gasca-Hurtado, 2018b) las áreas de gestión de proyectos de software que han sido intervenidas con la gamificación son gestión de la integración, gestión de recursos, gestión del alcance, gestión de la calidad y gestión del riesgo. Teniendo en cuenta estas áreas de gestión se presentan los resultados.

Tipo de estrategia	Áreas de gestión de proyectos de software					
	Recursos	**Calidad**	**Riesgos**	**Integración**	**Alcance**	**Total**
Herramienta de software	(Sukale & Pfaff, 2014) (Prause & Jarke, 2015a) (Diniz, Silva, Gerosa, & Steinmacher, 2017) (Snipes, Nair, & Murphy-Hill, 2014)		(Moreta, Gamboa, & Palacios, 2016)	(Jurado, Fernandez, & Collazos, 2015) (Herranz, Colomo-Palacios, & Seco, 2014) (Ašeriškis & Damaševičius, 2014) (Sharma & Kaulgud, 2016) (Sammut, Seychell, & Attard, 2011) (Parizi, 2016) (Sharma, Kaulgud, & Duraisamy, 2016)	(Busetta et al., 2017) (Snijders et al., 2015) (Ribeiro, Farinha, Pereira, & Mira da Silva, 2014) (Fernandes et al., 2012) (Lombriser et al., 2016)	17
Marco de trabajo	(Yilmaz & O'Connor, 2016)			(García, Pedreira, Piattini, Cerdeira-Pena, & Penabad, 2017) (Sammut et al., 2011) (Parizi, 2016)	(Piras, Giorgini, & Mylopoulos, 2016)	4
Modelo	(Butgereit, 2016)				(Lombriser et al., 2016)	2
Método				(Sharma et al., 2016) (Passos, Medeiros, Neto, & Clua, 2011)	(Kumar & Krishnamurthi, 2016) (Busetta et al., 2017)	4
Juego computadora	(Schafer, 2017) (Calderón, Ruiz, & O'Connor, 2017) (Calderón & Ruiz, 2016)			(Maxim, Kaur, Apzynski, Edwards, & Evans, 2016)		4
Juego escritorio	(Hof, Kropp, & Landolt, 2017)	(Soska, Mottok, & Wolff, 2017)		(Gasca-hurtado, Gómez-Alvarez, Muñoz, & Mejía, 2016)		3
Prototipo (propuesta)	(De Melo et al., 2014)				(Unkelos-Shpigel & Hadar, 2016)	2
Otro	(Akpolat & Slany, 2014)		(Moreta et al., 2016)			2
Total	12	1	2	13	10	38

Tabla 1 – Clasificación tipos de estrategias con respecto a las áreas de gestión

Con el fin de conocer el impacto esperado de la implementación de las estrategias de gamificación en las diferentes áreas de gestión, se hizo una clasificación de estrategias

teniendo en cuenta las siguientes categorías: 1) Motivación; 2) Compromiso; 3) Comunicación; 4) Rendimiento; 5) Colaboración; 6) Interacción equipo; 7) Formación/ Entrenamiento/Aprendizaje y 8) Participación. En la Tabla 2 se presenta dicha clasificación. Se tuvo en cuenta aquel impacto cuyo número de estudios fuera superior a 2.

Tipo de estrategia	**Impacto esperado**							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Herramienta software	11	8	2	5	7	1	1	4
Marco de trabajo	4	3		2		1		
Modelo	1	1		1			1	
Método	2	2		2	1	2		1
Juego de computadora	2						4	
Juego de escritorio	1	1	1				3	
Prototipo		1			2			
Otro	1	1						
TOTAL	**22**	**17**	**3**	**10**	**10**	**4**	**9**	**5**

1. Motivación; 2. Compromiso; 3. Comunicación; 4. Rendimiento; 5. Colaboración; 6. Interacción equipo; 7. Formación/Entrenamiento/ Aprendizaje; 8. Participación

Tabla 2 – Clasificación tipos de estrategias y el impacto esperado

Los estudios también fueron clasificados según la estrategia y el elemento de gamificación utilizado. De esta relación se obtuvieron los resultados que se muestran en la Tabla 3, donde se presenta la cantidad de estudios clasificados a partir de la relación indicada. Los elementos de gamificación que se tuvieron en cuenta son los que se clasificaron en (Machuca-Villegas & Gasca-Hurtado, 2018b). Es preciso indicar que un estudio se puede relacionar con uno o más elementos de gamificación.

Elemento de gamificación	**Estrategia de gamificación**							
	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**
Leaderboard	8	2	1	1			1	
Sistema de puntos	14	3	1	3	1		1	1
Insignias	6	2	1	1			1	
Barra de progreso	5		1	1				1
Recompensa	7	1	1	2	1		2	1
Retos	5	1	2	1				1
Logros	2	1		1				
Retroalimentación	3	1		1			1	
Ranking	4							
Juego serio					5	2		
Misiones	2	2						
Avatar	2	2	1					
Puntajes	2						1	
Dashboard	2	1		1				
Niveles	5	1	1	1	1			

A. Herramienta software; B. Marco de trabajo; C. Modelo; D. Método; E. Juego de computadora; F. Juego Escritorio; G. Prototipo; H. Otro

Tabla 3 – Clasificación tipos de estrategias y elementos de gamificación

Además, se relacionaron los elementos de gamificación y el impacto esperado por cada estrategia de gamificación identificada. Esta relación se muestra en la Tabla 4, donde se presenta la cantidad de estudios clasificados a partir de la relación indicada. Es preciso decir que un elemento de gamificación se puede relacionar con más de un tipo de impacto.

Elemento de gamificación	Impacto esperado							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Leaderboard	6	9	1	4	5			2
Sistema de puntos	15	12	1	9	6	3	3	5
Insignias	5	7		5	1	2		1
Barra de progreso	5	4	1	2	4		1	2
Recompensa	9	7	2	3	6	2	2	2
Retos	7	8		2	2		2	
Logros	3	3		3				
Retroalimentación	3	4		1	1		1	
Ranking	4	1	1	1	3			3
Juego serio	3	1	1				7	
Misiones	4	4		2		1		
Avatar	3	5		3		1		
Puntajes	2		2		3			2
Dashboard	2	1		4	1	2		
Niveles	5	7		5			1	

1. Motivación; 2. Compromiso; 3. Comunicación; 4. Rendimiento; 5. Colaboración; 6. Interacción equipo; 7. Formación/Entrenamiento/ Aprendizaje; 8. Participación

Tabla 4 – Clasificación elementos de gamificación y su relación con el impacto esperado

Otro aspecto que se analizó en este trabajo es la forma en que los estudios midieron el impacto de la aplicación de la gamificación. Esta clasificación refleja el grado de rigor con el cual se realiza la validación de estrategias. El 58.82% midió el impacto con algún método en particular y el 41.18% no evidencia un método de validación. En la Tabla 5. Se presentan los resultados de los métodos empleados.

Método para medir el impacto	Cantidad
Experimento	3
Cuestionarios	4
Caso de studio	9
Estudio preliminar o exploratorio	3
Prueba piloto	1

Tabla 5 – Métodos empleados para medir el impacto esperado

5. Discusión

5.1. Clasificación de las estrategias de gamifiación

Las estrategias de gamificación se han clasificado de acuerdo con las áreas de gestión de proyectos de software, el impacto esperado y los elementos de gamificación (ver Tablas 1-3). En estos resultados, prevalecen las herramientas de software como tipo de estrategia más utilizado. Los tipos de herramientas de software que predominan en los estudios analizados son aplicaciones web, plataformas y módulos. Este tipo de estrategia también se utiliza como medio para implementar las propuestas de marcos de trabajo, modelos o métodos que en algunos estudios fueron reportados.

Una de las áreas de gestión en donde se facilita el uso de cualquier tipo de estrategia es el área de gestión de recursos. Según los resultados, es el área que presenta más versatilidad en el uso de los tipos de estrategia identificados. En el área de gestión del alcance, en la que se han reportado trabajos relacionados con la elicitación de requisitos, se evidencia un amplio uso de diferentes tipos de estrategias. Sin embargo, a diferencia del área de gestión de recursos, en el área de gestión del alcance no se evidencian estrategias relacionadas con juegos. Para esta área hay una tendencia hacia las herramientas de software, modelos y marcos de trabajo. Esta tendencia puede estar relacionada con que la mayoría de los estudios encontrados en esta área están enfocados a la industria, más que a la academia. Al igual que las áreas de gestión de recursos y de alcance, el área de gestión de la integración también presenta diversidad en el uso de las estrategias empleadas.

En cuanto a los resultados de la clasificación de las estrategias de gamificación y el impacto esperado (Tabla 2), se destacan la motivación, el compromiso, el rendimiento, la colaboración y la formación como los factores de medición del impacto más empleados en los estudios. Sin embargo, no todos los estudios reportan resultados del alcance del impacto proyectado en sus estrategias de gamificación. Por otra parte, en los estudios se identifica el planteamiento de soluciones que llevan a fomentar este tipo de habilidades en profesionales o estudiantes. Además, se logró identificar la necesidad de promover la adopción de prácticas, herramientas, mejorar la productividad, brindar entrenamiento o formación al equipo de trabajo, motivar al desarrollador para la documentación de software, facilitar la participación de los usuarios en la elicitación de requisitos, mejorar el intercambio de información y comunicación, entre otras necesidades.

En el caso de la herramienta de software los trabajos asociados a ella, se inclinan hacia el desarrollo de la motivación, el compromiso y la colaboración, especialmente. En consecuencia, este tipo de estrategia es común emplearlo para promover habilidades conocidas como *habilidades blandas*. Por otra parte, se evidencia que los tipos de estrategias relacionados con los juegos se enfocan principalmente hacia la formación. Con los marcos de trabajo se busca desarrollar la motivación, el compromiso y el rendimiento, al igual que el método y el modelo en una menor escala.

En cuanto a la clasificación de las estrategias de gamificación y su relación con los elementos de gamificación (Tabla 3), se puede observar que el tipo de estrategia herramienta de software emplea todos los elementos de gamificación identificados, a excepción del juego serio. Esto deja ver que, una herramienta de software cuenta con un grado de flexibilidad alto para utilizar cualquier tipo de elemento de gamificación.

Generalmente, a través de ellas, se implementan las propuestas de gamificación. A su vez, los elementos de gamificación que predominan en las estrategias identificadas fueron: sistema de puntos, tabla de posiciones, insignias, recompensas, retos y niveles. Sin embargo, cuando se usa un juego como estrategia, se emplea generalmente un juego serio.

5.2. Relación impacto esperado y elementos de gamificación

Con el fin de identificar la relación entre el impacto esperado por la estrategia de gamificación y los elementos de gamificación, se realizó un mapeo que permitió identificar los elementos utilizados para alcanzar el impacto esperado con la implementación de la estrategia de gamificación propuesta (Tabla 4). Estos resultados muestran que el impacto asociado con la motivación puede ser trabajado con cualquiera de los elementos de gamificación. Es la única categoría de impacto que se mapeó con todos los elementos y que tiene más de una incidencia en sus trabajos. Para el caso del compromiso, todos los elementos, a excepción del puntaje, fueron utilizados por las estrategias de gamificación para medir el impacto esperado.

Para aquellas estrategias que buscaron mejorar el rendimiento en los participantes los elementos como: sistemas de puntos, insignias, niveles y tablas de posiciones son los que presentan una alta frecuencia de uso. Para la medición de factores como la colaboración, los elementos que más predominan para medir el impacto son el sistema de puntos, las recompensas, la tabla de posiciones y la barra de progreso.

Es preciso indicar que, las estrategias de gamificación que usaron más de un elemento de gamificación para implementar dicha estrategia, obtienen resultados más favorables cuando se mide el impacto. Las combinaciones más comunes entre elementos de gamificación reportadas en los estudios analizados son el uso de recompensas con sistema de puntos, desafíos o tabla de posiciones. También aquellas en las que se integra logros, con insignias, puntos, niveles y tablas de posiciones. Por otra parte, los elementos que se mantienen constantes y con una frecuencia de los estudios analizados mayor de 4, son sistema de puntos, recompensa, retos, tabla de posiciones, insignias, barras de progreso y niveles. Finalmente, se puede decir, que el sistema de puntos y las recompensas pueden influir en todas las categorías de impacto especificadas.

5.3. Métodos empleados para medir el impacto esperado

Otro de los objetivos de este trabajo se orienta en identificar los métodos de medición del impacto de las estrategias de gamificación. Los resultados de la Tabla 5 muestran que el método más utilizado fue el caso de estudio, seguido de los cuestionarios, experimentos y estudios preliminares.

Sin embargo, se destaca que el 41,18% de los estudios seleccionados no presentó una evidencia de la evaluación o medición del impacto de la estrategia de gamificación propuesta. Este porcentaje permite inferir que la medición o evaluación de las estrategias de gamificación es una línea de trabajo futuro por abordar.

6. Conclusiones

En este artículo se presenta un conjunto de estrategias de gamificación que han sido obtenidas del análisis de 34 estudios primarios obtenidos de la revisión sistemática de literatura de (Machuca-Villegas & Gasca-Hurtado, 2018b). Las estrategias han sido clasificadas de acuerdo con las áreas de gestión de proyectos de software descritas en el PMBOK. También, se ha hecho un análisis del impacto esperado y los elementos de gamificación utilizados en la realización de estrategias de gamificación en el área de gestión de proyectos. Adicionalmente, se estableció una relación entre el impacto esperado por la estrategia y los elementos de gamificación, y se identificó el método utilizado para medir dicho impacto.

El tipo de estrategia más utilizado para implementar las propuestas de gamificación en el contexto de gestión de proyectos de software es la herramienta de software, en la que se incluyen aplicaciones de software, plataformas o módulos. Generalmente, aquellos estudios que tienen como estrategia de gamificación a un marco de trabajo, un método o un modelo emplean una aplicación software o juegos para implementarlos.

El impacto esperado en las propuestas de gamificación de los estudios analizados se orientó en su gran mayoría hacia la promoción de la motivación, compromiso, rendimiento, colaboración y formación de los participantes, tanto en un ambiente académico como industrial. A su vez, los elementos de gamificación que más se utilizaron en la mayoría de las estrategias son sistema de puntos, tabla de posiciones, insignias, recompensas, retos y niveles.

Finalmente, con el objetivo de continuar trabajando en este campo de acción, se proyecta la realización de una propuesta doctoral en la que se plantee un modelo de gamificación apoyado en factores sociales y humanos que influyen en la productividad de equipos de desarrollo de software.

Referencias

Akpolat, B. S., & Slany, W. (2014). Enhancing Software Engineering Student Team Engagement in a High-Intensity Extreme Programming Course using Gamification. In *27th Conference on Software Engineering Education and Training (CSEE&T)* (pp. 149–153). IEEE.

Ašeriškis, D., & Damaševičius, R. (2014). Gamification of a Project Management System. In *The Seventh International Conference on Advances in Computer-Human Interactions Gamification* (pp. 200–207). https://doi.org/10.1145/2181037.2181040

Busetta, P., Kifetew, F. M., Munante, D., Perini, A., Siena, A., & Susi, A. (2017). Tool-Supported Collaborative Requirements Prioritisation. *Proceedings - International Computer Software and Applications Conference*, *1*, 180–189. https://doi.org/10.1109/COMPSAC.2017.243

Butgereit, L. (2016). Gamifying mobile micro-learning for continuing education in a corporate IT environment. *2016 IST-Africa Conference, IST-Africa 2016*, 1–7. https://doi.org/10.1109/ISTAFRICA.2016.7530597

Calderón, A., & Ruiz, M. (2016). Coverage of ISO/IEC 12207 Software Lifecycle Process by a Simulation-Based Serious Game, *155*, 59–70. https://doi.org/10.1007/978-3-642-21233-8

Calderón, A., Ruiz, M., & O'Connor, R. V. (2017). Coverage of ISO/IEC 29110 Project Management Process of Basic Profile by a Serious Game. In *European Conference on Software Process Improvement* (pp. 111–122).

De Melo, A. A., Hinz, M., Scheibel, G., Diacui Medeiros Berkenbrock, C., Gasparini, I., & Baldo, F. (2014). Version control system gamification: A proposal to encourage the engagement of developers to collaborate in software projects. In *International Conference on Social Computing and Social Media* (Vol. 8531 LNCS, pp. 550–558). Springer, Cham. https://doi.org/10.1007/978-3-319-07632-4_52

Diniz, G. C., Silva, M. A. G., Gerosa, M. A., & Steinmacher, I. (2017). Using gamification to orient and motivate students to contribute to oss projects. *Proceedings - 2017 IEEE/ACM 10th International Workshop on Cooperative and Human Aspects of Software Engineering, CHASE 2017*, 36–42. https://doi.org/10.1109/CHASE.2017.7

Dutra, A. C. S., Prikladnicki, R., & França, C. (2015). What do we know about high performance teams in software engineering? Results from a systematic literature review. In *41st Euromicro Conference on Software Engineering and Advanced Applications*. https://doi.org/10.1109/SEAA.2015.24

Engedal, J. Ø. (2015). Gamification - a study of motivational affordances. *Department of Computer Science and Media Technology*, *1*, 81.

Fernandes, J., Duarte, D., Ribeiro, C., Farinha, C., Pereira, J. M., & Da Silva, M. M. (2012). IThink : A game-based approach towards improving collaboration and participation in requirement elicitation. *Procedia Computer Science*, *15*, 66–77. https://doi.org/10.1016/j.procs.2012.10.059

García, F., Pedreira, O., Piattini, M., Cerdeira-Pena, A., & Penabad, M. (2017). A framework for gamification in software engineering. *Journal of Systems and Software*, *132*, 21–40. https://doi.org/10.1016/j.jss.2017.06.021

Gasca-hurtado, G. P., Gómez-Alvarez, M. C., Muñoz, M., & Mejía, J. (2016). Gamification Proposal for Defect Tracking in Software Development Process Gloria. In *Communications in Computer and Information Science* (pp. 212–224). https://doi.org/10.1007/978-3-662-43896-1

Hernández, L., Muñoz, M., Mejia, J., & Peña, A. (2016). Gamificación en equipos de trabajo en la ingeniería de software : Una revisión sistemática de la literatura Gamification in software engineering teamworks : A systematic literature review.

Herranz, E., Colomo-Palacios, R., & Seco, A. de A. (2014). Gamiware: A Gamification Platform for Software Process Improvement. *Communications in Computer and Information Science*, *425*, 13–24. https://doi.org/10.1007/978-3-662-43896-1

Herranz, E., Colomo-palacios, R., Seco, A. de A., & Sánchez-Gordón, M.-L. (2016). Towards a Gamification Framework for Software Process Improvement Initiatives : Construction and Validation, *22*(12), 1509–1532.

Hof, S., Kropp, M., & Landolt, M. (2017). Use of Gamification to Teach Agile Values and Collaboration. *Proceedings of the 2017 ACM Conference on Innovation and Technology in Computer Science Education - ITiCSE '17*, 323–328. https://doi.org/10.1145/3059009.3059043

Jurado, J. L., Fernandez, A., & Collazos, C. A. (2015). Applying gamification in the context of knowledge management. *Proceedings of the 15th International Conference on Knowledge Technologies and Data-Driven Business - i-KNOW '15*, *1*, 1–4. https://doi.org/10.1145/2809563.2809606

Kumar, B. S., & Krishnamurthi, I. (2016). Improving User Participation in Requirement Elicitation and Analysis by Applying Gamification Using Architect's Use Case Diagram, *49*, 471–482. https://doi.org/10.1007/978-3-319-30348-2

Lombriser, P., Dalpiaz, F., Lucassen, G., & Brinkkemper, S. (2016). Gamified requirements engineering: Model and experimentation. In *International Working Conference on Requirements Engineering: Foundation for Software Quality* (Vol. 9619, pp. 171–187). Springer, Cham. https://doi.org/10.1007/978-3-319-30282-9_12

Machuca-Villegas, L., & Gasca-Hurtado, G. P. (2018a). Gamification for improving Software Project: Systematic mapping in project management. In *2018 13th Iberian Conference on Information Systems and Technologies (CISTI)*. IEEE. https://doi.org/10.23919/CISTI.2018.8399415

Machuca-Villegas, L., & Gasca-Hurtado, G. P. (2018b). Gamification for Improving Software Project Management Processes: A Systematic Literature Review. In M. Jezreel, M. Muñoz, Á. Rocha, A. Peña, & M. Pérez-Cisneros (Eds.), *Trends and Applications in Software Engineering. CIMPS 2018. Advances in Intelligent Systems and Computing* (pp. 41–54). Springer, Cham. https://doi.org/ https://doi.org/10.1007/978-3-030-01171-0_4

Maxim, B. R., Kaur, R., Apzynski, C., Edwards, D., & Evans, E. (2016). An agile software engineering process improvement game. *2016 IEEE Frontiers in Education Conference (FIE)*, 1–4. https://doi.org/10.1109/FIE.2016.7757682

Moreta, L. L., Gamboa, A. C., & Palacios, M. G. (2016). Implementing a Gamified application for a Risk Management course. *2016 IEEE Ecuador Technical Chapters Meeting (ETCM)*, 1–6. https://doi.org/10.1109/ETCM.2016.7750858

Parizi, R. M. (2016). On the gamification of human-centric traceability tasks in software testing and coding. *2016 IEEE/ACIS 14th International Conference on Software Engineering Research, Management and Applications, SERA 2016*, 193–200. https://doi.org/10.1109/SERA.2016.7516146

Passos, E. B., Medeiros, D. B., Neto, P. A. S., & Clua, E. W. G. (2011). Turning real-world software development into a game. *Brazilian Symposium on Games and Digital Entertainment, SBGAMES*, 260–269. https://doi.org/10.1109/SBGAMES.2011.32

Pedreira, O., García, F., Brisaboa, N., & Piattini, M. (2015). Gamification in software engineering - A systematic mapping. *Information and Software Technology*, *57*(1), 157–168. https://doi.org/10.1016/j.infsof.2014.08.007

Piras, L., Giorgini, P., & Mylopoulos, J. (2016). Acceptance Requirements and Their Gamification Solutions. *Proceedings - 2016 IEEE 24th International Requirements Engineering Conference, RE 2016*, 365–370. https://doi.org/10.1109/RE.2016.43

Pirzadeh, L. (2010). Human Factors in Software Development : A Systematic Literature Review. *Master of Science Thesis in Computer Science and Engineering*, 368.

Prause, C. R., & Jarke, M. (2015a). Gamification for enforcing coding conventions. *Proceedings of the 2015 10th Joint Meeting on Foundations of Software Engineering - ESEC/FSE 2015*, 649–660. https://doi.org/10.1145/2786805.2786806

Prause, C. R., & Jarke, M. (2015b). Gamification for Enforcing Coding Conventions. *Proceeding of the 10th Joint Meeting of the European Software Engineering Conference and the ACM SIGSOFT Symposium on the Foundations of Software Engineering (ESEC/FSE 2015)*, 649–660. https://doi.org/10.1145/2786805.2786806

Project Management Institute. (2013). *A Guide to the Project Management Body of Knowledge*. *Project Management Journal* (Fifth). Pennsylvania: Project Manager Institude.

Ribeiro, C., Farinha, C., Pereira, J., & Mira da Silva, M. (2014). Gamifying requirement elicitation: Practical implications and outcomes in improving stakeholders collaboration. *Entertainment Computing*, *5*(4), 335–345. https://doi.org/10.1016/j.entcom.2014.04.002

Sammut, R., Seychell, D., & Attard, N. (2011). Gamification of Project Management within a Corporate Environment. *Proceedings of the 15th International Academic MindTrek Conference on Envisioning Future Media Environments - MindTrek '11*, 9. https://doi.org/10.1145/2181037.2181040

Schafer, U. (2017). Training scrum with gamification: Lessons learned after two teaching periods. *IEEE Global Engineering Education Conference, EDUCON*, (April), 754–761. https://doi.org/10.1109/EDUCON.2017.7942932

Sharma, V. S., & Kaulgud, V. (2016). Agile workbench: Tying people, process, and tools in distributed agile delivery. *Proceedings - 11th IEEE International Conference on Global Software Engineering, ICGSE 2016*, 69–73. https://doi.org/10.1109/ICGSE.2016.17

Sharma, V. S., Kaulgud, V., & Duraisamy, P. (2016). A gamification approach for distributed agile delivery. *Proceedings of the 5th International Workshop on Games and Software Engineering - GAS'16*, 42–45. https://doi.org/10.1145/2896958.2896966

Snijders, R., Dalpiaz, F., Brinkkemper, S., Hosseini, M., Ali, R., & Özüm, A. (2015). REfine: A gamified platform for participatory requirements engineering. *1st International Workshop on Crowd-Based Requirements Engineering, CrowdRE 2015 - Proceedings*, 1–6. https://doi.org/10.1109/CrowdRE.2015.7367581

Snipes, W., Nair, A. R., & Murphy-Hill, E. (2014). Experiences gamifying developer adoption of practices and tools. *Companion Proceedings of the 36th International Conference on Software Engineering - ICSE Companion 2014*, 105–114. https://doi.org/10.1145/2591062.2591171

Soska, A., Mottok, J., & Wolff, C. (2017). Pattern Oriented Card Game Development, (April), 1170–1177.

Standish Group. (2015). 2015 Chaos Report. Standish Group International Inc.

Steffens, F., & Marczak, S. (2016). *A Gamification Framework as a Collaboration Motivator for Software Development Teams*. Pontificial Catholic University of Rio Grande do Sul.

Sukale, R., & Pfaff, M. S. (2014). QuoDocs: Improving developer engagement in software documentation through gamification. *Proceedings of the Extended Abstracts of the 32nd Annual ACM Conference on Human Factors in Computing Systems - CHI EA '14*, 1531–1536. https://doi.org/10.1145/2559206.2581263

Unkelos-Shpigel, N., & Hadar, I. (2016). Inviting everyone to play: Gamifying collaborative requirements engineering. *5th International Workshop on Empirical Requirements Engineering, EmpiRE 2015 - Proceedings*, 13–16. https://doi.org/10.1109/EmpiRE.2015.7431301

Yilmaz, M. (2013). *A Software Process Engineering Approach to Understanding Software Productivity and Team Personality Characteristics: An Empirical Investigation Murat*. https://doi.org/10.1145/291712.291759

Yilmaz, M., & O'Connor, R. V. O. (2016). A Scrumban integrated gamification approach to guide software process improvement: a Turkish case study. *Tehnicki Vjesnik - Technical Gazette*, *23*(1), 237–245. https://doi.org/10.17559/TV-20140922220409

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 07/10/2018
Aceitação/Acceptance: 13/12/2018

Caracterización de factores que influyen en la baja producción científica de las universidades usando análisis de redes sociales

Jéssica Ponce[1], Ariosto Vicuña[1], Orlando Erazo[1], Eduardo Samaniego[1]

jessica_ponceord@hotmail.es, {avicuna, oerazo, esamaniego}@uteq.edu.ec

[1] Universidad Técnica Estatal de Quevedo, Av. Quito Km 1.5, 120301, Quevedo, Ecuador.

Pages: 156–169

Resumen: Con el paso de los años, las estructuras sociales se han transformado en combinaciones complejas dando lugar a un incremento en el análisis de sus conductas. Precisamente, un problema referente a ellas puede encontrarse en las instituciones de educación superior donde la baja producción científica tiene un impacto significativo en los indicadores de calidad. Ante esta situación, en este artículo se describe un estudio sobre las complejidades no visibles en la baja producción científica de universidades, aplicando el análisis de redes sociales como herramienta para descubrir patrones de dicha producción. Para el efecto, se construyeron cinco redes diferentes en base a información personal, de trabajo, de publicaciones de contenidos académicos, de dirección de tesis y de estudios de pregrado y postgrado de docentes universitarios. En base a la interpretación de cada una de estas redes se extrajeron los factores que tienen influencia en las publicaciones científicas realizadas.

Palabras-clave: redes sociales; producción científica; análisis de redes.

Characterization of factors that influence on the low scientific production of universities using social network analysis

Abstract: Over the years, social structures have been transformed into complex combinations, leading to an increase in the analysis of their behavior. Precisely, a problem concerning them can be found in higher education institutions where low scientific production has a significant impact on quality indicators. Given this situation, this article describes a study on the not visible complexities in the low scientific production of universities, applying the analysis of social networks as a tool to discover patterns of such production. To do it, five different networks were built based on personal information, work, academic content publications, thesis management, and undergraduate and postgraduate studies of university professors. Based on the interpretation of each of these networks, the factors that influence the scientific publications carried out were extracted.

Keywords*:* social networks; scientific production; network analysis.

1. Introducción

Con el devenir de los años, las estructuras sociales se han transformado en combinaciones complejas dando lugar a un incremento en el análisis de sus conductas. Básicamente, una estructura social es un conjunto limitado de elementos estructurales como individuos, grupos comunitarios, valores, instituciones y grupos asociativos. Estas estructuras presentan conductas susceptibles de ser analizadas. Este análisis resulta de interés porque permite entender el mundo desde la realidad y dar cuenta de ello. Para el efecto, se puede recurrir al uso de técnicas que ayuden en el estudio de la información en redes (como gestión del conocimiento, relaciones comunitarias, relaciones de productividad, entre otros).

Precisamente, el análisis de redes sociales (ARS) es una de esas técnicas que se puede utilizar para evaluar la actuación de estructuras sociales. Los estudios correspondientes se realizan entre actores, ya sean estos humanos, organizaciones económicas, profesionales, de diferentes clases sociales entre otros (Ronald L, 2004). Así, el ARS puede ser empleado para caracterizar de manera eficiente los diversos comportamientos de la estructura social observada. De manera similar, el ARS sirve para tratar de explicar fenómenos sociales cualitativos (Velásquez Rodríguez, 2013, pp. 145-146). Además, es posible graficar la forma en que se llevan a cabo los fenómenos sociales utilizando algoritmos diseñados para tal fin (como por ejemplo Kamada–Kawai) (Vargas Quesada & De Moya Anegón, 2007).

El ARS permite también responder de forma efectiva ante sucesos que pueden ocurrir dentro de un contexto. Esto ha permitido utilizarlo en diferentes áreas como seguridad, salud y educación. Jhon & Carrintong (Jhon & Carrintong, 2011) utilizaron el ARS para modelar organizaciones sociales delincuenciales. A partir de ese modelamiento analizaron los efectos de la delincuencia en los adolescentes, y en menor medida en el crimen de los adultos, basándose fundamentalmente en la teoría del control social. Por otra parte, el ARS ha sido aplicado en el campo de la salud pública para estudiar la transmisión de enfermedades (Ramos-Vidal, 2015). En este ámbito, las redes se usan para estudiar las conductas de salud y la estructura inter-organizativa de los servicios ofrecidos que producen información de notable importancia para descubrir problemas asociados a ese sector. En cuanto al uso educativo de las redes sociales, éstas pueden ayudar a analizar las actitudes y expectativas de alumnos universitarios (Espuny Vidal, et al., 2011). No obstante, las redes sociales podrían ser aprovechadas también para resolver otro tipo de problemas en el ámbito educativo.

Precisamente, un problema referente a las estructuras sociales que puede encontrarse en varias instituciones de educación superior (IESs) es la baja producción científica (Rodríguez Morales, et al., 2016) evidenciada fundamentalmente como publicaciones de artículos científicos y libros. Incluso, dicha producción puede considerarse como deficiente y estacionaria en algunos casos (Caballero Rico, et al., 2012). Esta "escasez" científica puede deberse a la baja colaboración entre pares (Gonzales Salda, et al., 2017), a una capacidad limitada para adquirir y utilizar el conocimiento, entre otras. Además, con el paso del tiempo la investigación se hace más compleja, multidisciplinaria, costosa y se distribuye geográficamente, lo que en algunos casos dificulta o impide su realización (Pérez Beltrán, et al., 2015). Finalmente, la falta de producción científica en las IESs tiene también un impacto significativo en los indicadores educativos de calidad.

Ante esta situación, en este artículo se describe un estudio sobre las complejidades no visibles en la baja producción científica de IESs, aplicando el ARS como herramienta para descubrir patrones de dicha producción dentro del entorno universitario. Para el efecto, se construyeron cinco redes diferentes con los datos recolectados en una IES. Los resultados permitieron determinar las relaciones existentes en los docentes que podrían tener una influencia en las publicaciones científicas realizadas. Aunque la propuesta del uso del ARS es aplicada a una universidad en particular, este estudio constituye un referente para realizar análisis similares en otras IESs, que contribuyan a tomar las decisiones adecuadas para mejorar la producción científica de sus docentes.

2. Metodología

2.1. Redes seleccionadas

Los estudios de redes sociales requieren que se defina la cantidad de redes a analizar. Esta cantidad es determinada por las características propias de los actores involucrados en el estudio; es decir, los límites del estudio se definen alrededor de los actores que constituyen la red (Hanneman & Riddle, 2005). Como los sujetos que constituyen las redes de interés son docentes universitarios, éstas deben corresponder a las características profesionales y personales de ellos. Además, estas características deben estar en concordancia con aquellos aspectos a considerar cuando se desea analizar la baja producción científica universitaria (García Zárate, 2003). Por lo tanto, las redes a ser analizadas en este trabajo son las siguientes: (1) Información personal: se estudian datos como estado civil, edad, hábitos de lectura de los docentes entre otros. (2) Trabajo: muestra si la carga horaria de los docentes, la realización de otros tipos de labores y las bases de datos de consulta para su labor de docencia tienen influencia en la escasa producción científica. (3) Dirección de tesis: permite estudiar la relación que hay entre la dirección de tesis y la no publicación de contenidos científicos. (4) Publicaciones de contenidos académicos: estudia variables como la elaboración de libros con ISBN y la no publicación de los mismos. (5) Estudios de pregrado y postgrado: muestra la importancia de estudios en los docentes para la realización de contenidos científicos.

2.2. Recolección de datos

El primer paso para la creación y análisis de las redes consistió en un estudio descriptivo transversal basado en encuestas para recolectar los datos necesarios. Las encuestas fueron diseñadas para obtener datos de los docentes tanto a nivel profesional como personal. Se consideraron variables como edad, enfermedades, acceso a internet, lugar de residencia, nivel de instrucción, libros, ponencias, investigaciones realizadas, artículos científicos, direcciones de tesis, entre otros. Estas variables permitieron crear los vínculos o lazos existentes entre los elementos involucrados para así explicar su comportamiento (Requena Santos, 1996).

La recolección de datos fue llevada a cabo en una universidad ecuatoriana, acreditada y categoriza en el tercero de cinco niveles[1]. Todos los profesores titulares de cada una

[1] Categoría vigente al momento de realizar el estudio. Actualmente, dicha universidad de encuentra en la segunda de cuatro categorías

de las cinco facultades existentes (Ciencias Empresariales, de la Ingeniería, Agrarias, Ambientales y Pecuarias) completaron las encuestas en papel. Además, como este estudio de redes sociales busca involucrar a todos los sujetos y sus relaciones no se puede mencionar una población y/o una muestra.

2.3. Creación de redes

Una vez recolectados los datos, se crearon las cinco redes seleccionadas. Esta etapa se llevó a cabo empleando la aplicación MTVAL (Matriz validadora) desarrollada por los autores. MTVAL (figura 1) permite la creación de redes, nodos y relaciones que vinculan los nodos en las redes sociales y genera un archivo compatible con otras aplicaciones afines como Pajek.

Como siguiente paso, la visualización de las redes se realizó utilizando Pajek (Batagelj & Mrvar, 1998) versión 4.10 (disponible en http://vlado.fmf.uni-lj.si/pub/networks/pajek/). Pajek es un software libre enfocado al análisis de redes sociales que incluye la opción de visualizarlas gráficamente. Este software ya ha sido aplicado antes en otras investigaciones. Por ejemplo, Pajek fue usado en un estudio sobre células madre que buscaba comparar "las líneas de investigación que ofrece el análisis de distintos tipos de descriptores, según su naturaleza documental, a partir de su aparición conjunta en los documentos" (Cantos Mateos, et al., 2013). En otro trabajo, Pajek se utilizó para analizar la importancia del uso de indicadores relacionales en los estudios de producción científica sobre grupos de investigación con el fin de determinar las dinámicas que aplican estos grupos (Rodríguez Gutiérrez & Gómez Velasco, 2017). Estos trabajos conducen a considerar que Pajek permite tener una mejor representación en redes grandes utilizando los algoritmos de visualización apropiados (Batagelj & Mrvar, 2011).

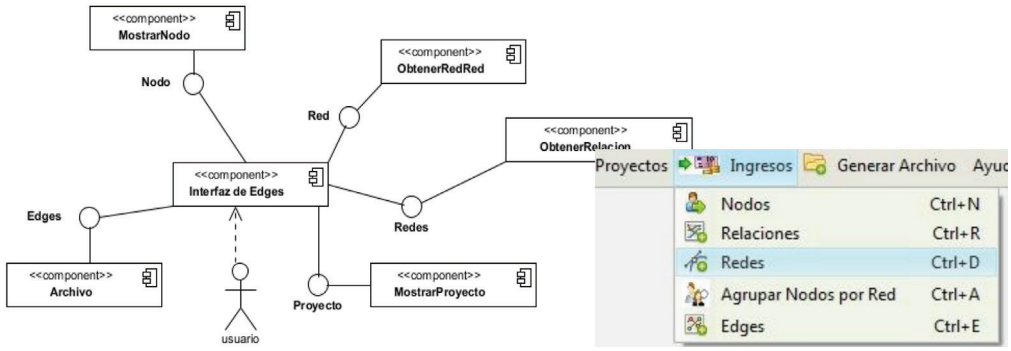


Figura 1 – MTVAL: Diagrama de componentes y ventana administración de MTVAL

Uno de los algoritmos implementados en Pajek es Kamada-Kawai (Kamada & Kawai, 1989). Este algoritmo está basado en grafos que permiten la visualización de diseños mostrando el comportamiento de una determinada situación (Vargas Quesada & De Moya Anegón, 2007). En nuestro caso, Kamada-Kawai permitió distribuir los nodos de la red de tal manera que aquellos que se encontraban conectados aparecieran más cerca unos de otros, con una ubicación hacia el centro de la red, mientras que los que

no estaban conectados tendieran a ir hacia el exterior. Esta característica ofrecida por Kamada-Kawai, junto con Pajek, contribuye a la realización de un mejor análisis de los actores que conforman la red (Gil Mendienta & Schmidt, 2002).

2.4. Indicadores

Para el análisis de la red propiamente dicho se emplearon varios indicadores. Primero, la **densidad (density)** de la red, que puede ser calculada sin necesidad de utilizar un software específico. Para ello se divide el número de relaciones (n) entre las posibles relaciones a existir (N) multiplicado por 100. Para obtener las posibles relaciones se multiplica el número total de nodos (n) por el número de nodos menos uno (Valázquez Álvarez & Aguilar Gallegos, 2005).

Otra medida utilizada es el **grado de centralidad (*Degree centrality*)** de la red**.** Este indicador permite saber el impacto que tiene un nodo central en la red que está siendo analizada basándose en la capacidad de conexión de dicho nodo. El nodo con mayor conexión es aquel que tiene mayor capacidad de influir en los demás (Yan & Ding, 2009). Para conocer la centralidad se miden las entradas y las salidas de cada nodo. Las entradas se calculan como la suma de las relaciones que los actores dicen tener sobre el resto, mientras que las salidas corresponden a la suma de las relaciones referidas hacia un actor por otros (Valázquez Álvarez & Aguilar Gallegos, 2005).

El tercer indicador aplicado es el **grado de intermediación (*betweenness*)**, que es la posibilidad que tiene un nodo o actor para controlar las comunicaciones entre pares de nodos. Se calculó usando los caminos geodésicos entre todos los pares posibles. Un camino geodésico es la ruta más corta que un actor debe seguir para llegar a otros nodos. El grado de intermediación se obtiene al contar las veces en que los nodos aparecen en los caminos geodésicos conectados entre sí (Valázquez Álvarez & Aguilar Gallegos, 2005). Con este indicador se analizó la importancia de los actores en las redes conociendo la proximidad entre ellos, lo que a su vez permitió evaluar la capacidad de comunicación existente entre los actores y la red (Polanco, 2008).

Finalmente, el **grado de cercanía (*closeness*)** se empleó para medir "la distancia de un nodo para alcanzar al conjunto de los otros (si es de salida) y de ser alcanzado por los otros (si es de entrada)" (Prat Forga & Cànoves Valiente, 2013). Los valores altos de una red indican una gran capacidad de los nodos al conectarse con los actores en la red (Valázquez Álvarez & Aguilar Gallegos, 2005).

3. Resultados e interpretación

En esta sección se presentan los resultados obtenidos por medio de la metodología descrita anteriormente. Al mismo tiempo se analiza la información obtenida en cada una de las redes anteriormente mencionadas.

3.1. Red información personal

La red de información personal (figura 2) muestra los nodos con mayor impacto en la no publicación de artículos científicos de los docentes. En general, se puede notar que la relación de la lectura, el sueldo, los problemas de salud y la edad con la producción

científica. La tabla 1 por su parte muestra los valores de enlaces, cercanías y frecuencia de los nodos que intervienen en la red de información personal. Se visualiza el número de los nodos, los nombres de los nodos y los valores de los indicadores. Estos últimos ponen de manifiesto la importancia y la influencia que tienen los nodos en la red de información personal. Además, el nodo más influyente para el indicador grado de centralidad es "no publica" con 163 enlaces lo que significa que la mayoría de docentes no realiza publicaciones científicas. Esto no se debe precisamente a que no "les guste leer" ya que en la columna cercanía este nodo tiene un valor alto (0,63643807); es decir, cuenta con una gran relación hacia los demás nodos. Por lo tanto, nuestros resultados sugieren que la no publicación está relacionada con los medios donde los docentes realizan sus lecturas (como periódicos y libros por encima de los artículos científicos).

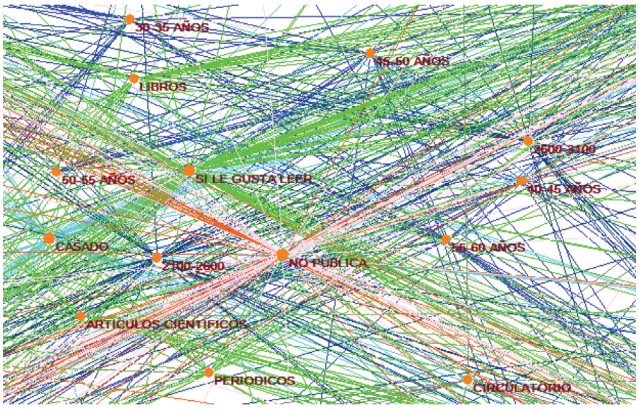


Figura 2 – Red información personal

Nº	Nodos	Centralidad	Cercanía	Intermediación
204	No publica	163	0,54375292	0,266763197
179	Si le gusta leer	127	0,63643807	0,300462297
165	Ingresos mensuales $2100-$2600	52	0,40584457	0,041436756
160	Casado	87	0,4849052	0,112223435
181	Lecturas en periódicos	39	0,37970543	0,017504406
182	Lecturas en libros	39	0,38229727	0,020349021
185	Lecturas en artículos científicos	35	0,38229727	0,023424902
157	55-60 años de edad	24	0,35113825	0,007845055
152	30-35 años de edad	22	0,35559714	0,01179288
154	40-45 años de edad	22	0,34786677	0,005312101
155	45-50 años de edad	22	0,34678979	0,008228019
169	Problema de salud Circulatorio	19	0,35113825	0,00545666
	Densidad de la red: 0,04205930 = 4,2%			

Tabla 1 – Valores de los indicadores para la red de información personal

Por otro lado, la columna intermediación indica que, después de los nodos “si publica” y “le gusta leer”, el nodo “casado” tiene una frecuencia de 0,112223435 en comparación con los demás nodos. Esto significa que aunque este nodo no presente una frecuencia alta en comparación con el nodo “le gusta leer”, tiene influencia en la no publicación de contenido científico. Finalmente, la densidad de la red es baja (4,2%), lo que sugiere que hay características específicas en la parte personal de los docentes que influyen en la no publicación de artículos científicos. Cabe notar que densidades como esta se encuentran normalmente en redes grandes debido a que existe una gran cantidad de conexiones entre los nodos con mayor influencia.

3.2. Red trabajo

La figura 3 muestra una porción de la red trabajo. Esta red pone de manifiesto la relación entre la carga horaria, la realización o no de otro tipo de labores y las fuentes de consulta para labores docentes con escasa producción científica.

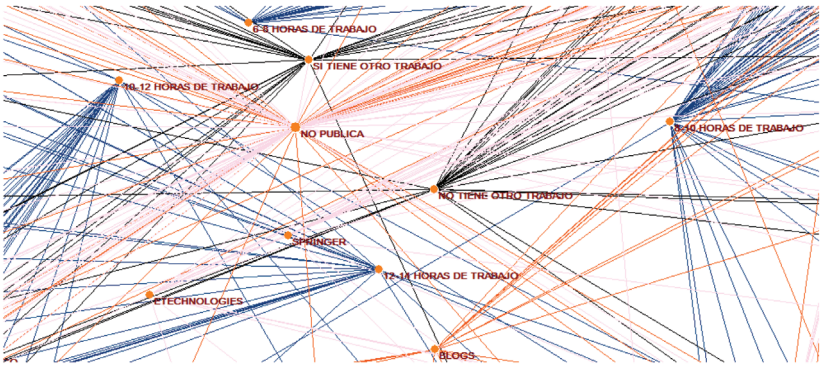


Figura 3 – Red trabajo

La tabla 2 detalla los valores que tienen los nodos en cada uno de los indicadores. El indicador grado de centralidad muestra que la carga horaria de “8-10 horas de trabajo” tiene 49 enlaces. Este valor indica que existe una asociación fuerte de dicha carga laboral con la no publicación de contenido científico. Además, existe una alta cercanía por parte del nodo “si tiene otro trabajo” (0,44897959), lo que puede significar que, al poseer otro trabajo, los docentes no tienen tiempo suficiente para realizar producciones de tipo científicas. También, una de las mayores fuentes de consultas para la labor de docencia se la realiza en los Blogs, con un valor en el indicador intermediación de 0,036102709, mismo que comparado con los valores de los nodos Springer y Etechnologies tiene una frecuencia alta. Esto permite llegar a la conclusión que las fuentes de consultas empleadas por los docentes influyen en la realización de publicaciones de carácter científico. Por último, la densidad de la red es de 2,6% lo que indica que las características con mayor peso de la red se corresponden entre ellas. En base a esto se puede decir que la carga horaria, como las labores de trabajo aparte de la docencia y las fuentes de consultas en la labor docencia, tienen un gran impacto en la escasa publicación de artículos científicos.

Nº	Nodos	Centralidad	Cercanía	Intermediación
174	No publica	163	0,56896552	0,646283736
154	8-10 horas de trabajo	49	0,39520958	0,16534511
151	No tiene otro trabajo	31	0,37078652	0,052562601
155	10-12 horas de trabajo	31	0,35044248	0,064717799
150	Si tiene otro trabajo	30	0,44897959	0,063702562
156	12-14 horas de trabajo	27	0,34920635	0,076044394
153	6-8 horas de trabajo	23	0,33616299	0,047484658
157	Blogs	19	0,34554974	0,036102709
175	Base de datos Springer	15	0,33962264	0,021465196
179	Base de datos Etechnologies	15	0,33962264	0,023747894
Densidad de la red: 0,02598853 = 2,6%				

Tabla 2 – Valores de los Indicadores para la red trabajo

3.3. Red dirección de tesis

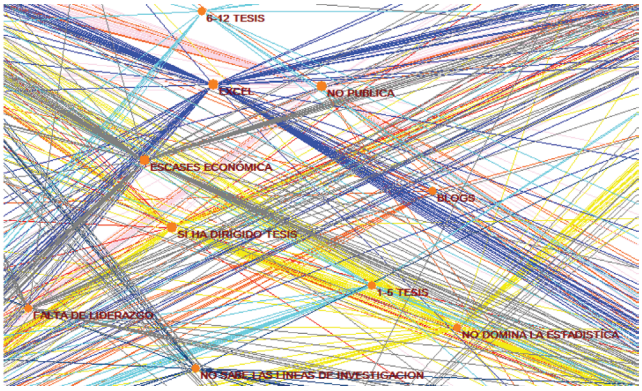


Figura 4 – Red dirección de tesis

En esta red (figura 4) se aprecia que la mayoría de docentes tiene una cantidad considerable de tesis de grado dirigidas, mismas que oscilan entre 1 y 12 tesis. Esta red demuestra también que los docentes no conocían las líneas de investigación de sus facultades al momento de la investigación. Asimismo, una gran parte de los docentes realizaban los cálculos estadísticos con herramientas no idóneas (como MS Excel). Además, entre las causas que los docentes consideran importantes para que no se haya desarrollado investigación están el aspecto económico, la falta de liderazgo y las publicaciones de docencia realizadas a través de blogs.

Los aspectos observados en la figura 4 se complementan con los valores de los indicadores recogidos en la tabla 3. El nodo con 163 enlaces corresponde a “no publica” lo que se

debe al gran número de docentes que no realizan publicaciones de contenido científico. Por otro lado, aunque el nodo referente a la dirección de tesis tiene un valor considerable (159 enlaces), los docentes no necesariamente conocen las líneas de investigación de su institución (cercanía de 0,3533123 con los demás nodos). En cuanto a la cantidad de tesis dirigidas, el indicador grado de centralidad muestra que la dirección de "1 a 5 tesis" (29 enlaces) tiene una mayor relación en comparación con la dirección de "6 a 12 tesis" (22 enlaces).

Nº	Nodos	Centralidad	Cercanía	Intermediación
174	No publica	163	0,56896552	0,646283736
151	Si ha dirigido tesis	159	0,54368932	0,298374463
152	No dominan la estadística	135	0,46666667	0,150353823
157	Excel	83	0,47457627	0,080173362
166	1-5 tesis	29	0,36721312	0,011270619
167	6-12 tesis	22	0,35555556	0,006783996
177	Escases económica	92	0,51141553	0,107639446
175	Falta de liderazgo	37	0,37966102	0,014230412
172	No sabe las líneas de investigación	24	0,3533123	0,005748372
183	Blogs	19	0,35109718	0,005028656
	Densidad de la red: 0,04607143= 4,6%			

Tabla 3 – Valores de los Indicadores para la red dirección de tesis

3.4. Red publicación de contenidos académicos

Esta red (figura 5) revela que la publicación de libros con ISBN es una deficiencia en la mayoría de docentes, característica importante para comprender la escasa publicación de contenido científico académico. Además, se puede ver que una cantidad considerable de docentes no han realizado publicaciones en revistas indexadas y no indexadas. De hecho, existen también docentes que no han publicado artículos científicos y mucho menos libros.

La información visualizada en la red es analizada a partir de los valores que tienen los indicadores de los nodos con mayor peso de acuerdo a lo presentado en la tabla 4. La densidad que presenta la red publicación de contenidos académicos es del 4,6%, el cual es un porcentaje bajo debido a que las relaciones entre los nodos involucrados en la baja producción científica tienen un alto grado de conexión. Por otro lado, los docentes que no tienen libros ni de artículos científicos publicados tienen una frecuencia de 0,158257727 correspondiente al indicador grado de intermediación. El nodo "No ISBN" presenta un valor alto de 143 enlaces en la columna centralidad; es decir, hay una gran cantidad de docentes sin publicaciones con ISBN. Además, el no disponer de publicaciones en revistas indexadas y no indexadas se ve reflejado en el valor del indicador grado de cercanía.

Figura 5 – Red de publicación de contenidos académicos

Nº	Nodos	Centralidad	Cercanía	Intermediación
168	No publica	163	0,563247271	0,187505112
193	No ISBN	143	0,742462312	0,263239886
195	No revista indexada	117	0,586856438	0,136707863
197	No revista no indexada	121	0,620284969	0,158257727
198	No publicación libro ni artículo científico	83	0,482783375	0,068553473
199	Si publicación libro y artículo científico	67	0,45163606	0,081299766
Densidad de la red: 0,04558144 = 4,6%				

Tabla 4 – Valores de los Indicadores para la red dirección de tesis

3.5. Red de estudio de pregrado y postgrado

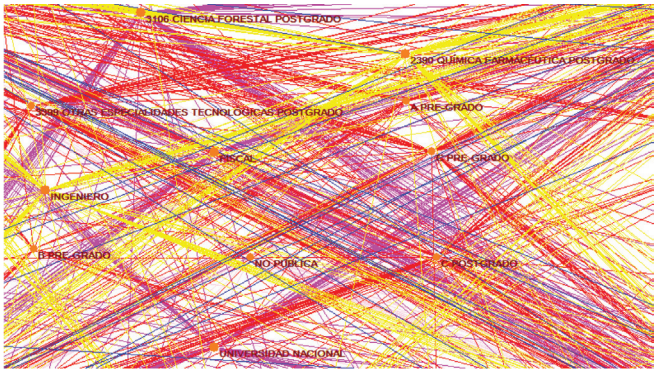


Figura 6 – Red de estudios pregrado y postgrado

La red de pregrado y postgrado, presentada en la figura 6, muestra la importancia de los estudios de tercer nivel en las publicaciones de contenidos científicos. En la relación

del tipo y categoría de universidades en las que realizaron sus estudios los docentes, se observa que en su mayoría estudiaron en universidades nacionales, estatales y que se encuentran en categoría C[2]. A esto hay que añadir que las áreas del conocimiento tales como "3399 otras especialidades tecnológicas (postgrado)", "3106 ciencia forestal" y "2390 química farmacéutica (postgrado)", son las que menos incidencia tienen en la publicación de los artículos científicos.

Nº	Nodos	Centralidad	Cercanía	Intermediación
168	No publica	70	0,40259208	0,03017556
192	Universidades públicas	125	0,56086921	0,13074529
196	C Pregrado	75	0,41833252	0,03853729
194	A Pregrado	34	0,3552888	0,00947807
195	B Pregrado	28	0,35214067	0,00728985
198	Ingeniero	116	0,50215005	0,09776414
252	Universidad Nacional	110	0,51326776	0,09553743
256	C Postgrado	76	0,42022829	0,03401496
270	Ciencia Forestal Postgrado	127	0,55527171	0,13372484
276	Química farmacéutica Postgrado	118	0,5142165	0,10555361
274	Otras especialidades tecnológicas Postg.	76	0,43130407	0,04835152
	Densidad de la red: 0,03859460 = 3,9%			

Tabla 5 – Definición de algunos estilos

La tabla 5 muestra los valores de los indicadores para los nodos que cuentan con mayor influencia en la red y que permiten entender las razones de la baja producción en los docentes. El número de enlaces para el nodo "Universidades públicas" es de 125 lo que significa que los docentes en su mayoría han estudiado en universidades estatales tipo C con un grado de cercanía (0,41833252) más alto con relación a las de categoría A y B.

Para los estudios de postgrado los indicadores muestran que para el nodo de "universidad nacional" hay una frecuencia alta (intermediación = 0,09553743), lo que significa que existe una gran cantidad de docentes que realizaron sus estudios de cuarto nivel en universidades categoría C. Por último, la red muestra que las áreas del conocimiento con menor número de publicaciones de artículos científicos son: Ciencia Forestal Postgrado (0,13372484), Química farmacéutica Postgrado (0,10555361) y otras especialidades tecnológicas (0,04835152).

[2] Categorización vigente al momento de realizar la investigación.

4. Conclusiones

En este artículo se presenta un estudio en el que se utilizan redes sociales para analizar la baja producción científica de docentes universitarios. Para ello se crearon cinco redes en base a información personal, de trabajo, de publicaciones de contenidos académicos, de dirección de tesis y de estudios de pregrado y postgrado de los docentes de una universidad de Ecuador. En base a la interpretación de cada una de estas redes se pueden extraer factores que influyen en la baja producción científica.

Primero, la red información personal revela que los docentes con menos publicaciones científicas son aquellos que presentan hábitos de lectura en periódicos o libros. La edad también es una característica relevante al momento de realizar contenidos científicos, es así que los docentes de 55 a 60 años tienen menos publicaciones que aquellos que están entre los 30 a 50 años. Además, por medio de la relación ubicación, familia, estado civil, edades, fuentes de ingresos y salud, se encontró que las características del contexto personal en que se desenvuelven los docentes tienen un rol importante en la creación de los contenidos científicos.

Segundo, los docentes con menos publicaciones a partir del estudio de la red trabajo son aquellos que tienen de 8 a 10 horas de trabajo mostrado como el nodo más influyente en la red. Las características "tienen otro trabajo" y realizan sus consultas para la labor docencia en "Blogs" también muestran una alta frecuencia para la no publicación de contenidos científicos de los docentes.

Tercero, en la red dirección de tesis los docentes que menos publican son aquellos que han dirigido tesis pero que en su mayoría no dominan la Estadística. A este grupo se suman aquellos que realizan los cálculos estadísticos en Excel e indican que las limitaciones económicas no permiten publicaciones de contenido científico.

Cuarto, la red publicación de contenidos académicos revela que los docentes con publicaciones sin ISBN son los que menos realizan publicaciones científicas.

Por último, los docentes con un menor número de contenidos científicos realizados, según la red de pregrado y postgrado, se encuentran entre los que terminaron sus estudios en universidades estatales de categoría C. Esta red también muestra la influencia de la carrera de pregrado cursada, tales como las de las áreas de Ciencia Forestal y Química Farmacéutica. Esta red pone de manifiesto además la influencia del poco conocimiento del uso de nuevas tecnologías para la realización de cálculos estadísticos en la ejecución de investigaciones científicas y académicas.

Aunque los factores mencionados fueron identificados a partir de los datos recolectados en una universidad, estos pueden tener una utilidad que va más allá del beneficio único de dicha institución. En primer lugar, este estudio muestra cómo emplear el análisis de redes sociales en un problema de la educación superior. Luego, los factores identificados podrían servir como referencia para que instituciones similares a la tomada como referente puedan tomar las decisiones necesarias para mejorar su producción científica. Sin embargo, dadas las diferencias que podrían existir entre universidades, sería útil realizar a futuro una comparación de los factores identificados en más de una universidad. Este estudio permitiría ratificar y/o eliminar/incorporar los

factores que inciden en la problemática de interés. Además, estas diferencias podrían también analizarse realizando los estudios respectivos en más de un periodo de tiempo y comparando tales resultados.

Referencias

Batagelj, V., & Mrvar, A. (1998). Pajek-program for large network analysis. *Connections, 21*(2), 47–57.

Batagelj, V., & Mrvar, A. (24 de septiembre de 2011). *Program for Analysis and Visualization of Large Networks*. Recuperado el 20 de julio de 2016, de http://vlado.fmf.uni-lj.si/pub/networks/pajek/doc/pajekman.pdf

Caballero, F. C., Uresti, R. M., & Ramírez de León, J. (2012). Análisis de la producción científica de la Universidad Autónoma de Tamaulipas y evaluación de su impacto en los indicadores educativos de calidad. *Scielo, 41*(161), 33–51.

Cantos Mateos, G., Zulueta, M. Á., Vargas Quesada, B., & Chinchilla Rodríguez, Z. (2013). Estudio comparativo sobre la visualización de redes de co-words a través de los descriptores del Science Citation Index y de Medline. *Actas I Congresso ISKO Espanha e Portugal* (págs.173-189). Portugal: Univ. do Porto.

Espuny, C., González, J., Lleixà, M., & Gisbert, M. (2011). Actitudes y expectativas del uso educativo de las redes sociales en los alumnos universitarios. *Revista de Universidad y Sociedad del conocimiento, 8*(1), 174.

García Zárate, Ó. A. (2003). *Hacia una nueva universidad en el Perú*. Lima: UNMSM.

Gil Mendienta, J., & Schmidt, S. (2002). *Análisis de redes: aplicaciones en ciencias sociales*. Mexico: UNAM.

Gonzales, J., Chavez, T., Lemus, K., Silva, I., Galvez, T., & Galvez, J. (2017). Producción científica en educación médica en Latinoamérica en Scopus, 2011-2015. *Edumed*, 1–7.

Hanneman, R., & Riddle, M. (2005). *Introduction to social network methods*. Obtenido de http://faculty.ucr.edu/~hanneman/

Jhon, S., & Carrintong, P. (2011). *Sage Handbook of Social Network*. Londres: SAGE Publications Ltd.

Kamada, T., & Kawai, S. (1989). An algorithm for drawing general undirected graphs. *Information processing letters, 31*(1), 7–15.

Pérez, J. E., Ureña, G. V., & Rodríguez, L. (2015). Social network analysis for the study of intellectual production in research groups. *Scielo, 37*(150), 124–142.

Polanco, X. (28 de Junio de 2008). Análisis de redes: introducción.

Prat, J. M., & Cànoves, G. (2013). La centralidad de las administraciones públicas en las redes sociales del turismo industrial. Una comparación entre el Bages -Berguedà y el Haut -Rhin. *Turismo y patrimonio cultural, 11*(4), 513–524.

Ramos-Vidal, I. (2015). Análisis de redes sociales: una herramienta efectiva para evaluar coaliciones comunitarias. *Rev. salud pública, 17*(3), 324.

Requena Santos, F. (1996). Redes sociales y cuestionarios. *Revista de Sociologia*(18), 191–201.

Rodríguez Gutiérrez, J. K., & Gómez Velasco, N. Y. (2017). Redes de coautoría como herramienta de evaluación de la producción. *Informática y documentación, 27*(2), 279–297.

Rodríguez Morales, A., Culquichicón Sánchez, C., & Gil Restrepo, A. F. (2016). Baja producción científica de decanos en facultades de medicina y salud de Colombia: ¿una realidad común en Latinoamérica? *Scielo, 58*(4), 402–403.

Ronald L, B. (2004). The Analisys of social Network. 505–526.

Valázquez Álvarez, A., & Aguilar Gallegos, N. (Junio de 2005). Manual introductorio al análisis de redes sociales. *Medidas de centralidad*. México.

Vargas, B., & De Moya, F. (2007). *Visualizing the Structure of Science*. Springer.

Velásquez, C. A. (Mayo de 2013). *Análisis de redes sociales del alumnado de la Universidad de San Carlos de Guatemala* (Vol. 315). Universidad Almería.

Yan, E., & Ding, Y. (2009). Applying centrality measures to impact analysis: A coauthorship network analysis. *Journal of the American Society for*.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 26/09/2018
Aceitação/Acceptance: 19/11/2018

Propuesta de un test heurístico de accesibilidad para sitios web basados en la norma NTC 5854

Daniela Pérez Oliveros[1], Gabierl E. Chanchi[1], Maria Isabel Vidal[1]

danielaperez@unimayor.edu.co, gchanchi@unimayor.edu.co; mvidal@unimayor.edu.co

[1]Institucion Universitaria Colegio Mayor del Cauca, Cra 7 #2-34, 190003, Popayán, Colombia

Pages: 170–182

Resumen: La accesibilidad web es un aspecto de gran relevancia de cara a garantizar el acceso a la información desde contextos de uso limitado y en personas con limitaciones físicas y/o cognitivas. En este sentido y con el fin de posibilitar la conducción de evaluaciones de accesibilidad en aplicaciones web que permitan determinar los aspectos funcionales a mejorar, no se ha evidenciado la existencia de un test que guíe el proceso de evaluación, teniendo en cuenta la norma de accesibilidad NTC 5854. Así, en este artículo se propone como aporte un test heurístico para la ejecución de evaluaciones heurísticas de accesibilidad en aplicaciones web, tomando como referencia la norma NTC 5854. El instrumento planteado define un conjunto de criterios de accesibilidad asociados a los cuatro principios de la norma: perceptible, operable, comprensible y robusto, los cuales al ser calificados por evaluadores expertos permiten obtener recomendaciones de accesibilidad de manera más adecuada.

Palabras Clave: Accesibilidad, Accesibilidad web, Norma NTC 5854, Test heurístico

Proposal of a heuristic accessibility test for web sites, according to the norm NTC 5854

Abstract: Web accessibility is an aspect of great relevance in order to guarantee access to information from contexts of limited use and in people with physical and/or cognitive limitations. In this sense and in order to facilitate the conduction of evaluations of accessibility in Web applications that allow to determine the functional aspects to improve, it has not been evidenced the existence of a test that guides the evaluation process, taking into account the Accessibility Standard NTC 5854. Thus, this article proposes as a contribution to a heuristic test for the execution of heuristic evaluations of accessibility in Web applications, taking as reference the standard NTC 5854. The proposed instrument defines a set of accessibility criteria associated with the four principles of the standard: perceptible, operable, comprehensible and robust, which, when qualified by expert evaluators, allows to obtain recommendations of Accessibility in a more appropriate way.

Keywords: Accessibility, web Accessibility, Norm NTC 5854, Test heuristic.

1. Introducción

La promoción e implementación de la política de acceso a la información en Colombia está a cargo de la secretaria de Transparencia de la Presidencia de la Republica, el Ministerio de las TIC, entre otras empresas gubernamentales, debido a esto se obliga a las entidades a garantizar que las personas en condición de diversidad funcional y los diferentes grupos étnicos y culturales accedan a información de interés. Así, se busca que el estado entregue a los colombianos la información pública que soliciten, de tal manera que tanto las entidades públicas y privadas además de responder a la solicitud expresa de información de los ciudadanos, también estén en la obligación de divulgarla proactivamente, respondiendo al requerimiento de forma proactiva, actualizada, accesible y comprensible (Digital, 2014).

Teniendo en cuenta lo anterior, en la actualidad la mayoría del contenido de las aplicaciones web disponibles en internet, presentan una gran barrera en el campo de la accesibilidad lo que cada vez más dificulta e imposibilita el acceso y la utilización de la misma por parte de los usuarios que presentan algún tipo de diversidad funcional o limitación de acceso. En este orden de ideas entre más software y aplicaciones web sean accesibles más personas podrán hacer uso de las mismas y así contribuyendo a la reducción de las brechas tecnológicas. Se calcula que más de 1000 millones de personas viven con algún tipo de diversidad funcional, es decir, alrededor del 15% de la población mundial (según estimaciones de la población mundial de 2010). Esta cifra es superior a la estimación anterior de la OMS, que data de los años setenta y rondaba el 10%. De acuerdo con la Encuesta Mundial de Salud, unos 785 millones (15,6%) de personas mayores de 15 años viven con algún tipo de diversidad funcional, mientras que, conforme al estudio Carga Mundial de Morbilidad, esa cifra ascendería a aproximadamente 975 millones (19,4%) (Organizacion mundial de la salud, Banco munduial, 2011). Al hablar de accesibilidad web se está haciendo referencia a un diseño web que permite a usuarios con diversidades funcionales o con alguna limitación de acceso tecnológico, escribir, entender, navegar e interactuar con la web. Entre las diversidades funcionales que aborda la accesibilidad se encuentran: limitaciones visuales, auditivas, físicas, cognitivas, etc., o aquellos que surgen debido al paso de los años. De este modo la accesibilidad busca que las personas de manera independiente a su diversidad funcional, pueden acceder a la información que las instituciones públicas, privadas, o todo tipo de organizaciones proveen (SeoColombia, 2009). La accesibilidad también permite a las organizaciones una realimentación sobre las nuevas necesidades y requerimientos de los usuarios, lo que implica un constante mejoramiento de las páginas web. A nivel mundial se han establecido normas para guiar el diseño, la construcción y poder evaluar la accesibilidad de las páginas web, y en Colombia, el Instituto Colombiano de Normas Técnicas y Certificación, ICONTEC publicó en 2011 la norma para la accesibilidad web NTC 5854, la cual establece los requisitos para que un sitio web sea accesible a la mayor cantidad posible de usuarios independientemente de su condición física (discapacidad visual, auditiva, cognitiva, etc.); calidad de conexión a internet o dispositivo de acceso (Icontec Internacional, 2011) (Barbosa, 2011).

En este orden de ideas, el Ministerio de Tecnología y las Comunicaciones, en su estrategia con el Gobierno en línea, ha dado gran importancia a la accesibilidad, al promover la ejecución y el cumplimiento de la norma NTC 5854 teniendo en cuenta sus

niveles de conformidad y así implementar en los diferentes portales de las instituciones gubernamentales considerando los diferentes temas abordados por dichos portales web, tales como: formularios de descarga, información en audio y video, acceso desde dispositivos móviles, consultas a bases de datos, servicios de interacción, trámites y servicios en línea, entre otros (MinTIC, 2014).

Con el objetivo de establecer los requisitos de accesibilidad que son aplicables a las páginas web, los cuales están agrupados en tres niveles de conformidad: A, AA, y AAA, la norma NTC 5854 fue desarrollada empleando como documento de referencia las pautas de accesibilidad para contenido web (WCAG) 2.0 provistas por el W3C (World Wide Web Consortium). La intención de la norma NTC 5854 es promover que el contenido de la web sea accesible a personas con diversidades funcionales: visuales, auditivas, físicas, de habla, cognitivas, de lenguaje, de aprendizaje o neuro-lógicas, permitiendo que pueden percibir, entender, navegar, interactuar y contribuir con los sitios web. No obstante, es importante destacar que la accesibilidad beneficia a todos los usuarios que visiten el sitio web, ya que hace que los contenidos se presentan de forma comprensible y adecuada para todas las personas (Tecnica, 2012).

A pesar de que existen normas como la NTC 5854, no se ha evidenciado la existencia de instrumentos concretos para la conducción de evaluaciones heurísticas de accesibilidad en aplicaciones web, los cuales permitan verificar el cumplimiento de las pautas y principios considerados en la norma. En el presente trabajo se propone un test heurístico de accesibilidad para sitios web, el cual incluye un conjunto de criterios de cumplimiento expresados en preguntas enfocadas a accesibilidad. Un test heurístico es entendido como una herramienta de apoyo para la ejecución de evaluaciones heurísticas, las cuales están centradas en la inspección del cumplimiento de un conjunto de heurísticas de usabilidad por parte de un grupo de expertos sobre un producto software (Jakob Nilsen, Robert. L Mack, 1994). Partiendo de esto se tiene en cuenta cada una de las pautas asociadas al principio de accesibilidad web definidos en la norma técnica colombiana NTC 5854. Este test heurístico fue planteado tomando como referencia la plantilla propuesta en el test heurístico de usabilidad de Torres Burriel, el cual define un conjunto de preguntas asociadas a criterios de usabilidad agrupadas por diferentes temáticas de usabilidad con el fin de que el evaluador valore cada pregunta con una valoración de 1 a 5 (Torres Burriel, 2008)

Mediante el test heurístico propuesto se realizó una evaluación heurística de accesibilidad para la evaluación al prototipo web www.lauvid.com, mediante el test heurística realizado se obtuvieron un conjunto de recomendaciones para mejorar el desarrollo, el desempeño y la accesibilidad del sitio web. El resto del articulo está organizado de la siguiente forma: en la sección 2 está la metodología propuesta, en la sección 3 se presenta los resultados y discusiones, en la sección 4 se dan a conocer las conclusiones y trabajo futuros y por último en la sección 5 se presenta las referencias para la investigación de este artículo.

2. Metodología Propuesta

Para el desarrollo y la validación del test heurístico de accesibilidad presente en este artículo, se generaron las siguientes fases a saber: exploración de los principios de la norma, obtención de heurísticas, definición de criterios, construcción del test heurístico y el caso de estudio

Fase 1 - Exploración de los principios de la norma: En esta fase se realizó el estudio exhaustivo y detallado de cada una de las pautas, principio, niveles de conformidad y prioridad presentes en la norma NTC 5854, tomando como referencia los conceptos y especificaciones de la norma ISO 9241 literal 171. Así mismo se tuvieron en cuenta algunas de las heurísticas de accesibilidad presentadas en el test heurístico de Torres Burriel. Lo anterior con el fin de retroalimentar y complementar la información obtenida a partir de las pautas de la norma NTC 5854.

Fase 2 Obtención de heurísticas: una vez realizado el de un estudio de la norma NTC 5854 y de las normas y test heurístico de accesibilidad relacionados, se procedió a la obtención de un conjunto de heurísticas de accesibilidad. Las heurísticas fueron obtenidas según los pautas de los diferentes principios de la norma NTC 5854.

Fase 3 y 4 Definición de criterios y construcción de los test heurísticos: Cada criterio, los cuales se presentan en término de preguntas puede ser valorados de 1 a 5, siendo 1 la mínima expresión de la heurística en el portal evaluado y 5 la máxima expresión de la heurística en el portal evaluado. La redacción de cada pregunta del test heurístico fue generada de modo que fuera claro, evitando ambigüedades por parte del evaluador. Así mismo mediante las preguntas asociadas a cada pauta de la norma, se generó un test heurístico la cual tomó como base el test heurístico propuesto por Torres-Burriel (Torres Burriel, 2008).

Fase 5 Caso de estudio: Una vez generado el test heurístico para la conducción de evaluaciones heurísticas, se procedió a escoger evaluadores con conocimientos en las áreas de Ingeniería de Software, Diseño de Software, Usabilidad - Accesibilidad, y Desarrollo de Software. Estos evaluadores aplicaron el instrumento desarrollado sobre el prototipo web la empresa LAUVID SAS, el cual se desarrolló en el marco del proyecto del grupo de investigación TARGET de la Facultad de Ciencias Sociales y Administrativas de la Institución Universitaria Colegio Mayor del Cauca.

2.1. Propuesta del Test Heurístico

Con base en el test heurístico de usabilidad propuesto por Torres-Burriel se realizó una adaptación de dicha plantilla, enfocándose en las pautas y principios de la norma NTC 5854 (Perceptible, Operable, Comprensible y Robusto) organizando la platilla de la siguiente: en la primera hoja se establecieron los datos personales de los evaluadores, una pequeña explicación sobre los niveles de conformidad y el rango de calificación por cada criterio, en la segunda hoja se encuentran las heurísticas del principio perceptible, en la tercera hoja las heurísticas del principio operable, en la cuarta hoja las heurísticas de comprensible y en la quita las heurísticas del principio robusto. Finalmente, en la hoja seis se encuentran el consolidado de los cuatro principios evaluados.

Como complemento a cada una de las preguntas asociadas a las heurísticas y principios de accesibilidad generados, se vincularon los niveles de conformidad y/o prioridad obtenidos a partir de cada una de las pautas de la norma exploradas. A continuación, se presentan las escalas de valoración propuestas para cada una de las preguntas y presentadas donde: 1: totalmente insatisfactorio, 2: poco satisfactorio, 3: neutral, 4: satisfactorio y 5: muy satisfactorio

Tal como se mencionó anteriormente, los niveles de prioridad que se consideraron para el test heurístico se describen a continuación.

A partir de los niveles de prioridad y como medio de verificación de los mismos, surgen los niveles de conformidad.

A: Todos los puntos de verificación de prioridad 1 se satisfacen.

AA: Todos los puntos de verificación de prioridad 1 y 2 se satisfacen.

AAA: Todos los puntos de verificación de prioridad 1,2 y 3 se satisfacen

Teniendo en cuenta la norma NTC 5854, partiendo de cada principio y de las pautas asociados a estos, se generaron un conjunto de criterios expresados en preguntas según de los cuales se obtuvo el test heurístico desarrollado. A modo de ejemplo del test generado en el presente trabajo se muestran a continuación algunas de las preguntas obtenidas a partir de las pautas pertenecientes a los principios de la norma NCT 5854 (Ver Tabla 1).

Principio	Nivel de Conformidad	Preguntas
PERCEPTIBLE	AAA	¿El portan web cuenta con niveles de ajuste que permitan modificar el tamaño del texto presente?
	AA	¿Los videos exhibidos dentro del portal presentan subtítulos indicando lo que sucede a lo largo del video?
OPERABLE	AAA	¿Se puede navegar desde el teclado por las diferentes pestañas o menús que se encuentren en el portal web?
	AAA	¿El portal web sigue con actividad si se presenta algún error (falla en la conexión a internet) sin que haya pérdida de la información con la que se está trabajando?
COMPRESIBLE	A	¿El portal web proporciona etiquetas o instrucciones en los campos a la hora de ingresar los datos?
	AA	¿El portal web mantiene la misma consistencia en cuento a los colores de los componentes a la navegación?
ROBUSTO	A	El contenido del portal web, es compatible con las herramientas actuales y futuras (por herramientas se entiende los navegadores Microsoft Internet Explorer, Google Chrome, Mozilla Firefox)
	A	¿El portal web emplea los estándares actuales como HTML5, CSS, JavaScript?

Tabla 1 – Preguntas que componen los principios

2.2. Caso de estudio

En el presente artículo se realizó un test heurístico basada en la norma NTC 5854, la cual plantea los requerimientos de accesibilidad que deben cumplir los portales web. El test generado fue aplicado en la ejecución del prototipo web Lauvid SAS desarrollada en el marco del proyecto "GESTIÓN COMERCIAL Y DE MERCADOS DIGITAL PARA MIPYMES AGROINDUSTRIALES, ESTUDIO DE CASO EMPRESA LAUVID SAS, PIENDAMO CAUCA.", con el fin de proporcionar un conjunto de recomendaciones de accesibilidad que permita mejorar la accesibilidad de dicho prototipo web.

LAUVID SAS tiene por objeto social la transformación y comercialización de hortalizas, frutas, verduras, aromáticas y condimentarías. Su origen radica en la necesidad de hacer empresa en el campo. Por esta razón su compromiso principal es con los campesinos, especialmente mujeres y jóvenes vulnerables ubicados en la zona de influencia, generando desarrollo y tejido.

3. Resultados Y Discucuion

La evaluación heurística, se realiza a través del criterio de los evaluadores expertos, cinco en total, a quienes luego de utilizar y navegar en el sitio web se les presenta el test heurístico el cual se plantea en criterios de cumplimiento expresados en pregunta agrupas en los principios establecidos en la norma: perceptibilidad, operatividad, comprensión y robustez, una vez los expertos realizan la evaluación del sitio web, esta es tabulada y analizada estadísticamente. Es importante resaltar que por tratarse de información aportada por expertos, el tratamiento estadístico a realizar en el presente artículo no es el clásico a través del cálculo de parámetros como la media aritmética y la desviación estándar, ya que este tipo de análisis es útil para muestras grandes y cuando se pretende realizar una caracterización de la población por medio de la muestra o la estimación de parámetros poblacionales, una vez aclarado lo anterior, y teniendo la puntuación dada por los expertos de 1 a 5, donde 1 se considera totalmente insatisfactorio y 5 muy satisfactorio, a cada una de las preguntas que componen cada uno de los criterios, se calcula el promedio para cada pregunta y se analiza el aporte dada por estas al cumplimiento del criterio, adicionalmente hemos establecido que cuando el promedio calculado de las respuestas de cada pregunta es inferior a cuatro , se ha definido una zona de no cumplimiento con este aspecto (pregunta) por lo cual se debe recomendar a los programadores responsables de la Página WEB, la corrección de este aspecto, y de esta manera dar cumplimiento a los criterios establecidos en la norma.

3.1. Análisis de los resultados en el principió perceptible

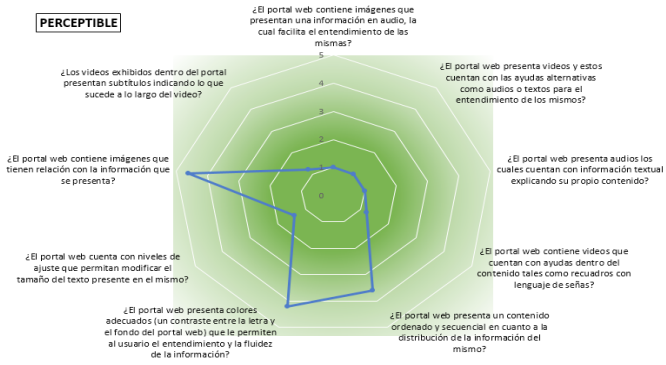


Imagen 1 – Diagrama radial, de valoración de expertos a cada una de las preguntas que componen el principio de Perceptible.

En esta sección se presenta los resultados obtenidos de la evaluación heurística realizada al prototipo web de LAUVID S.A.S para el principio perceptible. Tal como se puede

apreciar en el diagrama radial de la imagen 1, de acuerdo al juicio de los expertos, los aspectos relacionados con las preguntas P1, P2, P3, P4, P7 y P9 en promedio cada una presenta un nivel no satisfactorio, por lo cual deben ser corregidas en forma prioritaria, máximo cuando los aspectos P1, P2, y P3, están clasificados como nivel de conformidad A.

3.2. Análisis de los resultados obtenidos en el principió perceptible

En esta sección se presenta los resultados obtenidos de la evaluación heurística realizada al prototipo web de LAUVID S.A.S para operable.

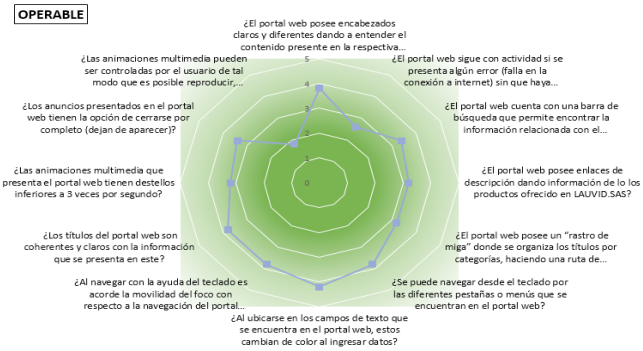


Imagen 2 – Diagrama radial, de valoración de expertos a cada una de las preguntas que componen el principio de Operable.

Tal como se aprecia en la figura 2, los aspectos que conforman el principio operable se encuentran en la gran mayoría muy cerca de la zona de conformidad, a excepción de los aspectos O2 y O12, es importante resaltar la prioridad que le debe otorgar al aspecto O12 ya que este aspecto es importante para lograr la conformidad A.

3.3. Análisis de los resultados obtenidos en el principió Comprensible

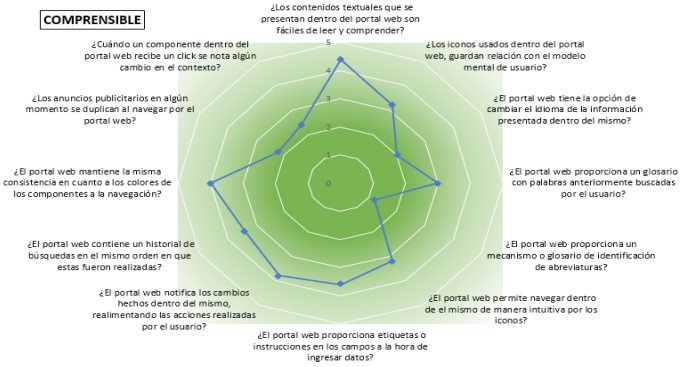


Imagen 3 – Diagrama radial, de valoración de expertos a cada una de las preguntas que componen el principio de Comprensible

En esta sección se presenta los resultados obtenidos de la evaluación heurística realizada al prototipo web LAUVID S.A.S para comprensible. En la imagen se observa que se deben corregir los aspectos C2, C3, C4, C5, C11 y C12, dando prioridad a los aspectos C3 y C11 por estar incluidos en el nivel de conformidad A.

3.4. Análisis de los resultados obtenidos en el principió Robusto

En esta sección se presenta los resultados obtenidos de la evaluación heurística realizada al prototipo web de LAUVID S.A.S para Robusto. Se puede observar en la imagen 4 que este criterio a juicio de los expertos, presenta un buen comportamiento, ya que no se presentan aspectos con valoraciones bajas, aunque es importante mejorar los aspectos R3 y R5, por encontrarse en zona de cuidado con los promedios más bajos de este aspecto.

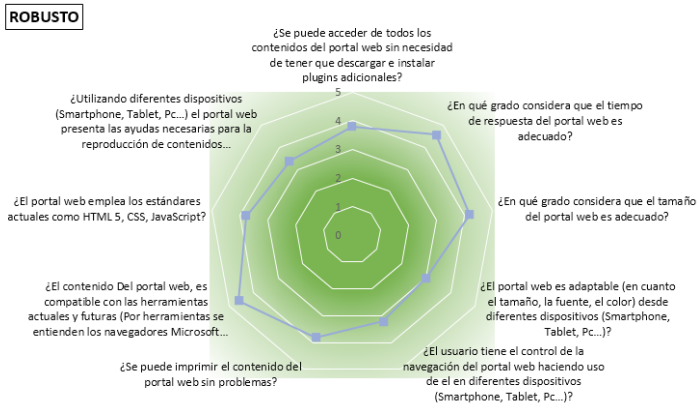


Imagen 4 – Diagrama radial, de valoración de expertos a cada una de las preguntas que componen el principio de Robusto

3.5. Criterio de conformidad

De acuerdo a los Criterios de Conformidad establecidos en la norma NTC 5854 y teniendo como criterio que la conformidad se da para aquellos criterios cuya valoración promedio sea igual a superior a cuatro, es decir que se clasifique como satisfactorio, se presenta el siguiente resultado para cada nivel de conformidad.

Criterio	A	AA	AAA
PERCEPTIBLE	2.160 No Conforme	1.825 No conforme	2.133 No Conforme
OPERABLE	3.280 No Conforme	3.280 No Conforme	3.350 No Conforme
COMPRENSIBLE	2.750 No Conforme	3.200 No Conforme	3.030 No Conforme
ROBUSTO	4.200 Conforme	4.300 Conforme	3.820 No Conforme

Tabla 2 – Valoracion de comformidad por nicel establecido en la norma NTC 5854

3.6. Coeficiente de concordancia

Se considera importante buscar un indicador que permita medir la concordancia de los juicios emitidos por los evaluadores, es por esto que revisando la bibliografía al respecto, se encuentra que uno de los índices utilizados para medir la concordancia entre los juicios de los evaluadores es el denominado índice de Kappa. La concordancia es un término más global que hace referencia a la medida en que dos o más codificadores están de acuerdo entre ellos, lo cual se entiende como la proporción de acuerdos entre el número total de codificadores (Juan Jose Torres Gordillo, Victor Hugo Perea Rodrigez, 1994). Inicialmente el estudio de concordancia fue realizado por Cohen en 1960. Este autor formuló el cálculo para la obtención de la probabilidad en la concordancia entre dos codificadores. Dicha probabilidad es conocida como coeficiente Kappa de Cohen (caracterizado con el símbolo k), el cual se define como un estudio de concordancia entre dos investigadores que escudriñan un sitio web en este caso. Como se dio a conocer anteriormente al ser una probabilidad, toma su valor en el intervalo [0, 1]. Ésta se representa en la ecuación 1.

Ecuación 1. Formula Kappa de Cohen

$$K = \frac{P_o - P_c}{1 - P_c}$$

Dónde:

Po: es la concordancia observada realmente y se calcula sumando las marcas que representan la concordancia y dividiendo por el número total de ellas.

Pc*:* es la proporción esperada por azar y se calcula sumando las probabilidades de acuerdo por azar para cada categoría.

Fleiss en 1981 generalizó la aplicación del índice Kappa de Cohen para medir el acuerdo entre más de dos codificadores u observadores para datos de escala nominal y ordinal. Por tanto, dado que el estudio considera cinco evaluadores en el proceso de codificación, se empleó Kappa de Fleiss, ya que este parte de la misma fórmula que propone Cohen, pero generalizada para más de dos codificadores. El coeficiente Kappa de Fleiss añade el cálculo del sesgo del codificador (precisión-error) y el cálculo de la concordancia (calibración) (Juan Jose Torres Gordillo, Victor Hugo Perea Rodrigez, 1994). El cual se puede calcular con la siguiente ecuación 2:

Ecuación 2. Formula Kappa de Fleiss

$$\overline{K} = \frac{1 - nm^2 - \sum_{i=1}^{n}.\sum_{j=1}^{r} x_{ij}^2}{nm\,(m-1)\sum_{j=1}^{r}.\overline{P}_j\overline{Q}_j}$$

Dónde:

n: Número de observaciones en la muestra estudiada (n= 5)

m: Identifica el número de evaluadores (m=210);

x_ij: Define el número de veces que se registros la categoría i en la heurística j;

r: Indica el número de heurísticas en que se compone el sistema nominal (r =42)

p: Es la proporción de acuerdos positivos entre evaluadores;

q: Es la proporción de acuerdos negativos (no acuerdos) entre evaluadores (1 - p).

Para el cálculo del Coeficiente de Kappa de Fleiss, se organiza la información en dos tablas, la tabla 1 es un arreglo que contiene las valoraciones que otorga cada evaluador a las heurísticas, y a la vez se calcula el número de acuerdos positivos, es decir cuando los evaluadores coinciden en la calificación de una o varias heurísticas. Se debe tener en cuenta que para cada heurística, los juicios de evaluadores organizados en todas las posibles parejas de evaluaciones arroja un total de 10 combinaciones a saber { (1,2), (1,3), (1,4), (1,5), (2,3), (2,4), (2,5), (3,4), (3,5), (4,5)}, con el número de acuerdos positivos (p) y acuerdos negativos (q), donde q se puede calcular como 10-p, donde se calcula la proporción de acuerdos positivos y la proporción de no acuerdos.

La proporción de acuerdos positivos ($\bar{P}$) se calcula dividiendo el número de acuerdos positivos sobre el total de posibles parejas ($\bar{P}$= p/10), de igual forma se calcula ($\bar{Q}$=q/10) y por último se obtiene el producto ($\bar{P}\bar{Q}$) para cada heurística y finalmente la sumatoria de estos productos. En la tabla 5 se muestra un arreglo donde se identifica el número de veces que la categoría i, es seleccionada en la heurística j :

Heurística	Ctg1	Ctg2	Ctg3	Ctg4	Ctg5	Xij^2
¿El portal web sigue con actividad si se presenta algún error (falla en la conexión a internet) sin que haya pérdida de la información con la que se está trabajando?	0	3	1	1	0	11

Tabla 3 – Números de veces que la categoría i, es seleccionada en la heurística j

De la heurística anterior se interpreta que, tres de los evaluadores marcaron la categoría 2, uno marco la categoría 3 y uno marco la categoría 4, mientras que las categorías 1 y 5 no fueron marcadas por ninguno de los evaluadores. De este modo X^2 i,j= $(0^2+3^2+1^2+1^2+0^2$ =11. Finalmente se suman todos los valores de la última columna cuyo resultado se denomina la sumatoria de cuadrados.

Con los datos arrojados en los dos arreglos anteriores se da aplicación a la ecuación Ec2 anteriormente enunciada para calcular el índice de Kappa de Fleiss, así:

Ecuación 3. Ecuación 2 remplazando valores obtenidos

$$1 - \frac{5(210^2) - 518}{5 * 210 * (210 - 1) * 7.04} = 1 - 0.136 = 0{,}864$$

Para la interpretación del Índice Kappa de Fleiss calculado, es necesario contar con criterios previos que permitan enjuiciar aquello que es objeto de evaluación. De esta manera Fleiss ofrece una clasificación de los Kappas con la cual ayuda a interpretar los

coeficientes obtenidos. Este autor caracteriza como regulares los Kappas que se haya entre 0.40 y 0.60, buenos de 0.61 a 0.75, y excelentes por encima de 0.75.

Por su parte, Altman en 1991 propone una clasificación algo más amplia. Los coeficientes registran valores que van desde 0 a 1, siendo 0 el valor donde hay mayor desacuerdo entre investigadores y 1 el punto donde encontramos mayor acuerdo. Su clasificación indica que los Kappas pueden ser pobres (0 a 0.20), débiles (0.21 a 0.40), moderados (0.41 a 0.60), buenos (0.61 a 0.80) y muy buenos (0.81 a 1.00)(Juan Jose Torres Gordillo, Victor Hugo Perea Rodrigez, 1994).

Bien sea que se utilice la interpretación de Fleiss o la de Altman , se puede concluir que la concordancia entre los juicios emitidos por los evaluadores frente a las heurísticas objeto de la evaluación es excelente bajo el criterios de Fleiss y muy buena bajo el criterio de Altman, esto demuestra que bajo la hipótesis de que los juicios emitidos fueron de manera independiente, los resultados indican cuan acordes son los juicios emitidos por los evaluadores sobre el mismo fenómeno, en este caso la accesibilidad de la página web analizada, es también importante resaltar que para este caso se da conocer que tan consistentes son el método y los instrumentos utilizados. De este modo, la validez de la evaluación heurística realizada es adecuada para el sitio LAUVID SAS y puede considerarse pertinentes las recomendaciones de accesibilidad obtenidas sobre el sitio.

3.7. Recomendaciones

Como resultado del test heurístico aplicada al portal web LAUVID SAS, se generaron un conjunto de recomendaciones clasificadas por principio, las cuales surgieron a partir de los errores de accesibilidad detectados en el portal

3.7.1. *Principio perceptible*

- El portal web cuenta con imágenes, pero estas deben ir acompañadas de un audio de apoyo, para facilitar su entendimiento.
- El portal web cuenta con audios y videos, pero estos deben ir acompañados de un texto de apoyo, para complementar la comprensión de su contenido.

3.7.2. *Principio operable*

- Ante una falla en la conexión a internet, no debe perderse la información obtenida anteriormente en el sitio.
- El portal web debe contar con una barra de búsqueda, la cual debe tener una ubicación de rápido y fácil acceso para los usuarios del portal.

3.7.3. *Principio comprensible*

- El portal web debe tener la opción de cambiar el idioma de la información presentada dentro del mismo.
- El portal web debe tener un glosario de términos técnicos y abreviaturas.

3.7.4. *Principio Robusto*

- Las imágenes presentadas dentro el portal deben ajustarse al tamaño del dispositivo en el cual se esté navegando.
- Los enlaces del portal web deberán funcionar correctamente, es decir no estar rotos o no funcionar.

4. Conclusiones Y Trabajos Futuros

El aporte principal del presente trabajo fue la generación de un test heurístico de accesibilidad puesto que no se evidenció la existencia de instrumentos concretos para la ejecución de la evaluación heurística de acuerdo a la norma NTC 5854. De este modo, el instrumento propuesto pretende servir de referencia para desarrolladores y diseñadores con el fin de posibilitar la construcción de sitios web accesibles y el proceso de evaluación de los mismos.

En este artículo se presentó como aporte adicional el análisis estadístico realizado a los resultados del test heurístico aplicado sobre el portal de Lauvid.com. En este análisis se hizo uso de diagramas radiales asociados a cada principio de accesibilidad de la norma. De este modo gracias a los diagramas realizados es más fácil comprender y entender donde se presentan más fallas dentro del portal web, a nivel de cada uno de los principios contemplados por la norma de accesibilidad NTC 5854

Durante esta investigación se tuvo la oportunidad de explorar y adaptar los diferentes estudios y criterios sobre las evaluaciones heurísticas, con el cual permite hacer una inspección y conocer si cada elemento de la interfaz de usuario sigue los principios establecidos de accesibilidad en la norma colombiana NTC 5854. Se realizó una serie de preguntas con cada principio para poder evaluar dicho portal web y así poder ver las fallas que presenta a la hora de navegar dentro del mismo.

A modo de trabajo futuro se pretende mejorar la funcionalidad del test heurístico generado, en el sentido que permita seguir avanzando en el conocimiento sobre accesibilidad con base a la norma Colombiana NTC 5854.

Referencias

Barbosa, A. (15 de 11 de 2011). COLNODO Y PROGRAMA GOBIERNO EN LINEA.

Digital, C. C. (11 de marzo de 2014). Colombia Digital. Obtenido de Colombia Digital: https://colombiadigital.net/actualidad/noticias/item/6686-derecho-a-la-informacion-como-derecho-fundamental.html

Icontec Internacional. (2011). Icontec Internacional.

Jakob Nilsen, Robert. L Mack. (1994). Usability Inspection Methods. New York, NY.

Juan Jose Torres Gordillo, Victor Hugo Perea Rodrigez. (1994). Calculo de la fiabilidad y concordancia entre codificadores de un siste de categorias para el estudio del FORO ONLINE en E-LEARNING. Revista de investigacion eduactiva, 94–101.

LAUVID. (19 de 04 de 2017). LAUVID. Obtenido de http://www.lauvid.com/productos/nuestra-empresa/

MinTIC. (07 de 05 de 2014). MinTIC. Obtenido de http://www.mintic.gov.co/portal/604/w3-article-6053.html

Organizacion mundial de la salud, Banco munduial. (2011). Discapacidad: Una preocupacion Mundial . En Informe mundial sobre la discapacidad (pág. 295). Japon : OMS, Prganizacion Mundial de la Salud.

SeoColombia. (09 de 03 de 2009). SEOCOLOMBIA. Obtenido de https://www.seocolombia.com/blog/que-es-la-accesibilidad-web/

Tecnica, N. (2012). NORMA TECNICA 5854. Obtenido de http://ntc5854.accesibilidadweb.co/index.php/15-explicacion-tecnica/descripcion-general/160-introduccion-a-la-norma-ntc-5854?layoutMode=default

Torres Burriel. (12 de 12 de 2008). torresburriel estudio. Obtenido de torresburriel estudio: http://www.torresburriel.com/weblog/2008/12/12/pautas-de-accesibilidad-para-el-contenido-en-la-web-20-wcag20/

UNET06. (Marzo de 2006). Usability Net. Obtenido de Usability Net: http://www.usabilitynet.org/tools/r_international.html

Wendy Chisholm, Gregg Vanderheiden, Ian Jacobs, w3c. (2000). Diseño Accseible de paginas web. Consejeria de Trabajo y Politica Social.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 23/10/2018
Aceitação/Acceptance: 03/12/2018

Granja inteligente: Definición de infraestructura basada en internet de las cosas, IpV6 y redes definidas por software

Edwin Barrientos-Avendaño[1], Dewar Rico-Bautista[2], Luis Anderson Coronel-Rojas[3], Fabian Ranulfo Cuesta-Quintero[4]

ebarrientosa@ufpso.edu.co, dwricob@ufpso.edu.co, lacoronelr@ufpso.edu.co, fcuestaq@ufpso.edu.co

[1] Ingeniería de sistemas. Maestría en ingeniería de sistemas y computación. Departamento Sistemas e Informática. Grupo de ingeniería en innovación, tecnología y emprendimiento (GRIITEM). Universidad Francisco de Paula Santander Ocaña. Ocaña, Colombia.

[2] Ingeniería de sistemas. Maestría en ciencias computacionales. Departamento Sistemas e Informática. Grupo (GRIITEM). Universidad Francisco de Paula Santander Ocaña. Ocaña, Colombia.

[3] Ingeniería de sistemas. Maestría en práctica pedagógica. Departamento Sistemas e Informática. Grupo (GRIITEM). Universidad Francisco de Paula Santander Ocaña. Ocaña, Colombia.

[4] Ingeniería de sistemas. Maestría en telecomunicaciones móviles. Departamento Sistemas e Informática. Grupo (GRIITEM). Universidad Francisco de Paula Santander Ocaña. Ocaña, Colombia.

Pages: 183–197

Resumen: En la universidad Francisco de Paula Santander Ocaña (UFPS Ocaña) existe una granja agrícola-pecuaria. Los estudiantes de los programas de zootecnia e ingeniería ambiental realizan prácticas en laboratorios de aves, cerdos, conejos, bovinos, entre otros. Las nuevas tecnologías se puedan integrar sin afectar el medio ambiente, generando diversos cambios en la forma de aprendizaje. Se proponen soluciones para el mejoramiento continuo de los procesos, que al final se hace en pro de la sostenibilidad, calidad de vida y una granja inteligente. Este artículo se contextualiza desde el concepto de internet de las cosas (IoT), su seguridad, la interoperabilidad de los dos protocolos IPV4 e IPV6 y redes definidas por software. Teniendo en cuenta las arquitecturas y tecnologías existentes, se propone una infraestructura de granja inteligente que permita dar soporte a todos los procesos.

Palabras-clave: IoT; IPv6; Arquitectura; Universidad inteligente; redes definidas por software.

Smart farm: Defining of infrastructure based on internet of things, IpV6 and software defined networks

Abstract: At the Francisco de Paula Santander Ocaña University (UFPS Ocaña) there is an agricultural and livestock farm. The students of the zootechnics and environmental engineering programs practice in laboratories of birds, pigs, rabbits, bovines, among others. The new technologies can be integrated without affecting

the environment, generating diverse changes in the way of learning. Solutions are proposed for the continuous improvement of processes, which in the end is done for sustainability, quality of life and an intelligent farm. This article is contextualized from the concept of Internet of Things (IoT), its security, the interoperability of the two protocols IPV4 and IPV6 and software defined networks. Taking into account existing architectures and technologies, an intelligent farm infrastructure is proposed to support all processes.

Keywords:IoT; IPv6; architecture; smart university;software defined networks.

1. Introducción

La Universidad Francisco de Paula Santander Ocaña (UFPS Ocaña) es el principal claustro universitario de la región del Catatumbo, Norte de Santander, Colombia. Se ofrecen múltiples programas académicos enfocados a desarrollar el sector productivo de la región. La universidad posee una granja agrícola, pecuaria y ambiental donde los estudiantes de zootecnia e ingeniería ambiental realizan prácticas en laboratorios de aves, cerdos, conejos, bovinos, entre otros. El programa de ingeniería de sistemas desarrolla proyectos en la granja. Impactan en el sector agropecuario y ambiental de la región. De igual maneraestán dando soporte a todos los procesos llevados a cabo en los laboratorios; basados en internet de las cosas (IoT) y procesos de virtualización y así acercarlos al concepto de inteligente (*smart*).

La tecnología se convierte en un elemento fundamental de la universidades, sin la cual no se podría llevar a cabo ninguna de sus funciones esenciales (Gómez, Jimenez, Gumbau, & Llorens, 2017). Para mejorar la gestión de sus recursos las universidades diseñan su direccionamiento estratégico en función de unas necesidades propias, pero no se tienen en cuenta los requerimientos de orden nacional que demandan su innovación y competitividad en áreas como la infraestructura tecnológica. Sus planes de desarrollo consideran la tecnología como soporte, pero no como una necesidad estratégica para su funcionamiento. Se delega su implementación para suplir necesidades específicas de software y hardware, pero no se evalúa la capacidad actual para aportar a los objetivos estratégicos de las universidades.

Al ser el pilar básico del desarrollo de las comunidades del futuro les permiten dotar de inteligencia a todos sus ámbitos y generar servicios que proporcionen una mejora calidad de vida a los ciudadanos y una mayor sostenibilidad a las ciudades y universidades (MACIÁ, 2017). Sin embargo, no convendría concebirla sólo como un elemento táctico, no deberían planificarse de manera aislada, sino que tendrían que formar parte de la planificación global de la universidad (Fernández Martínez & Llorens Largo, 2016). Se proyecta desarrollar infraestructuras bajo el protocoloIPV6(Rico-Bautista, Medina-Cárdenas, & Santos Jaimes, 2008; S. & R., 2007),que permitan implementar dispositivos IoT generando dentro de la universidad entornos inteligentes de forma segura logrando integrarlas a la red actual de la universidad, buscando la interoperabilidad de los dos protocolos IPV4 e IPV6(Lobo & Rico-Bautista, 2012; Rico-Bautista & Alvernia-Acevedo, 2017; Rico-Bautista, Medina-Cárdenas, & Rojas-Osorio, 2016), y dando soporte a todos los procesos de seguridad bajo el esquema de redes definidas por software.

El artículo se presenta en cuatro (4) partes que son: (i) Revisión de antecedentes, (ii) Conceptualización, (iii) Resultados obtenidos hasta el momento, así como de la

proyección de los trabajos a realizar, (iv) Definición de conclusiones acerca del objeto de estudio.

2. Antecedentes

Los sistemas pueden integrar estrechamente componentes físicos con virtuales (Jindarat & Wuttidittachotti, 2015), (Caria, Schudrowitz, Jukan, & Kemper, 2017). Esto representa una prioridad de investigación y desarrollo en el área de las TIC, concentrado en diferentes tecnologías, donde se resalta *Internet ofThings*(Chieochan, Saokaew, & Boonchieng, 2017). Diferentes trabajos (Dumitrache, Sacala, Moisescu, & Caramihai, 2017), (Taylor et al., 2013) proponen arquitecturas con un énfasis principal en el área agrícola (granja inteligente), abordando temas como la sostenibilidad y la adaptabilidad a los cambios ambientales y de mercado tratando de optimizar su rendimiento, manteniendo las sinergias entre sus componentes (Schader et al., 2016), (Odara, Khan, & Ustun, 2015). Esto se logra, por ejemplo, a través de sensores que se utilizan para supervisar en el entorno variables como: luminosidad, temperatura y humedad; al igual que con los actuadores tales como: bombas de agua, ventiladores y luces para el cultivo de plantas. Para ello se propone una arquitectura de sistema y soluciones relevantes, donde se integran con éxito diferentes módulos relacionado con sistemas de detección, comunicación y análisis de datos en un sistema completo, que no sólo monitorea el entorno de la granja sino también realiza la automatización remota y el control del usuario (Minh et al., 2017).

Ejemplos de tecnologías utilizadas en la solución de problemas en granjas se encuentran las siguientes: Redes inalámbricas de sensores (*WSN*) que están emergiendo rápidamente como una tecnología de red potencial para el uso en varias situaciones de emergencias (Bonevska, Jankova, Mateska, Atanasovski, & Gavrilovska, 2009). *PotatoScanner*, puede medir la temperatura superficial de las plantas y se controla a través de una aplicación de *Android*(Gernert& Wolf, 2017). Las tecnologías *ZigBee* permiten la identificación de plagas en los cultivos, la sequía o el aumento de humedad (Zhu, Lin, Xiaojing, & Yuanguai, 2012). En el área de horticultura se encuentran estudios de estimaciones y seguimiento al crecimiento de cultivos. La medición del volumen es importante para evaluar el valor económico de un árbol (Verma, Lamb, Reid, & Wilson, 2016)(Ryu et al., 2015). *Ifarma*, por ejemplo, es un sistema de información agrícola cuyo objetivo es planificar, supervisar y mantener un registro de todas las actividades agrícolas durante la temporada de cultivo (Paraforos et al., 2017). Para lograr un óptimo resultado asociado al crecimiento de los cultivos un factor influyente es el sistema de riego, siendo a su vez, uno de los problemas más importantes por el costo del agua y la energía para su implementación (Xie, Huang, Chi, & Zhu, 2017), (Jacob, 2017).

3. Conceptualización

3.1. Internet de las cosas base de un entorno *Smart*.

Internet de las cosas o IoT (Internet ofThings), se refiere a la conexión de objetos tecnológicos o que sean electrónicos a Internet, este concepto se deriva del avance de la tecnología y a la necesidad de compartir y controlar las cosas que nos rodean,

incluyendo las necesidades de árboles y plantas (Li, Xu, & Zhao, 2015), (Luvisi & Lorenzini, 2014). Algunos de los campos en los que IoT se presenta fuertemente son: la adopción generalizada de redes basadas en el protocolo IP (Airehrour, Gutierrez, & Ray, 2016), la economía en la capacidad de cómputo, la miniaturización, los avances en el análisis de datos (Danieletto, Bui, & Zorzi, 2013) y el surgimiento de la computación en la nube, salud, herramientas de aprendizaje, seguridad (Flauzac, Gonzalez, & Nolot, 2015), optimización de procesos (Aziz, 2016), agricultura (Stočes, Vaněk, Masner, & Pavlík, 2016), entre otros. IoT posee una arquitectura macro de tres niveles o capas (Hurlburt, Voas, & Miller, 2012). En el nivel básico despliega diferentes sensores, cada uno de ellos es una fuente de información, y diferentes tipos de sensores capturan diversos contenidos y formatos de información. El nivel intermedio integra varias redes cableadas e inalámbricas para transferir la información de las cosas con precisión. El nivel más superior es de la aplicación que consta de tres capas (capa de tecnología, capa de middleware, capa de aplicación) y se compone de aplicaciones que exportan toda la funcionalidad de los sistemas para el usuario final (Rodríguez Molano et al., 2015).

Internet de las cosas es un tema emergente de importancia técnica, social y económica (Whitmore, Agarwal, & Da Xu, 2015). Tener sensores u objetos dispersos que generan información desde cualquier lugar, requiere la interconexión de estos objetos heterogéneos a través de Internet (Jara, Moreno-Sanchez, Skarmeta, Varakliotis, & Kirstein, 2013). Cobran importancia los Objetos inteligentes (*smartobjects*), objetos físicos con un sistema embebido que le permite procesar información y comunicarse con otros dispositivos y realizar acciones con base en una acción o evento determinado (Ghaleb, Subramaniam, Zukarnain, & Muhammed, 2016). No obstante, todos estos sistemas complejos presentan retos tales como: 1) la interoperabilidad y los estándares debido a las diferencias entre software y hardware utilizado por cada uno en los diferentes procesos; 2) el análisis avanzado de enormes cantidades de datos producidos por la red de sensores, que requieren de un tratamiento que permita su posterior análisis (Danieletto et al., 2013); 3) la seguridad entendido como los ataques a dispositivos conectados a Internet, el temor a la vigilancia y las preocupaciones relacionadas con la privacidad (Jeon & Jeong, 2016) y 4) aspectos relacionadas con las economías emergentes (Bleda, Jara, Maestre, Santa, & Gómez Skarmeta, 2012; Byun et al., 2016).

Los esfuerzos de muchos gobiernos y organizaciones de mejorar la conectividad a Internet han generado una gran cantidad de dispositivos conectados a este, creciendo de forma exponencialmente. Estudios realizados mencionan que para el año 2020 la cantidad de dispositivos conectados en total será de 50 mil millones (Martinez, Mejia, & Munoz, 2016), mientras (Yu, Sekar, Seshan, Agarwal, & Xu, 2015) argumentan que, para el mismo año, habrá 25 mil millones de dispositivos sólo de IoT. Estas cifras ponen en manifiesto una preocupación a nivel de seguridad ya que muchos fabricantes de hardware y software de dispositivos tales como teléfonos móviles, sensores, cámaras de vigilancia, GPS, consolas de videojuegos, neveras, televisores, relojes; cada uno de ellos corriendo aplicaciones como redes sociales, georeferenciación, sistemas transaccionales, alertas, y demás; cada una de dichas aplicaciones genera datos en una amplia gama de formatos (texto, gráficos, imágenes, videos, sonidos, etc) los cuales pueden llegar a ser interceptados o robados por agentes maliciosos quienes explotan las vulnerabilidades valiéndose de las debilidades de fabricación tanto en hardware como en software y en la forma como estos dispositivos se comunican con ellos mismos y con el exterior (Katz. Matías David, 2013).

Un informe de la *Hewlett Packart Enterprise* sobre IoT revela cifras importantes a tener en cuenta ya que en la implementación de un sistema IoT muchas organizaciones pueden estar expuestas a pérdida de información o crear agujeros de seguridad por las características débiles que presentan los dispositivos donde el 70% no cifran sus comunicaciones, las cuentas de usuarios válidas pueden ser identificadas en un 70%, los dispositivos que por facilidad para los usuarios finales en su administración se les agrega interfaz gráfica, el 60% presentan vulnerabilidades a distintos ataques como secuencias de comandos en sitios cruzados (XSS) (Rico-Bautista, Parra-Valencia, & Guerrero, 2017), (Sánchez-Torres, Rodríguez-Rodríguez, Rico-Bautista, & Guerrero, 2018). Estudios realizados han permitido identificar que las vulnerabilidades presentes en los dispositivos IoT se focalizan en las técnicas de cifrado tanto en el almacenamiento como en la comunicación por falta de estándares de cifrado y descifrado, debido a que no existen algoritmos ligeros de cifrado que permitan implementarse en dispositivos con poca capacidad de procesamiento, generando fuga de información, pérdida de confidencialidad, comunicaciones no protegidas, rastreo de paquetes, entre otros (Martinez et al., 2016).

3.2. Modelo universidad inteligente: Eje *Smart environment*.

El concepto de *Smart university*(Rico-Bautista et al., 2017), al igual que el de *smartcity*, está fuertemente anclado al de la potencia de las TI, haciendo especial énfasis en el concepto del "internet de las cosas" (*Internet ofthings*) y de las actuales *smartnetworks*, las cuales interconectan sistemas y personas y, además, estimulan la innovación para facilitar un conjunto de objetivos en beneficio de todos.

Partiendo de lo expresado por Maestre (Góngora, 2016) que:

"Una ciudad inteligente es un territorio caracterizado por el uso intensivo de las tecnologías, principalmente de información y comunicación, para promover la colaboración, la innovación y la eficiencia para lograr el mejoramiento de la calidad de vida de los ciudadanos y la sostenibilidad de las ciudades a través del mejoramiento de los servicios hacia los ciudadanos" (p.6).

Ahora desde este concepto de *Smart city*, trasladado a toda su comunidad y a su entorno, Maciá, Berná, Sánchez, Lozano, & Fuster (MACIÁ, 2017) generan como resultado el concepto de *Smart university*, el cual mantiene el mismo objetivo principal de mejorar la calidad de vida de su comunidad aplicando de forma global, intensiva y sostenible las TI bajo el principio de servicio a la ciudadanía. Es por eso por lo que se trata de un entorno controlado, en el cual las políticas económicas, de sostenibilidad, construcción o desarrollo deciden de forma local y no dependen de factores externos, aunque siguen criterios similares a las políticas y estrategias nacionales o internacionales. Para lograr todo es necesario apoyarse en el poder de las TI que existen actualmente, ya que su correcto uso será el pilar fundamental para el éxito del proyecto. La infraestructura actual permite disponer de sistemas digitales conectados de manera global, capaces de soportar aplicaciones, servicios (Cárdenas & Bautista, 2012; Constanza & Cárdenas, 2009; Medina-Cárdenas, Areniz-Arévalo, & Rico-Bautista, 2016; Medina & Rico, 2008), plataformas y sensores, todos ellos funcionando en conjunto bajo un ecosistema o marco estandarizado de normas y protocolos capaces de actuar de manera integral y a su vez desacoplada, desempeñando un papel imprescindible en la generación correcta de conocimiento e interconexión.

Las principales ventajas de emplear TI son:

- Fácil acceso a los datos
- Adquisición de información
- Procesamiento rápido de la información
- Transmisión rápida
- Almacenamiento y análisis de la información
- Interactividad

Todos los paradigmas de las tecnologías emergentes proporcionarán una serie de características a cada una de las capas definidas en un modelo arquitectural genérico *Smart network* y proporcionando servicios a distintos niveles. Se puede dividir estos modelos en dos capas y a su vez, en una serie de niveles que proporcionan una división de responsabilidades. De los ejes básicos de una *Smart university*, anteriormente citados, nos centramos en *Smart*environment (Entorno y eficiencia). Este eje se centra en el uso de las tecnologías verdes *Green IT*, para desarrollar entornos inteligentes, donde sean capaces de optimizar los recursos naturales, preservar y proteger el medio ambiente.

4. Resultados y discusión

4.1. Infraestructura de comunicación granja

De acuerdo a Johnson et al. citado por (Banica, Burtescu, & Enescu, 2017), identifica siete categorías de tecnologías que impulsan la innovación en la educación: Tecnologías de consumo, estrategias digitales, tecnologías habilitadoras, tecnologías de Internet, tecnologías de aprendizaje, medios sociales y tecnologías de visualización.

- El acceso de los estudiantes a los materiales de aprendizaje desde cualquier ordenador u otro dispositivo conectado a Internet Xie et al. citado por (Banica et al., 2017);
- El uso de dispositivos inteligentes (equipados con aplicaciones de hardware y software) en las denominadas aulas inteligentes Xie et al. citado por (Banica et al., 2017);
- La recopilación y análisis de un importante volumen de datos procedentes de sensores y dispositivos portátiles con mayor facilidad y la supervisión de las capacidades y logros de los estudiantes Xie et al. citado por (Banica et al., 2017);
- El desarrollo de software social educativo dentro de un contexto de IoT, refiriéndose a los Servicios de Redes Sociales (SNS), Wikis, Weblogs como apoyo a actividades de colaboración.

Para proponer la infraestructura de red que sea la más adecuada para la organización de una *SmartFarm* basada en IoT para la universidad Francisco de Paula Santander Ocaña e ir generando espacios inteligentes bajo un entorno IPV6 (Carvajal Escobar, Santos Jaimes, & Rico Bautista, 2013; Lobo & Rico-Bautista, 2012), teniendo en cuenta la interoperabilidad con el protocolo IPv4 y los riesgos de seguridad a los que están expuestos dichos dispositivos, basados en las vulnerabilidades y amenazas presentes para dicha tecnología. En la Fig. 1, en el contexto de IoT se plantea una Intranet IPV6 (Carvajal Escobar et al., 2013; Jara et al., 2013; Lobo & Rico-Bautista, 2012; Rico-Bautista et al., 2016) para soportar todo lo relacionado con la granja inteligente.

Se puede observar la propuesta para la organización de la infraestructura de comunicación, donde se muestra una red corporativa actualmente soportando todos los procesos de negocio de la Universidad Francisco de Paula Santander Ocaña. Se ha planteado enrutar todo el tráfico a través de la tecnología SDN y un servidor de actualizaciones que esté monitoreando al exterior sobre posibles actualizaciones por parte de la casa fabricante de los dispositivos IoT, sin exponer directamente los dispositivos a internet con el objetivo de contrarrestar algunas vulnerabilidades presentes en el proceso de descarga de actualizaciones. En la intranet se plantean una red privada virtual VPN que permita sacar todo el tráfico IoT hacia la red corporativa después de haber realizado la translación a IPv4.

Implementación de listas de control de acceso sobre dispositivos virtualizados basados en Mininet el cual es un emulador de red que ejecuta una colección de dispositivos finales, switches, routers y enlaces en un solo core de Linux. Se utiliza la virtualización ligera para hacer que un solo sistema parezca una red completa. Los programas que se ejecutan pueden enviar paquetes a través de lo que parece ser una interfaz de red real, con una velocidad de enlace y con retardo. La relación entre Mininet y las SNDs se establece por su similitud al trabajar con nodos que físicamente no existen, pudiendo en ambos casos definir sus configuraciones por software. Una característica importante de Mininet es que permite emular redes definidas por software con todas sus características. En particular, trabaja con el protocolo Openflow, usado por las SDNs hacia y desde el entorno inalámbrico, lugar donde fluirá toda la información de los dispositivos inteligentes, una central que recepcione la información de las diferentes zonas agrícolas y pecuarias asegurando la información

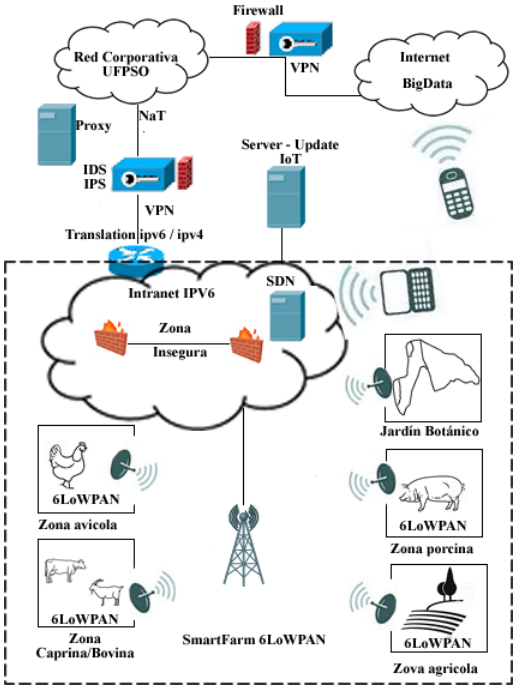


Figura 1 – Infraestructura de comunicación granja inteligente. Fuente: Autores

recolectada. Con la información suministrada de los sistemas IoT de la *SmartFarm* en la nube será sometida a un proceso de BigData(Sun, Song, Jara, & Bie, 2016) para la generación de nuevo conocimiento.

4.2. Arquitectura de seguridad soportada por redes definidas por software.

Las tres capas que constituyen la arquitectura de las redes definidas por software (SDN) interactúan a través de una interfaz de programación de aplicaciones APIs que permiten una mayor flexibilidad, fiabilidad y seguridad, elementos transcendentales en las nuevas redes emergentes. (Ver Fig. 2).

Capa de infraestructura. Es la capa más baja de SDN donde se encuentran los dispositivos físicos o virtuales como (switch, routers) los cuales están conectados a través de una interfaz abierta que permite el switcheo y envío de paquetes en una conexión de red. Estas conexiones se hacen a través de medios de transmisión diferentes, incluyendo, cobre, redes inalámbricas, y fibra óptica. Si bien las preocupaciones básicas están asociadas a esta capa ya que dependen de estos dispositivos para el correcto funcionamiento de las SDN.

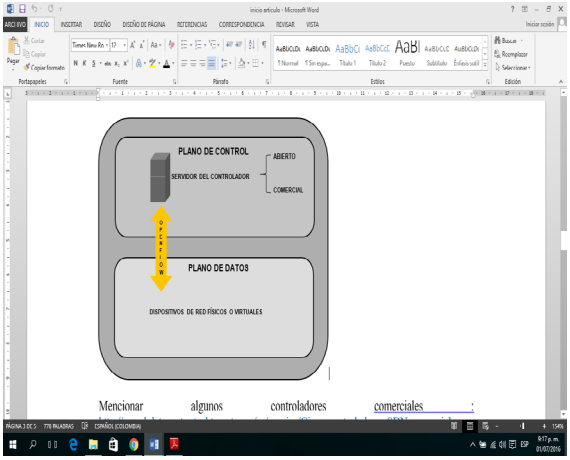


Figura 2 – Gráfica de la arquitectura de la SDN. Fuente: Autores

Capa de control. Es la capa central de la arquitectura de *SDN,* permite controlar un conjunto abstracto de recursos de plano de control, es decir, es la entidad que controla y configura los nodos de red para dirigir correctamente los flujos de tráfico. El controlador SDN elimina la inteligencia de conmutación y encaminamiento de datos de los nodos que realizan dicha función, pasando al controlador SDN, que toma esas decisiones y selecciona el mejor camino para el tráfico.

Capa de aplicación. Consiste en las aplicaciones de negocio de los usuarios finales, que utilizan servicios de comunicación de SDN a través de las API hacia arriba (northbound) de la capa de control, tales como REST, JSON, XML, etc., permite a los servicios y aplicaciones simplificar y automatizar las tareas de configuración, provisión y

gestionar nuevos servicios en la red, ofreciendo a los operadores nuevas vías de ingresos, diferenciación e innovación.

En el proceso de establecer una arquitectura de seguridad para una red corporativa ya sea una empresa de servicios, producción o una universidad, se hace necesario establecer una arquitectura que permita identificar los puntos críticos de la red cuando esta implementa dispositivos o sistemas IoT para controlar o gestionar ciertos procesos dentro de la empresa. Se tiene como objetivo establecer un esquema de *SmartFarm,* gestionando toda la seguridad y tráfico de la red a través de SDN, por lo tanto, en la Universidad Francisco de paula Santander Ocaña se hace necesario realizar un análisis sobre la forma como se conectan y se comunican los dispositivos IOT y la forma como compartirán la información recopilada hacia la nube. Todo esto enmarcado en controladores SDN buscando una mayor flexibilidad y la fiabilidad de la red.

Todos los componentes están siendo incorporados a través de diferentes proyectos de investigación, (ver tabla 1).

PROYECTO	DESCRIPCIÓN
Guía interactiva realidad aumentada la UFPS Ocaña.	Esta guía permitirá mostrar a toda la comunidad universitaria y visitantes información importante a tener en cuenta tanto en lo académico como en lo administrativo, como localización de salas de computo, aulas, laboratorios, oficinas administrativas, puntos de encuentro ante situaciones de emergencia, parqueaderos, zonas de estudio, restaurantes, granja y jardín botánico.
Sistema de información porcino para la UFPS Ocaña.	Este sistema contendrá toda la hoja de vida de machos, hembras y crías, con sus respectivas fases en el desarrollo: inseminación, manejo del semen, control de vacunas, seguimiento y control a la gestación.
Software interactivo de aprendizaje de bovinos en la UFPS Ocaña.	El software permitirá de una manera interactiva aprender sobre los bovinos en la Universidad Francisco de Paula Santander Ocaña
Aplicación móvil para el seguimiento y control de la alimentación y producción lechera de las cabras.	La aplicación CAPRINAPP es una herramienta que permitirá la toma de decisiones mediante una entrada de datos rápida y un sencillo pero eficaz método para informes. A través de la herramienta se podrá hacer seguimiento individual de las cabras (crecimiento, producción láctea, enfermedades…), permitirá generar reportes que ayudarán a la toma de decisiones.
Desarrollo de un sistema de geolocalización para bovinos: caso estudio granja UFPS Ocaña.	Este sistema permitirá hacer un seguimiento del desplazamiento de cada uno de los bovinos de la granja, el sistema permitirá monitorear cuando un bovino se encuentra en calor o está estresado a través de su recorrido, también permitirá ubicar en tiempo real la ubicación geográfica del bovino.
Sistema de aforo inteligente de pasto: caso de estudio granja UFPS Ocaña.	A través de este sistema de aforo inteligente, los potreros no sufrirán la degradación de las pasturas a causa de los manejos de pastoreo extensivo tradicional, el sistema permitirá determinar el potrero con la mayor cantidad de pasto acta para el consumo bovino y además reportará el crecimiento de pasto de cada potrero.

Tabla 1 – Listado de proyectos en curso

4.3. Prototipos de aplicación infraestructura granja inteligente

La granja inteligente es muy macroproyecto, que abarcará diferentes procesos, propios de una granja, con el ánimo de administrar de manera eficiente los recursos que se dan y que se llevan de manera manual. Unos de los prototipos que se están desarrollando es la aplicación CAPRINAPP, la cual, es una herramienta que permite la toma de decisiones mediante una entrada de datos rápida y un sencillo pero eficaz método para informes. A través de la herramienta se podrán hacer seguimiento individual de las cabras (crecimiento, producción láctea, enfermedades), permite generar reportes que ayuda a la toma de decisiones. El poder permitir realizar estos reportes, implica el poder determinar qué cantidad de comida le corresponde a cada una de las cabras, y por ende se podrá proporcionar la cantidad exacta de alimento para aquella que produzca más leche. Así mismo, el desarrollo de un sistema de geolocalización para bovinos: caso estudio granja UFPSO actualmente en la granja se realiza de manera manual el procedimiento de recoger las vacas para llevarlas hacia los corrales después de haber pastoreado, a veces se pierden algunas de ellas por ser tan grande la granja. A través de la herramienta se podrá hacer un seguimiento del desplazamiento de cada uno de los bovinos. un valor agregado es que la herramienta suministrara el recorrido de la vaca y a su vez se podrá determinar si esta estresada (mas recorrido de lo normal), lo que predice que la vaca esta en calor . (ver fig.3)

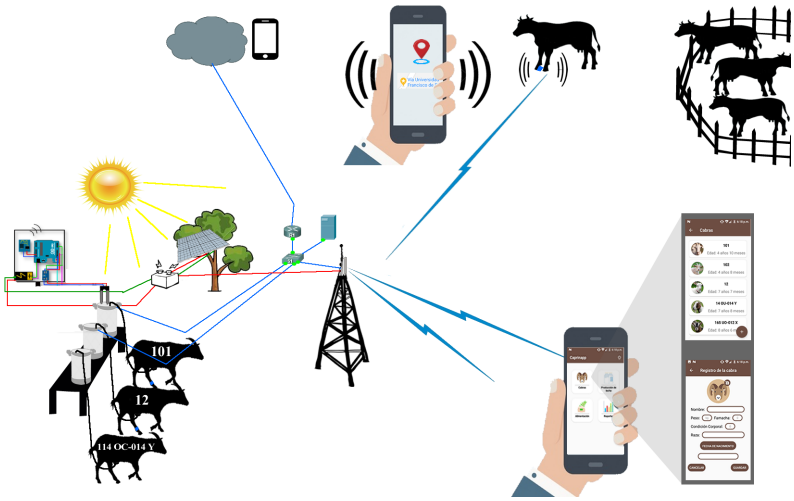


Figura 3 – Prototipos de aplicación granja inteligente.

5. Conclusiones

La forma como el internet de las cosas (IoT) ha venido creciendo a nivel mundial incursionando en todos los sectores de nuestra vida social, personal y laboral destacando la industria, medicina, educación, servicios en el hogar, vigilancia, transporte y agricultura y demás, nos permitió establecer una visión de cómo podemos implementar un sistema IoT que administre y gestione los distintos procesos de la granja agropecuaria con la que cuenta la Universidad Francisco de Paula Santander Ocaña.

Este artículo muestra el punto de partida para la construcción de una granja inteligente bajo un entorno IPV6 de forma segura minimizando el riesgo de comprometer activos de información de la organización.

Un valor agregado a la infraestructura granja inteligente son las redes definidas por software ya que permitirán llevar el control en lo referente a la seguridad a las capas superiores pues separa el hardware a través de la virtualización de los recursos.

Agradecimientos

A la UFPS Ocaña, mediante la División de Investigación y Extensión (DIE) vincula a docentes, administrativos y estudiantes para que participen en la ejecución y desarrollo de proyectos de investigación.

Referencias

Airehrour, D., Gutierrez, J., & Ray, S. K. (2016). Secure routing for internet of things: A survey. Journal of Network and Computer Applications, 66. https://doi.org/10.1016/j.jnca.2016.03.006

Aziz, B. (2016). A formal model and analysis of an IoT protocol. Ad Hoc Networks, 36, 49–57. https://doi.org/10.1016/J.ADHOC.2015.05.013

Banica, L., Burtescu, E., & Enescu, F. (2017). THE IMPACT OF INTERNET-OF-THINGS IN HIGHER EDUCATION. Scientific Bulletin-Economic Sciences, 16(1), 53–59.

Bleda, A. L., Jara, A. J., Maestre, R., Santa, G., & Gómez Skarmeta, A. F. (2012). Evaluation of the impact of furniture on communications performance for ubiquitous deployment of wireless sensor networks in smart homes. Sensors (Switzerland), 12(5), 6463–6496. https://doi.org/10.3390/s120506463

Bonevska, S., Jankova, A., Mateska, A., Atanasovski, V., & Gavrilovska, L. (2009). Early Fire Detection with WSN. Telecommunications Forum TELFOR, 1482–1485.

Byun, J., Kim, S., Sa, J., Kim, S., Shin, Y. T., & Kim, J. B. (2016). IoT(internet of things) based smart city services for the creative economy. International Journal of Smart Home, 10(7), 185–192. https://doi.org/10.14257/ijsh.2016.10.7.19

Cárdenas, Y. C. M., & Bautista, D. W. R. (2012). Mejores prácticas de gestión para la calidad de los servicios en tecnologías de información. Gerencia Tecnológica Informática.

Caria, M., Schudrowitz, J., Jukan, A., & Kemper, N. (2017). Smart farm computing systems for animal welfare monitoring. In 2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO) (pp. 152–157). https://doi.org/10.23919/MIPRO.2017.7973408

Carvajal Escobar, J. E., Santos Jaimes, L. M., & Rico Bautista, D. W. (2013). Movilidad en IPV6: simulación con Network Simulator. INGE CUC; Vol. 9, Núm. 2 (2013): (Julio - Diciembre). Retrieved from http://revistascientificas.cuc.edu.co/index.php/ingecuc/article/view/2

Chieochan, O., Saokaew, A., & Boonchieng, E. (2017). IOT for smart farm: A case study of the Lingzhi mushroom farm at Maejo University. In 2017 14th International Joint Conference on Computer Science and Software Engineering (JCSSE) (pp. 1–6). https://doi.org/10.1109/JCSSE.2017.8025904

Constanza, Y., & Cárdenas, M. (2009). Modelo de gestión basado en el ciclo de vida del servicio de la Biblioteca de Infraestructura de Tecnologías de Información (ITIL). Revista Virtual Universidad Católica Del Norte, (27), 1–21.

Danieletto, M., Bui, N., & Zorzi, M. (2013). RAZOR: A compression and classification solution for the internet of things. Sensors (Switzerland), 14(1), 68–94. https://doi.org/10.3390/s140100068

Dumitrache, I., Sacala, I. S., Moisescu, M. A., & Caramihai, S. I. (2017). A Conceptual Framework for Modeling and Design of Cyber-Physical Systems. STUDIES IN INFORMATICS AND CONTROL, 26(3), 325–334. https://doi.org/10.24846/v26i3y201708

Fernández Martínez, A., & Llorens Largo, F. (2016). Gobierno de las TI para universidades. Retrieved from http://tic.crue.org/wp-content/uploads/2016/04/gobierno_de_las_TI_para_universidades.pdf

Flauzac, O., Gonzalez, C., & Nolot, F. (2015). New security architecture for IoT network. In Procedia Computer Science (Vol. 52, pp. 1028–1033). Elsevier Masson SAS. https://doi.org/10.1016/j.procs.2015.05.099

Gernert, B., & Wolf, L. (2017). Poster: PotatoScanner - Using a field sprayer as mobile DTWSN node. CHANTS 2017 - Proceedings of the 12th Workshop on Challenged Networks, Co-Located with MobiCom 2017, 47–49. https://doi.org/10.1145/3124087.3124102

Ghaleb, S. M., Subramaniam, S., Zukarnain, Z. A., & Muhammed, A. (2016). Mobility management for IoT: a survey. Eurasip Journal on Wireless Communications and Networking. https://doi.org/10.1186/s13638-016-0659-4

Gómez, J., Jimenez, T., Gumbau, J., & Llorens, F. (2017). UNIVERSITIC 2017 Análisis de las TIC en las Universidades Españolas.

Góngora, G. P. M. (2016). Revisión de literatura sobre ciudades inteligentes: una perspectiva centrada en las TIC. Ingeniare, 19(19), 137–149. Retrieved from http://www.unilibrebaq.edu.co/ojsinvestigacion/index.php/ingeniare/article/viewFile/705/639%0Ahttp://www.unilibrebaq.edu.co/ojsinvestigacion/index.php/ingeniare/article/view/705

Hurlburt, G. F., Voas, J., & Miller, K. W. (2012). The Internet of Things: A Reality Check. IT Professional, 14(June), 56–59. https://doi.org/10.1109/MITP.2012.60

Jacob, N. K. (2017). IoT Powered Portable Aquaponics System, 1–5.

Jara, A. J., Moreno-Sanchez, P., Skarmeta, A. F., Varakliotis, S., & Kirstein, P. (2013). IPv6 addressing proxy: Mapping native addressing from legacy technologies and devices to the internet of things (IPv6). Sensors (Switzerland), 13(5), 6687–6712. https://doi.org/10.3390/s130506687

Jeon, J., & Jeong, S.-R. (2016). Designing a Crime-Prevention System by Converging Big Data and IoT. Journal of Korean Society for Internet Information, 17(3), 115–128. https://doi.org/10.7472/jksii.2016.17.3.115

Jindarat, S., & Wuttidittachotti, P. (2015). Smart farm monitoring using Raspberry Pi and Arduino. In 2015 International Conference on Computer, Communications, and Control Technology (I4CT) (pp. 284–288). https://doi.org/10.1109/I4CT.2015.7219582

Katz. Matías David. (2013). Redes y seguridad. Alfaomega Grupo Editor, (Mexico), 87.

Li, S., Xu, L. Da, & Zhao, S. (2015). The internet of things: a survey. Information Systems Frontiers, 17(2), 243–259. https://doi.org/10.1007/s10796-014-9492-7

Lobo, J., & Rico-Bautista, D. (2012). IMPLEMENTACIÓN DE LA SEGURIDAD DEL PROTOCOLO DE INTERNET VERSIÓN 6. Gerencia Tecnológica Informática, 11(29), 35–46. Retrieved from https://biblat.unam.mx/es/revista/gerencia-tecnologica-informatica/articulo/implementacion-de-la-seguridad-del-protocolo-de-internet-version-6

Luvisi, A., & Lorenzini, G. (2014). RFID-plants in the smart city: Applications and outlook for urban green management. Urban Forestry and Urban Greening, 13(4), 630–637. https://doi.org/10.1016/j.ufug.2014.07.003

MACIÁ, F. (2017). SMART UNIVERSITY. Hacia una universidad más abierta. (Alfaomega & Marcombo, Eds.) (Primera).

Martinez, J., Mejia, J., & Munoz, M. (2016). Security analysis of the Internet of Things: A systematic literature review. In 2016 International Conference on Software Process Improvement (CIMPS) (pp. 1–6). https://doi.org/10.1109/CIMPS.2016.7802809

Medina-Cárdenas, Y. C., Areniz-Arévalo, Y., & Rico-Bautista, D. (2016). Modelo estratégico para la gestión tecnológica en la organización: plan táctico de la calidad (ITIL & ISO 20000). (Instituto Tecnológico Metropolitano, Ed.). Fondo Editorial ITM. https://doi.org/https://doi.org/10.22430/9789585414006

Medina, Y. C., & Rico, D. W. (2008). Modelo De Gestión De Servicios Para La Universidad De Pamplona: Itil. Scientia Et Technica, XIV(39), 314–319. Retrieved from http://www.redalyc.org/articulo.oa?id=84920503056

Minh, Q. T., Phan, T. N., Takahashi, A., Thanh, T. T., Duy, S. N., Thanh, M. N., & Hong, C. N. (2017). A cost-effective smart farming system with knowledge base. ACM International Conference Proceeding Series, 2017–Decem, 309–316. https://doi.org/10.1145/3155133.3155151

Odara, S., Khan, Z., & Ustun, T. S. (2015). Integration of Precision Agriculture and SmartGrid technologies for sustainable development. In 2015 IEEE Technological Innovation in ICT for Agriculture and Rural Development (TIAR) (pp. 84–89). https://doi.org/10.1109/TIAR.2015.7358536

Paraforos, D. S., Vassiliadis, V., Kortenbruck, D., Stamkopoulos, K., Ziogas, V., Sapounas, A. A., & Griepentrog, H. W. (2017). Multi-level automation of farm management information systems. Computers and Electronics in Agriculture, 142(September), 504–514. https://doi.org/10.1016/j.compag.2017.11.022

Rico-Bautista, D., & Alvernia-Acevedo, S. A. (2017). ANÁLISIS DE UNA RED EN UN ENTORNO IPV6: UNA MIRADA DESDE LAS INTRUSIONES DE RED Y EL MODELO TCP/IP. Revista Colombiana Tecnologías de Avanzada, 1(29), 81–91. https://doi.org/https://doi.org/10.24054/16927257.v29.n29.2017.2490

Rico-Bautista, D., Medina-Cárdenas, Y. C., & Rojas-Osorio, J. A. (2016). PENTESTING EMPLEANDO TECNICAS DE ETHICAL HACKING EN REDES IPv6. Revista Ingenio UFPSO, 11(70–96), 1. Retrieved from http://revistas.ufpso.edu.co/index.php/ringenio/article/view/391/223

Rico-Bautista, D., Medina-Cárdenas, Y. C., & Santos Jaimes, L. M. (2008). Ipsec De Ipv6 En La Universidad De Pamplona. Scientia Et Technica, 2(39), 320–325. https://doi.org/http://dx.doi.org/10.22517/23447214.3239

Rico-Bautista, D., Parra-Valencia, J. A., & Guerrero, C. D. (2017). IOT: UNA APROXIMACION DESDE CIUDAD INTELIGENTE A UNIVERSIDAD INTELIGENTE. Revista Ingenio UFPSO, 13(1), 9–20.

Rodríguez Molano, J. I., Montenegro marín, C. E., Cueva Lovelle, J. M., Molano, J., Marin, C., & Cueva, J. (2015). Introducción al Internet de las Cosas. Redes de Ingeniería, 6(7), 53–59. https://doi.org/10.14483/udistrital.jour.redes.2016.1.a04

Ryu, M., Yun, J., Miao, T., Ahn, I. Y., Choi, S. C., & Kim, J. (2015). Design and implementation of a connected farm for smart farming system. In 2015 IEEE SENSORS (pp. 1–4). https://doi.org/10.1109/ICSENS.2015.7370624

S., L. M., & R., D. W. (2007). IPV6 en la Universidad de Pamplona: Estado del arte. Scientia Et Technica, XIII, 415–420. Retrieved from http://www.redalyc.org/articulo.oa?id=84903771

Sánchez-Torres, B., Rodríguez-Rodríguez, J. A., Rico-Bautista, D. W., & Guerrero, C. D. (2018). Smart Campus: Trends in cybersecurity and future development. Revista Facultad de Ingeniería, 27(47). https://doi.org/10.19053/01211129.v27.n47.2018.7808

Schader, C., Baumgart, L., Landert, J., Muller, A., Ssebunya, B., Blockeel, J., … Stolze, M. (2016). Using the Sustainability Monitoring and Assessment Routine (SMART) for the Systematic Analysis of Trade-Offs and Synergies between Sustainability Dimensions and Themes at Farm Level. SUSTAINABILITY, 8(3). https://doi.org/10.3390/su8030274

Stočes, M., Vaněk, J., Masner, J., & Pavlík, J. (2016). Internet of Things (IoT) in Agriculture - Selected Aspects. Agris On-Line Papers in Economics and Informatics, VIII(1), 83–88. https://doi.org/10.7160/aol.2016.080108

Sun, Y., Song, H., Jara, A. J., & Bie, R. (2016). Internet of Things and Big Data Analytics for Smart and Connected Communities. IEEE Access, 4, 766–773. https://doi.org/10.1109/ACCESS.2016.2529723

Taylor, K., Griffith, C., Lefort, L., Gaire, R., Compton, M., Wark, T., … Trotter, M. (2013). Farming the Web of Things. IEEE Intelligent Systems, 28(6), 12–19. https://doi.org/10.1109/MIS.2013.102

Verma, N. K., Lamb, D. W., Reid, N., & Wilson, B. (2016). Comparison of Canopy Volume Measurements of Scattered Eucalypt Farm Trees Derived from High Spatial Resolution Imagery and LiDAR. REMOTE SENSING, 8(5). https://doi.org/10.3390/rs8050388

Whitmore, A., Agarwal, A., & Da Xu, L. (2015). The Internet of Things—A survey of topics and trends. Information Systems Frontiers, 17(2), 261–274. https://doi.org/10.1007/s10796-014-9489-2

Xie, T., Huang, Z., Chi, Z., & Zhu, T. (2017). Minimizing amortized cost of the on-demand irrigation system in smart farms. CySWATER@CPSWeek, 43–46. https://doi.org/10.1145/3055366.3055370

Yu, T., Sekar, V., Seshan, S., Agarwal, Y., & Xu, C. (2015). Handling a trillion (unfixable) flaws on a billion devices. In Proceedings of the 14th ACM Workshop on Hot Topics in Networks - HotNets-XIV (pp. 1–7). https://doi.org/10.1145/2834050.2834095

Zhu, X., Lin, Y., Xiaojing, Z., & Yuanguai, L. (2012). ZigBee implementation in intelligent agriculture based on Internet of Things. Proceedings of the 2nd International Conference on Electronic and Mechanical Engineering and Information Technology, EMEIT 2012, 23, 1842–1846. https://doi.org/10.2991/emeit.2012.408

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 17/09/2018
Aceitação/Acceptance: 07/11/2018

Modelo de arquitectura Cloud Federada utilizando recursos oportunistas de universidades

Diego Rincon[1], Cesar Diaz[2], Olmer Garcia[2], Holman Bolivar[1]

darincon@ucatolica.edu.co, cesaro.diazb@utadeo.edu.co, olmer.garciab@utadeo.edu.co, hdbolivar@ucatolica.eu.co

[1] Universidad Católica de Colombia, Bogotá, Colombia

[2] Universidad de Bogotá Jorge Tadeo Lozano, Bogotá, Colombia

Pages: 198–211

Resumen: Este proyecto tiene como finalidad el diseño de una arquitectura, que pueda proveer una plataforma tecnológica capaz de prestar un servicio interinstitucional o federado (entre organizaciones), de uso de recursos computacionales de tipo oportunista. Para cumplir el objetivo general, se evaluaron tecnologías e implementaciones existentes, con el fin generar un modelo de arquitectura a través de un método iterativo e incremental, para después realizar una prueba de concepto por medio de un prototipo, que permitió la validación del modelo propuesto.

Palabras-clave: Computación en la nube, Computación oportunista, Nube Federada.

Federated Cloud architecture model using opportunistic university resources

Abstract: The main objective for this project is to design an architecture that's able to offer a technological platform capable of providing an inter-institutional or federate service (between organizations) using a cloud-computing paradigm to deliver opportunistic resources. To accomplish the main purpose of this project, two technologies and types of implementation available were studied and evaluated, in order to generate an architectural model through an iterative and incremental method, consequently a concept test were done employing a prototype that allows the validation of the suggested model.

Keywords*:* Cloud computing, Opportunistic Cloud, Federated Cloud.

1. Introducción

El Cloud marca una tendencia tan grande que en el mercado existen proveedores que prestan servicios de infraestructura, de forma masiva(Foster, Zhao, Raicu, & Lu, 2008), sin embargo cuando dos o más organizaciones que tienen recursos computacionales se unen, dan paso a lo que se conoce como un Cloud Federada, donde cada miembro

se acoge a las regulaciones de la federación. Lo que también conlleva en nuevos retos computacionales descritos en (Goiri, Guitart, & Torres, 2010).

En este sentido un modelo de arquitectura de Cloud federado, desde el punto de vista de miembro se convierte en un esquema organizacional y tecnológico, robusto que permite a cualquier miembro aumentar su capacidad de respuesta y análisis de datos en un momento específico, ya sea para la resolución de problemas de análisis o el despliegue de un servicio de primer nivel (Gonzalez Guerrero, Castro, & Rincon Yañez, 2012).

El modelo de cómputo oportunista, consiste en la donación del tiempo que las máquinas se encuentran sin uso para fortalecer una infraestructura de cómputo existente (Celesti, Fazio, Villari, & Puliafito, 2012), sin embargo estos recursos necesitan de una capa de software adicional que tenga la habilidad de iterar entre la funcionalidad de escritorio y la publicación de recursos en el pool al cual está asignado (Kertesz, Kecskemeti, & Brandic, 2014). Si combinamos la arquitectura Cloud federada junto con un modelo de alistamiento de recursos de cómputo oportunista (Rochwerger et al., 2009) podemos, como primera medida aumentar el poder computacional de cada miembro de la federación.

Dentro de las instituciones educativas, específicamente en las instituciones de educación superior (IES), existen recursos computacionales que no están siendo utilizados las 24 horas del día, sin embargo, muchos de ellos permanecen encendidos lo que acarrea a las instituciones costos de funcionamiento y mantenimiento de los mismos(Rincon Yanez, 2016).

El modelo desarrollado dentro de este proyecto de investigación busca la creación de un mecanismo en el que las instituciones que forman parte de una federación no solo apoyen con recursos de cómputo de una manera fácil entre sí, sino que también los recursos de las instituciones se usen de una forma más óptima. Adicionalmente el modelo busca crear un uso consiente de los recursos computacionales al aprovechar los tiempos muertos de los recursos que se encuentran ociosos, reduciendo la huella de carbono al no necesitar la adición de más recursos a la infraestructura existente.

El propósito del desarrollo de este modelo de arquitectura Cloud federada utilizando metodologías oportunistas para el alistamiento de los recursos computacionales es proveer un servicio de infraestructura tipo IaaS (Infraestructure as a Service) a las entidades de educación superior de forma libre, segura y escalable. Esto permitirá a las instituciones de educación superior pertenecientes a la federación aumentar la escala de resolución de problemas al obtener más recursos de cómputo, así como un ahorro en los costos de operación ya que el aprovechamiento de la infraestructura existente y ociosa se verá explotado generando un uso útil para la misma institución o algún miembro de la federación.

El presente artículo, primero, describe las características de la computación oportunista y las herramientas que se usan para el proyecto. Seguido el trabajo muestra el proceso de desarrollo del modelo que permite describir los actores que interactúan en la arquitectura. Tercero es presentada la arquitectura de software y datos propuesta, por último, se presentan los resultados de una prueba de conceptos y las conclusiones.

2. Computación oportunista

Cuando se utiliza los servicios de computó en las salas de computo de las universidades, los estudiantes en promedio utilizan solo un porcentaje del procesador para sus tareas diarias; la computación oportunista trata de utilizar el restante de procesador, sin quitar rendimiento al usuario principal que utiliza el recurso en ese momento (C. O. Diaz, Castro, Villamizar, Pecero, & Bouvry, 2013).

Para que este tipo de computación funcione, es necesario un sistema administrador de recursos locales, que controla la carga de procesamiento previniendo que los trabajos compitan por recursos de cómputo limitado. Típicamente, un sistema manejador de recursos comprende un manejador de recursos, un programador y *scheduler*.

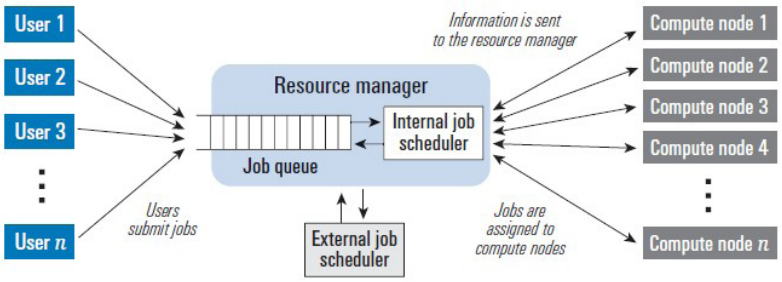


Figura 1 – Esquema de Manejador de Recursos de LSF(Iqbal, Gupta, & Fang, 2005).

Usualmente, el manejador de recursos (Figura 1) controla varios procesos en los nodos maestros y en los nodos de trabajo, incluyendo un proceso programador, que típicamente se ejecuta en los nodos maestros. El manejador de recursos también fija un sistema de colas para los usuarios que envían trabajos, los usuarios pueden consultar el manejador de recursos para determinar el estado de sus trabajos.

A continuación, son presentados los conceptos de las dos tecnologías que se estudiaron que pueden realizar las tareas anteriormente descritas (Baldión Albarracín & Celis Baracaldo, 2015) presenta un análisis comparativo de UnaCloud Y Boinc en un ambiente oportunista donde se realizó un análisis DOFA (debilidades, oportunidades, fortalezas y amenazas) de las herramientas.

2.1. BOINC

El *berkeley open infrastructure for network computing (BOINC)* es un sistema de middleware abierto que es soportado para la computación voluntaria y oportunista. BOINC tiene como objetivo permitir a los investigadores de aprovechar la enorme potencia de procesamiento de múltiples ordenadores personales de todo el mundo.

La Figura 2 presenta la arquitectura del sistema, que está compuesto por programas que se ejecutan en los voluntarios coordinados por un servidor por proyecto.

El cliente BOINC tiene implementado dos políticas de planificación, una hace referencia a la planificación de los recursos de hardware como CPU y memoria, y la segunda al trabajo a buscar, es decir la cantidad de trabajo que es solicitado por un proyecto.

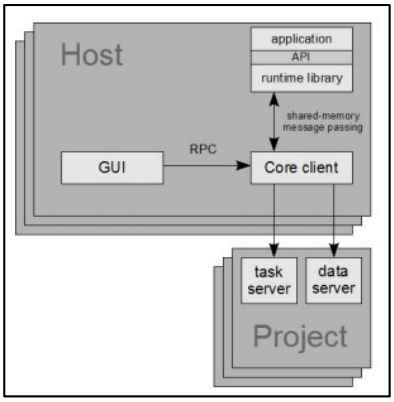


Figura 2 – Arquitectura BOINC (Besso & Striglio Raúl, 2011)

Estas políticas tienen impacto en el rendimiento de los proyectos que son basados en BOINC y de los cuales se evidencia que:

- Las estimaciones proporcionadas para la finalización de un trabajo proporcionado por un proyecto, pueden ser en orden de magnitud.
- Los usuarios pueden cambiar la planificación de los recursos que comparte.
- Los proyectos también pueden estar fuera de línea o contar con periodos muertos donde no se utilizan los recursos.
- Las unidades de trabajo componen una tarea, cada unidad de trabajo es independiente y es enviada desde el servidor hasta el cliente para su procesamiento.

Las unidades de trabajo que están compuestas de códigos ejecutables, parámetros y archivos de entrada las cuales son enviadas al servidor una vez que se hayan procesado.

2.2. UnaCloud

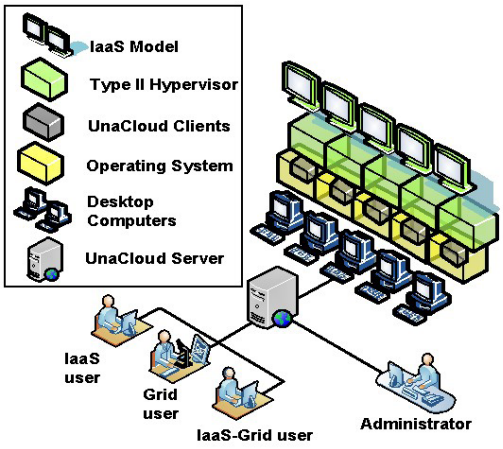


Figura 3 – Arquitectura UnaCloud(Rosales, Castro, & Villamizar, 2011).

Es una infraestructura tipo IaaS que en ambientes oportunistas permite el aprovechamiento de los ciclos de ocio de los computadores para la contribución a proyectos que requieran un mayor nivel de procesamiento.

La arquitectura de UnaCloud (**Figura 3**) incluye un sistema de información Web y una infraestructura computacional basada en computadores de escritorio, cuya integración es capaz de articular una infraestructura computacional oportunista, que provee bajo demanda servicios computacionales fundamentales, sistemas operativos y aplicaciones(Rosales et al., 2011).

3. Desarrollo del modelo

Dentro del desarrollo del proyecto se han integrado los conceptos computación en la nube, computación oportunista y federación, los cuales apuntan a una macro integración de servicios que traspasan las barreras de una institución.

La federación actuará como controlador de recursos y entidad que gestiona la "confianza" entre todos los participantes de esta, incluyendo la autenticación inter-dominios de los usuarios pertenecientes a todas las instituciones pertenecientes al conglomerado. Dentro de la federación se han identificado tres (3) actores tal y como se describen en Figura 4, los cuales tienen funciones específicas, estas funciones están basadas en la combinación del modelo cliente/servidor para que el usuario consuma el servicio y publicador/suscriptor y para que la institución participante publique su servicio ante el control de la federación (Coulouris, Dollimore, Kindberg, & Blair, 2012).

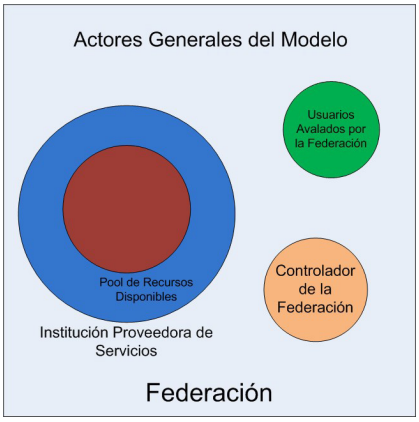


Figura 4 – Actores Generales del modelo.

Los primeros actores son los proveedores de servicio de cómputo. La institución tiene la tarea de donar su capacidad de cómputo inactiva a la federación. Para esto la institución proveedora debe implementar un Oportunistic Local Resource Manager (OLRM) y un servicio adicional conocido como Service Provider (SP). El SP es un publicador de servicio hacia la federación y a su vez es un intermediario con el proveedor de identidad(Seltsikas & Currie, n.d.), sin este componente el servicio no será avalado por el controlador de

la federación, de esta manera se crea una relación de confianza entre el Proveedor de Identidades y el controlador.

Los segundos son los usuarios del servicio, los cuales pertenecen a una institución de la federación, los cuales están regidos por políticas de la institución a la que pertenecen que incluyen horarios de uso, cuotas de uso y servicios disponibles. La institución auspiciante debe tener un directorio que utilice un esquema de información inetOrgPerson (Smith, 2000), este esquema debe estar en un servicio responda al protocolo LDAP.Un usuario necesitará un complemento, este complemento está definido por eduGain el documento GN4-1(Hämmerle et al., 2016).

El tercer actor es el controlador de la federación que es un actor el cual implementa un esquema de Platform as a Service (PaaS), esta plataforma tiene como función la gestión y el control de los participantes en la federación, tanto como los proveedores de servicio, como los proveedores de identidad. Una de las funciones del controlador de la federación es actuar como descubridor de servicios hacia los usuarios y de entidad registrante hacia los proveedores de servicio.

Este modelo se utiliza como base en la conformación de una iniciativa de Cloud Federada para Colombia compuesta inicialmente por instituciones universitarias en distintas ciudades que conformaran la infraestructura inicial compartiendo sus salas de computo (Cesar O. Diaz, Gomez, Castro, Barrios, & Bolivar, 2016) se puede encontrar.

4. Arquitectura general del modelo

La arquitectura se encuentra dividida en un modelo multi-capas, estas capas están separadas dadas las responsabilidades de los actores, garantizando un esquema descentralizado. La Figura 5 presenta la arquitectura propuesta.

4.1. Componentes de la arquitectura

Dentro de los componentes de la arquitectura propuesta se pueden observar que dos de los actores tienen infraestructura o componentes tecnológicos, necesarios para el funcionamiento del modelo propuesto. Los componentes por actor son:

Institución Prestadora de Servicio: Pool de Recursos son recursos computacionales pertenecientes a la institución que realizan su operación de forma oportunista, **Directorio Institucional** es el listado de políticas, Usuarios y Grupos, pertenecientes a la institución auspiciante, bajo un esquema de publicación con los requisitos técnicos de metadatos. **Proveedor de Identidades** es el componente tecnológico perteneciente a la federación en custodia de la Institución auspiciante, que tiene como fin mediar la relación de confianza entre los usuarios institucionales y la Federación en General y por ultimo el **Proveedor de Servicios** que es el componente tecnológico perteneciente a la federación en custodia de la Institución Prestadora del Servicio, que tiene como propósito la publicación del servicio para los usuarios autenticados.

Manejador de la Federación: Sistema de Información es la interfaz visual que el Usuario del Servicio va a utilizar, adicionalmente es el puente directo entre los componentes tecnológicos pertenecientes a la federación, en custodia de la Institución y los servicios de control del Manejador de la Federación, **Descubrimiento de Servicios** es el Servicio en el

cual los proveedores de servicio y de identidad se registran para poder establecer relaciones de confianza entre los directorios institucionales y los servicios de otras instituciones, componente de **Políticas y Acuerdos** es un servicio en el que se encuentran almacenadas las políticas, acuerdos institucionales y de uso de los servicios registrados en la federación, **Monitoreo de Políticas de Participante** es el servicio en el que se crean, guardan y almacenan los accesos, registros y tareas que el Usuario del Servicio utiliza en la federación y para finalizar los **Registros y Logs** que se usa para permitir el registro de eventos en la federación, con el propósito de realizar auditorías y controles de seguridad tanto a los usuarios como a las Instituciones prestadoras de Servicios y Auspiciantes.

4.2. Procesos Principales

Los procesos son acciones que los componentes del modelo están en capacidad de realizar con el fin de prestar algún servicio al usuario final del mismo de la federación. Estos procesos están ligados directamente al paso de información desde el fondo de la arquitectura hasta el sistema de información que es el punto de contacto con el Usuario del Servicio.

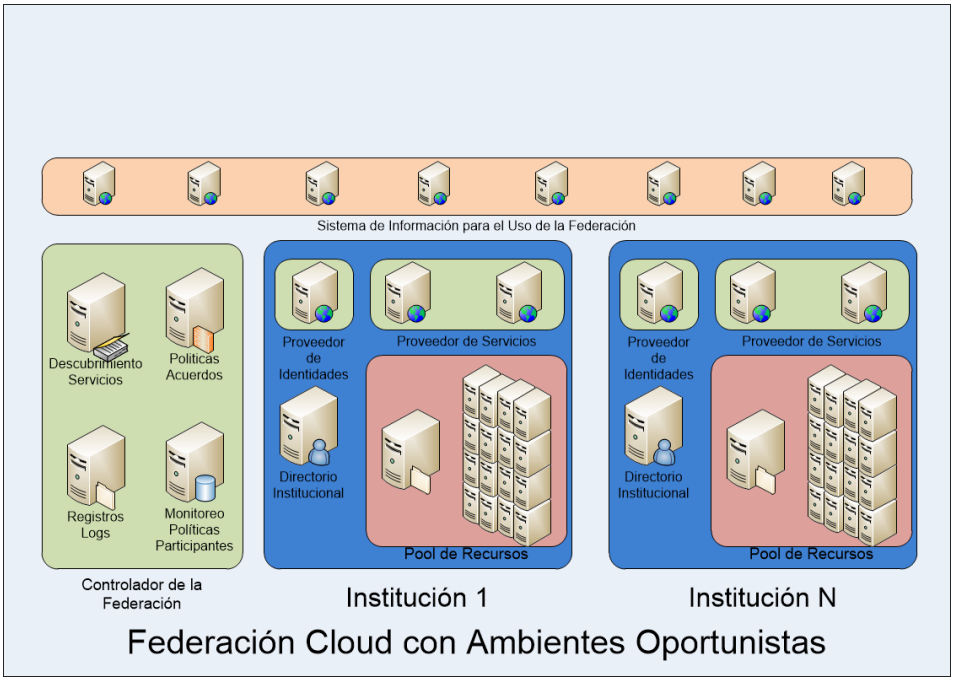


Figura 5 – Arquitectura General del Modelo

Dentro de la arquitectura descrita existen innumerable cantidad de procesos, estos se pueden clasificar en los componentes donde están ubicados y con que otros componentes se comunican, a continuación hay una lista de aquellos más representativos, según su componente de origen:

- Adaptador de Proveedor de Servicios: Consulta de estatus de cola, Consulta de estado de tareas, Envío de tareas, Consulta de resultado de Tareas.
- Repositorio de Políticas y Acuerdos Institucionales: Asignación de Políticas y acuerdos por proveedor de servicio, Consulta de Políticas por Institución.
- Sistema de Registro de Políticas de Participante: Creación de Perfil de usuario, Actualización de información de Tareas e interfaces al proveedor de servicio, Actualización de Controles y Dashboard de Usuario.
- Sistema de Registro y Control de Eventos: Captura de eventos de Usuarios, y de Proveedores de Servicio, Captura de eventos de Proveedores de Identidad, Captura de eventos del Servicio de Descubrimiento.
- Sistema de Información del Usuario: Consulta de Perfiles de usuario y de interfaces de Consumo, Consulta de Estado de Dashboard, Despliegue de información de usuario y de Políticas Institucionales y de Usuario.

Sin embargo se identifican 3 procesos los cuales son los más complejos, y ya que incluyen la mayor cantidad de componentes en su flujo de trabajo se han catalogado como Principales, estos procesos son:

4.2.1. Inicio de Sesión de Usuario

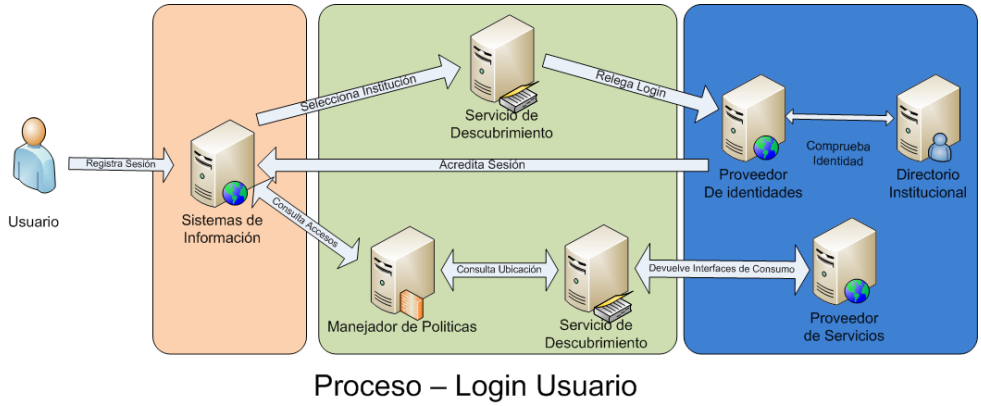


Figura 6 – Proceso de inicio de sesión del usuario

El Proceso de Inicio de sesión del usuario, descrito en la **Figura 6**, tiene como finalidad, validar las credenciales de un usuario de la federación, este puede pertenecer a cualquier institución validada por la misma. El Sistema de Información de la Federación le proveerá al Usuario del Servicio una interfaz web con el fin de poder contactar a su institución auspiciante.

4.2.2. Creación del Perfil de Usuario

El Proceso de la creación de un perfil de usuario se da una sola vez, al momento en el que el usuario inicia sesión por primera vez en la federación. El proceso se describe en la **Figura 7**.

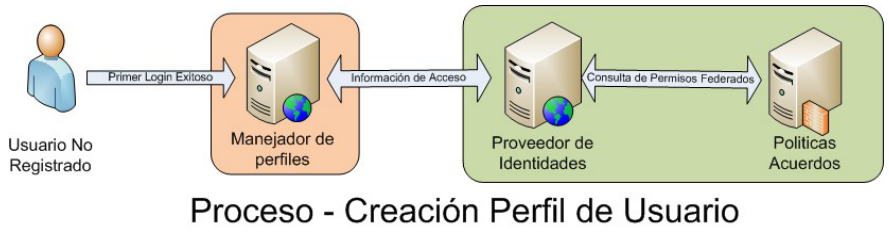


Figura 7 – Proceso de creación de perfil del usuario

4.2.3. Consulta y Despliegue de Información

El Proceso de Consulta y Despliegue de Información, descrito en la **Figura 8**, tiene como propósito el despliegue de información de un Dashboard de los recursos y tareas pertinentes del Usuario del Servicio.

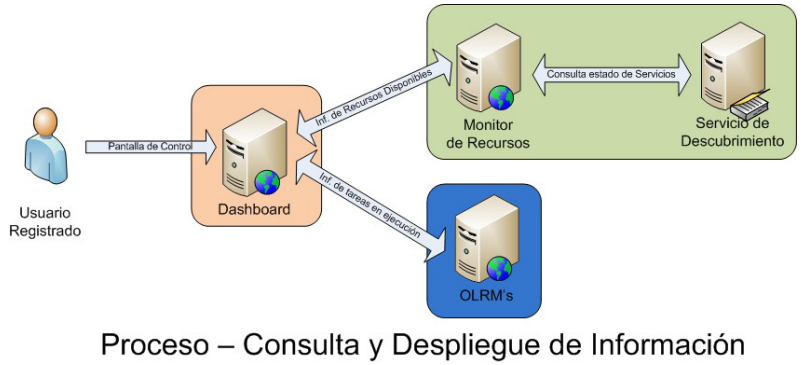


Figura 8 – Proceso de Consulta y despliegue de Información.

4.3. Componentes de la Federación

Los componentes generan los servicios pertenecientes al modelo, estos servicios no son todos propuestos por el modelo, algunos son sistemas existentes y requeridos por el mismo para su correcto funcionamiento. Los servicios no definidos por esta arquitectura son los siguientes: Pool de Recursos Oportunistas, Directorio Institucional, Proveedor de Identidades, Proveedor de Servicios, Sistema de Descubrimiento de Servicios

Estos componentes se encuentran dispersos en la arquitectura multi-capa propuesta y tienen como funcionalidad brindar las funcionalidades de control, acceso, seguridad y monitoreo de recursos.

4.3.1. Sistema de información

Como se había dicho antes, es la interfaz visual que el Usuario del Servicio va a utilizar. El esquema de servicio de este componente hacia el usuario federado es de tipo Software as a Service (SaaS). La arquitectura básica del componente es presentada en la **Figura 9**.

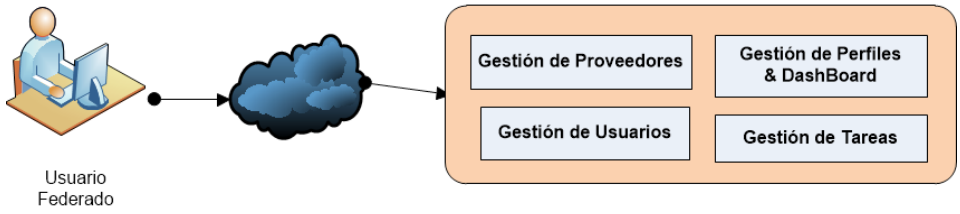


Figura 9 – Arquitectura Componente - Sistema de Información.

4.3.2. Políticas y Acuerdos

El componente de políticas y acuerdos funciona como un repositorio de información especializada de toda la federación. Este repositorio de información está disponible para la consulta de los participantes (componentes y usuarios), así como la gestión de las políticas de los usuarios responsables de las instituciones y de los servicios. En la Figura 10 es presentado el modelo de datos de este componente.

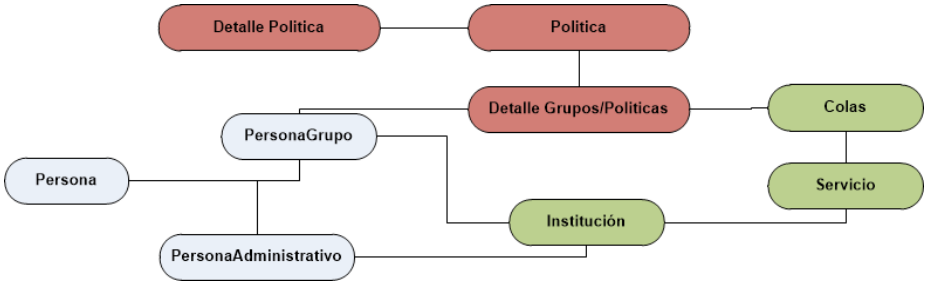


Figura 10 – Modelo de Datos Componente - Políticas y Acuerdos.

Entre esta información está comprendida la información general y políticas de uso de las instituciones proveedoras, así como las políticas de cada auspiciante de usuarios separadas en grupos de usuarios. El modelo de prestación de servicio de este componente está orientado a una arquitectura Platform as a Service (PaaS), funcionando como una plataforma de gestión para los usuarios responsables de los prestadores de servicio y una plataforma de consumo de información estándar para los componentes de la federación.

4.3.3. Registros y Logs

El componente de Registro y Log es el encargado de la gestión de los eventos del Modelo Federado. Actúa como puente entre el sistema de información y las Plataformas de Datos de los componentes descritos, así como también entre los Datos y los Componentes no definidos por este Modelo. El modelo de datos de este componente es presentado en la **Figura 11**.

5. Resultados

Esta sección presenta los resultados preliminares de la arquitectura a través de un prototipo que integro BOINC con los procesos Inicio de Sesión de Usuario y consulta

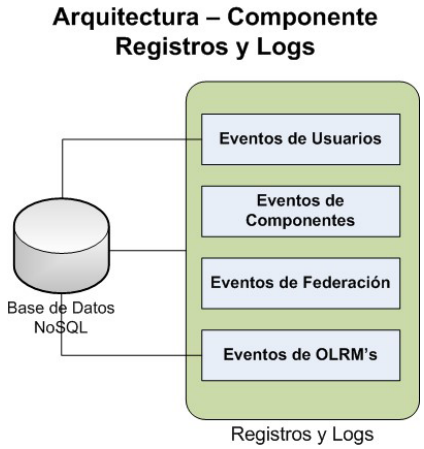


Figura 11 – Arquitectura y Modelo de Datos Componente - Registro y Logs.

y despliegue de información. Para la prueba de estos módulos se tuvo en cuenta los componentes sistema de información y monitoreo de políticas de usuario como se muestra en la **Figura 12**.

En la prueba de concepto fue necesario el despliegue de componentes con el fin de poder garantizar las funcionalidades de la federación, estos componentes fueron, **Directorio Institucional (Institución X), Proveedor de Identidades (Institución X), Proveedor de Servicios (Institución Y), Manejador Local de Recursos Oportunistas (BOINC Institución Y), Descubridor de Recursos**.

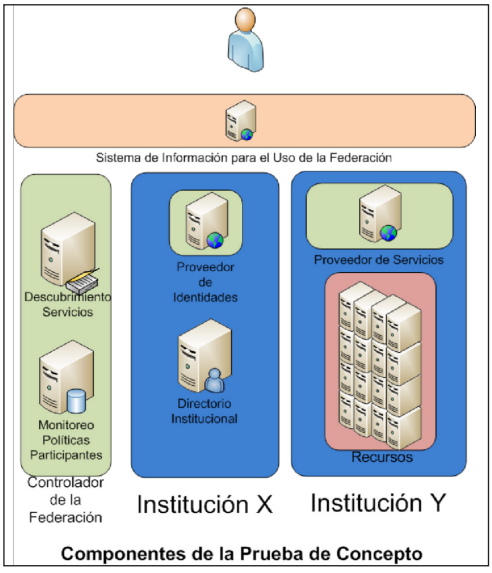


Figura 12 – Componentes del prototipo

En la puesta en marcha de la prueba de concepto de este proyecto se pudo probar los conceptos diseñados de manejo de recursos oportunistas entre ámbitos federados, alcanzando la superación de la etapa de concepción a satisfacción. Durante el desarrollo del prototipo, se puede observar, que la arquitectura propuesta puede utilizarse, no solo para compartir recursos oportunistas (OLRM), sino que también para otro tipo de servicios.

Durante la actividad de identificación de funcionalidades del modelo se plantearon un grupo estas, que se podían considerar como una extensión o abstracción del OLRM, se puede considerar a estos como meta-planificadores. Sin embargo, como parte del objetivo principal era realizar una extensibilidad de los OLRM más allá de la institución a la que pertenece y de esta manera brindar funcionalidades adicionales, entre las que se destacan: un Esquema de Gobernanza definido, estableciendo responsabilidades y límites dentro de la administración de componentes, esquema de autonomía en la administración de usuarios de la institución, manteniendo control de los accesos centralizados por institución, Control de políticas por instituciones y usuarios, Registro de tareas, modelado por usuario federado, Convergencia web, facilitando usabilidad hacia el usuario y extensibilidad de funcionalidades hacia y desde componentes de la arquitectura usando un único componente como punto de entrada de los usuarios.

6. Conclusiones

Uno de los retos más importantes que tienen los profesionales de infraestructura es la integración de servicios, debido a la variedad de tecnologías e implementaciones de diferentes estándares. La orquestación de servicios entre diferentes sectores incluyendo diferentes dominios administrativos como las instituciones, corresponde un reto superior. Sin embargo, la gobernanza de tecnologías se convierte en una herramienta fundamental en esquemas federados.

Los Sistemas informáticos cada vez se hacen más complejos, esto ha creado las arquitecturas multicapas, que buscan delegar responsabilidades en los componentes, sin embargo, la especialización de los componentes es una tendencia, en donde la característica específica y la simpleza se transfieren directamente en la arquitectura propuesta. La implementación de la arquitectura propuesta genera un gran esfuerzo, sin embargo, la creación de una federación como institución mundial y neutral, brinda cierto grado de transparencia, al realizar la publicación de servicios, así como permitir el consumo de los mismos por sus usuarios.

Según las funcionalidades añadidas en el proceso iterativo de diseño, el trabajo futuro para un modelo como el descrito en este documento, es lo suficientemente amplio como para alcanzar una definición de línea de trabajo alrededor de esta problemática, basada en el desarrollo de servicios masivos utilizando la arquitectura cloud federada definida.

Referencias

Baldión Albarracín, L. P., & Celis Baracaldo, L. M. (2015). *Análisis comparativo de las herramientas UnaCloud y BOINC en un ambiente oportunista*. Universidad Católica de Colombia. Retrieved from https://repository.ucatolica.edu.co/handle/10983/2777

Besso, L., & Striglio Raúl. (2011). *FuD-BOINC. Implementación de una capa de distribución del framework FuD utilizando como middleware a BOINC*. Universidad Nacional de Rio Cuarto. Retrieved from https://docplayer.es/5917123-Fud-boinc-implementacion-de-una-capa-de-distribucion-del-framework-fud-utilizando-como-middleware-a-boinc-lucas-besso-raul-striglio.html

Celesti, A., Fazio, M., Villari, M., & Puliafito, A. (2012). Virtual machine provisioning through satellite communications in federated Cloud environments. *Future Generation Computer Systems*, *28*(1), 85–93. https://doi.org/10.1016/J.FUTURE.2011.05.021

Coulouris, G., Dollimore, J., Kindberg, T., & Blair, G. (2012). *Distributed systems : concepts and design*. Addison-Wesley.

Diaz, C. O., Castro, H., Villamizar, M., Pecero, J. E., & Bouvry, P. (2013). Energy-aware VM Allocation on an Opportunistic Cloud Infrastructure. In *2013 13th IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing* (pp. 663–670). IEEE. https://doi.org/10.1109/CCGrid.2013.96

Diaz, C. O., Gomez, C. E., Castro, H. E., Barrios, C. J., & Bolivar, H. D. (2016). Federated Campus Cloud Colombian Initiative. In *2016 16th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid)* (pp. 712–717). IEEE. https://doi.org/10.1109/CCGrid.2016.48

Foster, I., Zhao, Y., Raicu, I., & Lu, S. (2008). Cloud Computing and Grid Computing 360-Degree Compared. In *2008 Grid Computing Environments Workshop* (pp. 1–10). IEEE. https://doi.org/10.1109/GCE.2008.4738445

Goiri, I., Guitart, J., & Torres, J. (2010). Characterizing Cloud Federation for Enhancing Providers' Profit. In *2010 IEEE 3rd International Conference on Cloud Computing* (pp. 123–130). IEEE. https://doi.org/10.1109/CLOUD.2010.32

Gonzalez Guerrero, E., Castro, H. E., & Rincon Yañez, D. A. (2012). Grid Colombia: Accomplishments and Challenge of a NGI in a Latin American. In *Proceedings of the Joint GISELA-CHAIN Conference* (Vol. 1, p. 10).

Hämmerle, L., Sabatino, R., Lenggenhager, T., Mantovani, M. L., Pilt, P., Toom, L., … Reetz, J. (2016). Comparison of Authentication and Authorisation Infrastructures for Research. Retrieved from https://publikationen.bibliothek.kit.edu/1000054481

Iqbal, S., Gupta, R., & Fang, Y.-C. (2005). Planning Considerations for Job Scheduling in HPC Clusters. *Dell Power Solutions*, 4.

Kertesz, A., Kecskemeti, G., & Brandic, I. (2014). An interoperable and self-adaptive approach for SLA-based service virtualization in heterogeneous Cloud environments. *Future Generation Computer Systems*, *32*, 54–68. https://doi.org/10.1016/J.FUTURE.2012.05.016

Rincon Yanez, D. A. (2016). *Prototipo de un modelo de arquitectura Cloud Federada utilizando recursos oportunistas*. Pontificia Universidad Javeriana.

Rochwerger, B., Breitgand, D., Levy, E., Galis, A., Nagin, K., Llorente, I. M., ... Galan, F. (2009). The Reservoir model and architecture for open federated cloud computing. *IBM Journal of Research and Development*, *53*(4), 4:1-4:11. https://doi.org/10.1147/JRD.2009.5429058

Rosales, E., Castro, H., & Villamizar, M. (2011). UnaCloud: Opportunistic Cloud Computing Infrastructure as a Service. *Conference on Cloud Computing, GRIDs, and Virtualization*, *2*(2), 8.

Seltsikas, P., & Currie, W. L. (n.d.). Evaluating the application service provider (ASP) business model: the challenge of integration. In *Proceedings of the 35th Annual Hawaii International Conference on System Sciences* (pp. 2801–2809). IEEE Comput. Soc. https://doi.org/10.1109/HICSS.2002.994341

Smith, M. (2000). *Definition of the inetOrgPerson LDAP Object Class*. Retrieved from https://tools.ietf.org/html/rfc2798

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 22/10/2018
Aceitação/Acceptance: 02/12/2018

Principios de la Web Semántica y Computación Afectiva en un Ecoturismo Sustentable mediante el Desarrollo de Aplicación Web Educativa

Pablo Alejandro Quezada-Sarmiento[1,2]

paquesar@alumni.uv.es, paquezadasa@uide.edu.ec

Universitat de Valencia, Facultad de Filosofía y Ciencias de la Educación, Doctorando del Programa de Doctorado en Educación, Av. Blasco Ibáñez 30, 46010 Valencia, España

Universidad Internacional del Ecuador – Dirección de Investigación y Posgrados, Quito, Ecuador

Pages: 212–221

Resumen: La Web se ha convertido en un recurso necesario de uso diario, los beneficios que ofrece, al ser una fuente de conocimiento y colaboración, dan lugar a nuevas iniciativas como los datos enlazados, cuyo propósito es vincular los datos esparcidos por la Web por medio de relaciones semánticas. El presente artículo da a conocer los principios la web semántica y su aplicabilidad en los contextos del ecoturismo y educativo. El aplicativo fue desarrollado por estudiantes de la asignatura de programación web de la Universidad Internacional del Ecuador en marco del proyecto de investigación de computación afectiva y herramientas de la Nube como un aporte educativo a la mejora de las destrezas de programación.

Palabras-clave: Computación Afectiva; Educación; Ecoturismo; Web Semántica; Economía; Software

Principles of the Semantic Web and Affective Computing in a Sustainable Ecotourism through the Development of Educational Web Application

Abstract: The Web has become a necessary resource for daily use, the benefits it offers, as it is a source of knowledge and collaboration, giving rise to new initiatives such as linked data, whose purpose is to link the data scattered through the Web through relationships semantics among these. This article shows the principles of the semantic web and its applicability in the contexts of ecotourism and educational. The application was developed by students of web programming course of the International University of Ecuador in the framework of the research project of affective computing and tools of the cloud as an educational contribution to the improvement of the programming skills.

Keywords: *Affective Computing; Education; Ecotourism; Semantic Web; Economics; Software*

1. Introducción

Las nuevas tecnologías y paradigmas computacionales pueden ser empleados en diversos campos, entre ellos el educativo lo cual sirve para el diseño, búsqueda, presentación, intercambio y reutilización de material debido a que la tecnología permite almacenar, organizar, replicar, difundir, transformar y ser accesible, lo que conlleva al ahorro de tiempo y recursos (Soto et al.,2017).

En el área educativa se ha visto la necesidad de crear espacios donde no existan limitantes de tiempo ni capacidad; es así, que la educación usa diversas plataformas para el proceso de enseñanza - aprendizaje. A nivel de la educación superior, especialmente en los campos de la Ingeniería Informática, el mundo académico muestra un entusiasmo significativo por el desarrollo de diversas competencias relacionadas a las competencias tecnológicas de ello la necesidad del adecuado uso de herramientas complementarias como son las herramientas de la nube especialmente enfocadas a la mejora de destrezas de programación que a su vez estén alineado a los estándares o cuerpos de conocimiento que rigen a dichas disciplinas y al contexto de la computación afectiva (Quezada & Mengual 2018).

El uso de diversas herramientas tecnológicas en educación superior integra competencias y habilidades, y une a docentes, investigadores y estudiantes, los cuales generan conocimiento. En la actualidad se requieren, por tanto, modelos educativos abiertos y flexibles al servicio de los estudiantes, que potencien el desarrollo de competencias aplicables a nivel profesional y que se vinculen con las diferentes áreas de conocimiento (Quezada, 2018).

En la actualidad existen un sinnúmero de iniciativas en el ámbito del turismo inteligente, pero cada una adaptada a su realidad y requerimientos específicos, analizando el contexto de explotación, los actores implicados y las barreras tecnológicas, sociales y de negocio que les afectan (Stone & Stone,2011).Las TICs y la Web Semántica son los medios para imaginar una nueva forma de satisfacer necesidades y concebir nuevos servicios que permitan la innovación en un sector como el ecoturístico, muy representativo en muchas economías nacionales (Hans, & Steen, 2014).

En un contexto en donde la oferta es cada vez más amplia, los nuevos servicios serán un elemento diferenciador para todos los agentes que de alguna manera se benefician del turismo. Puntos de información completos y accesibles, servicios de valor añadido disponibles en los hoteles, información de calidad en recintos históricos o sistemas de apoyo para guías turísticas (Meyer & Dorothea, 2004) , son algunos de los aspectos en las que pueden contribuir las tecnologías de localización y sobre las que se pueden implementar servicios innovadores que sin duda atraerán cada vez más a un turismo de calidad (Pompl, Wilhelm & Lavery ,1993).El ecoturismo es una alternativa posible para salir del emplazamiento e ingresar como protagonista que permite igualdad de condiciones de todos los involucrados. Está, entre otras razones, ha hecho que el turismo cobre importancia en los últimos años dando paso a que varios países incluyan dentro de sus políticas públicas, el desarrollo turístico sostenible de sus territorios (Smith & Randy, 2003). Es importante siempre incluir la perspectiva de la comunidad porque el ecoturismo no puede tener éxito sin el apoyo de la comunidad. Para un desarrollo turístico exitoso, los líderes comunitarios y los planificadores turísticos deben considerar

el turismo como una "*industria comunitaria local*". Es entonces cuando el ecoturismo se constituye en una alternativa de desarrollo económico y social que se integra fácilmente a las actividades productivas diarias de una determinada población, contribuye al empoderamiento del territorio por parte de los actores involucrados, e incrementa las oportunidades de empleo y el crecimiento económico en el lugar en donde se implementa (Quezada et al.,2018). Además, cuenta con un punto de partida efectivo para las poblaciones o comunidades rurales porque les permite impulsar su economía, pero sin dejar de lado la conservación del medio ambiente natural y etnográfico, logrando un desarrollo económico sostenible de la población involucrada (Ruschmann, 1997).

Zamora Chinchipe se ubica al sureste del Ecuador, en la región Amazónica, esta provincia ocupa una superficie de 10.556 km², posee una orografía montañosa, se encuentra situada alrededor de 970 metros sobre el nivel del mar, entre la confluencia de los ríos Zamora, Bombuscaro y Jamboe. Debido a su ubicación estratégica es poseedora de un alto grado de biodiversidad, razón por la cual es conocida como "*Ciudad de aves y Cascadas*". En referencia a la oferta de ecoturismo existente en la región están "Napo Galeras", "Comunidad Kichwa de Añangu", "Napo wildlife Center", "Guayzimi", "Comunidad Shayme", "Yankuam Lodge" que son proyectos que han tenido una buena aceptación por parte de las comunidades involucradas y receptoras como también por parte de los turistas que se destacan por esta modalidad de turismo. Por esta razón es que en Zamora Chinchipe (Ecuador) se pueden encontrar comunidades que han decidido ser parte de esta actividad como es el caso de Kiim, pueblo Shuar, donde se encuentra una reserva ecológica importante, con fauna y flora endémica, además de la riqueza cultural milenaria amazónica asociada a esta comunidad, poseen un Centro Shuar en el que ofrecen diferentes actividades turísticas (Quezada et al.,2018). Otra iniciativa de turismo comunitario existente en la zona es impulsada por la comunidad Napurak, la cual también es una etnia Shuar, que labora conjuntamente con los comuneros para poder aplicar esta nueva alternativa de desarrollo, permitiendo que generen ingresos económicos (Akkawi, 2010). El sector turístico ofrece sus servicios a través de nuevos intermediarios que cumplen una doble función; la de distribuidor de productos turísticos y la de gestor de servicios a empresas relacionadas con el sector turístico, lo que provoca que el internet ya no solo sea una fuente de información sino una herramienta de gestión. Debido a la integración de las TIC's se facilita el vivir estas experiencias, con el desarrollo y avances de la tecnología surgen nuevas actividades turísticas que permiten incrementar el valor añadido del servicio y satisfacer las nuevas y actuales demandas que tienen los turistas de hoy en día (Granda, 2017).

Es importante a la vez mencionar que "*cada vez más lugares rurales cuentan con las TIC's necesarias para lograr atraer a más visitantes, puesto que, gracias a estas herramientas, se han dado a conocer ante la mirada del mundo, lo cual ha generado un flujo de visitantes*" además la inclusión de las TIC´s soporta la idea de un turismo más sostenible (Vila & Vila, 2014).

La ayuda que el comercio electrónico, web brindan a la actividad turística permite que zonas rurales, lugares lejanos o muy remotos se den a conocer, ya sea esto por parte de los actores involucrados o por parte de los turistas de esa zona en cuestión (Gössling, 2017).

2. Metodología

Para el desarrollo del aplicativo web educativo se utilizó la metodología de desarrollo de software Scrum la cual es un marco de trabajo para el desarrollo así como un método de gestión de proyectos que se adapta a cualquier tipo de proyecto y está basada en un enfoque iterativo e incremental y de revisiones continuas.

Scrum es una metodología ágil, orientada a la gestión de proyectos, en el que se puede utilizar varios procesos y técnicas, “El marco de trabajo Scrum consiste en los Equipos Scrum, roles, eventos, artefactos y reglas asociadas (Quezada et al., 2017)”, (Soto et al,2017).

Scrum es un marco de trabajo incremental e iterativo para el desarrollo de proyectos, que se organiza en iteraciones de 1 a 4 semanas, llamados Sprint, los sprint son de un tiempo fijo, finalizan en una fecha determinada así no se haya terminado el sprint y se desarrollan uno después del otro. Al inicio de cada Sprint un equipo elige los requisitos del cliente de una lista priorizada, el equipo tiene reuniones diarias para informar el progreso del proyecto y una vez finalizado el Sprint, se presenta lo que se ha construido a los interesados del proyecto, ya que Scrum hace hincapié en productos funcionales listos para entregar (Quezada et al., 2017).

Se trabajó conjuntamente docentes, estudiantes e interesados en un entorno idóneo para el desarrollo del aplicativo adaptándose a las iteraciones de los Sprint y lecciones aprendidas enfocados en el producto funcional. La cooperación del cliente fue fundamental para comprender los principios del ecoturismo y verlos plasmados en el aplicativo web educativo.

Referente a las prácticas ágiles se realizó una planificación distribuida, con una adecuada descripción de las funcionalidades reflejadas en las historias de usuario. Dentro los elementos fundamentales fueron los entregables rápidos, así como las pruebas realizadas por los diferentes interesados.

Figura 1 – Grupo Control de Desarrollo de aplicativo Web

Dentro de la implementación Scrum se definieron los roles respectivos:

Scrum Master: Conocedor de Scrum (Docente tutor de la materia).

Product Owner: Persona que tiene la visión del producto a desarrollar.

Team Developer: Equipo encargado de desarrollar el producto (equipo de trabajo).

El grupo control estuvo integrado por estudiantes del componente educativo de Programación Web; quienes aplicaron los principios de la web semántica, computación afectiva bajo el marco de la metodología de desarrollo Scrum y principios de ecoturismo (Figura 1).

El conocimiento de las concepciones de los estudiantes sobre cómo se aprende la programación es un aspecto importante para los docentes. Con este conocimiento, los docentes pueden proporcionar actividades que ayudarán a sus estudiantes a avanzar a través de los niveles del espacio de resultados, desde menos exigentes cognitivamente y menos efectivos (permitiendo a los estudiantes ser buenos programadores) hasta más exigentes cognitivamente y más efectivos.

El conocimiento ya existe; el objetivo es establecer un consenso sobre el subconjunto del núcleo del conocimiento que caracteriza a la disciplina de la ingeniería informática. De la misma manera "*La articulación de un conjunto de conocimientos es un paso esencial hacia el desarrollo de una profesión, ya que representa un amplio consenso respecto a lo que un profesional de la ingeniería informática debe conocer, en especial en las destrezas computacionales de programación*" (Quezada et al.,2018).

3. Resultados

Dentro del proceso de desarrollo se estableció la arquitectura de software; la misma puede ser observada en la figura 2.

Uno de los objetivos de la presente investigación fue desarrollar el aplicativo con principios de la web semántica y computación afectiva que satisfaga la problemática de cubrir un ecoturismo inteligente y cuyo resultado el buscador semántico tal como lo muestra la figura. 2, donde se visualiza el resultado de una búsqueda como es los grafos, los filtros, categorías, etc. Así mismo en la figura 3 se puede apreciar el respectivo diagrama de clases, desarrollados por los estudiantes.

En la figura 4 se puede visualizar el diagrama de despliegue de la aplicación desarrollada.

Otro de los objetivos de la presente investigación fue desarrollar el aplicativo con principios de la web semántica que satisfaga la problemática de cubrir un ecoturismo inteligente y cuyo resultado el buscador semántico tal como lo muestra la figura 5, donde se visualiza el resultado de una búsqueda como es los grafos, los filtros, categorías, etc.

Dentro de la solución implementada se estableció diferentes parámetros de configuración esto con el fin garantizar la mayor accesibilidad de la aplicación, las cuales se listan a continuación:

Lenguaje: El usuario podrá elegir el idioma en el que desea obtener resultados.

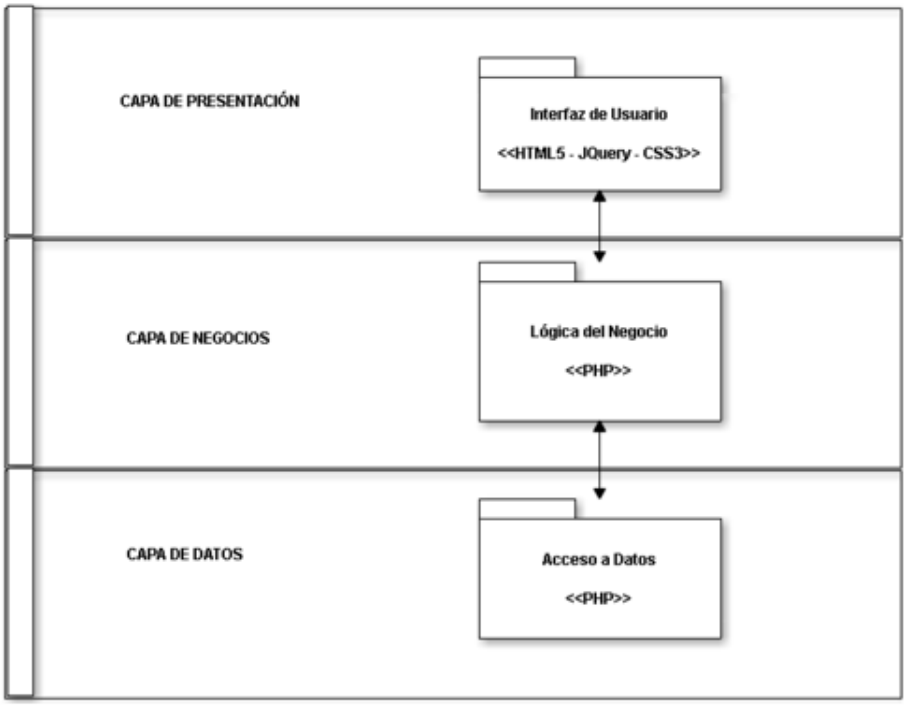


Figura 2 – Arquitectura de la aplicación semántica.

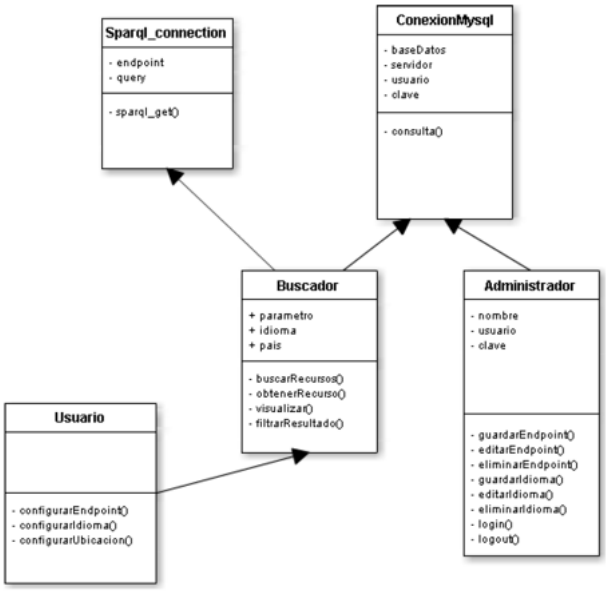


Figura 3 – Diagrama de Clases.

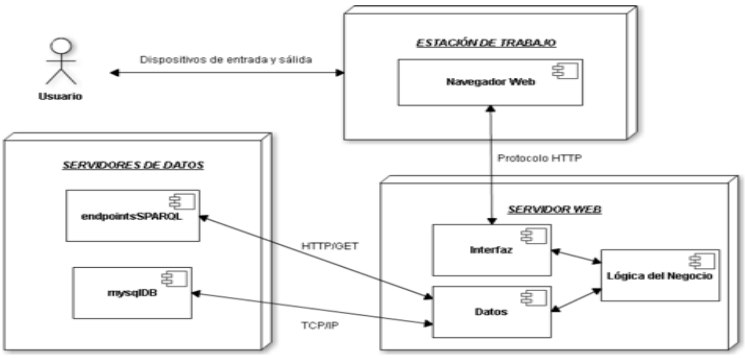


Figura 4 – Diagrama de despliegue.

Figura 5 – Interfaz del buscador semántico.

Buscar sobre: Al ser un buscador por coincidencia de palabras el usuario puede decidir sobre qué etiquetas orientar sus búsquedas.

Endpoints: La fuente de datos podrá ser cambiada para realizar búsquedas sobre una fuente en particular, así como también la opción de agregar una propia fuente de datos por medio de la dirección URL de un Endpoints.

Búsqueda Exacta: Esta opción le permite al usuario realizar búsquedas que contengan como parámetro el país de donde se conectan (Quezada & Ramirez, 2017) (Ver Figura 6)

Figura 6 – Interfaz del Buscador de Ecoturismo.

4. Conclusiones

El turismo ecológico constituye en el motor de desarrollo económico y social para las localidades y da lugar para que los visitantes puedan experimentar y disfrutar de otras formas de vivir y conocer el turismo; entre estas actividades la gastronomía como factor cultural, mediante la alimentación pueden conocer las costumbres tradiciones y cultura de estas localidades de tal forma que se fomenta el turismo vivencial que permite que los turistas se involucren con las comunidades, este tipo de turismo también fomenta el respeto por la naturaleza es así que gracias al desarrollo de las Tics, en los últimos años se ha conformado como el nuevo actor turístico, puesto a que la internet permite que se pueda llegar con la información a miles de usuarios que puedan acceder a la información en tiempo real y también permite que se conozca los lugares más recónditos de la tierra.

La información se presenta en tiempo real permitiendo al usuario conocer sobre las ventajas y características de los lugares turísticos lo cual facilita el acceso de la información a los usuarios y en tal virtud se pueda generar mayor difusión sobre los atractivos.

Mediante la aplicación de la web semántica y computación afectiva se logró mitigar la problemática identificada en el sector del ecoturismo, al desarrollar el buscador semántico y aplicar los principios de las metodologías agiles

El desarrollo del aplicativo semántico permitió generar un trabajo colaborativo con los profesionales en formación del componente de ingeniería web, así mismo se profundizó en la enseñanza del paradigma orientado a objetos aportando en el contexto del proyecto computación afectiva y herramientas de la nube.

Referencias

Akkawi, M. H. (2010). Resident Attitudes towards Tourism Development in Conservative Cultures: The Case of Qatar. University of Waterloo, Ontario, Canada.

Antonson, Hans, & Jens Kr Steen Jacobsen. (2014). "Tourism development strategy or just brown signage? Comparing road administration policies and designation procedures for official tourism routes in two Scandinavian countries." Land Use Policy, Vol. 36, pp. 342-350.

Gössling, S. (2017). Tourism, information technologies and sustainability: an exploratory review. Journal of Sustainable Tourism, 25(7), 1024–1041.

Goodwin (1996), cit. en Fiona Burton (1998), ìCan ecotourism objectives be achieved? î, Annals of Tourism Research, vol.25, n m. 3, p. 756.

Granda Rojas, O. E. (2017). Proyecto de equipamiento arquitectónico ecoturístico comunitario para el centro shuar Kiim, parroquia la paz, cantón Yacuambi, provincia de Zamora

Meyer, Dorothea (2004). "Tourism routes and getaways: key issues for the development of tourism routes and gateways and their potential for pro-poor tourism." Londres: Overseas Development Institute.

Pompl, Wilhelm and Patrick Lavery (1993). Tourism in Europe: structures and developments. England: Cab International.

Quezada-Sarmiento, P. A., Chango-Cañaveral, P. M., López-Criollo, J., Pacheco-Viteri, F. A., & Enciso, L. (2017). Design of a wireless network of temperature and lighting sensor for gastronomic laboratories under the principles of agile scrum methodology. [Diseño de una Red de sensores inalámbricos de temperatura e iluminación para laboratorios Gastronómicos bajo principios de la metodología ágil Scrum] Espacios, 38(46)

Quezada-Sarmiento, P. A., Chango-Cañaveral, P. M., Benavides-Cordova, V. M., Ramirez-Coronel, R. L., & Enciso, L. (2017). System monitoring for bridges structure. [Desarrollo de un Sistema de Monitoreo para estructuras de puentes] Espacios, 38(61)

Quezada-Sarmiento, P. A., & Ramirez-Coronel, R. L. (2017). Develop, research and analysis of applications for optimal consumption and visualization of linked data. Iberian Conference on Information Systems and Technologies, CISTI, 10.23919/CISTI.2017.7975964

Quezada-Sarmiento, Pablo Alejandro and Mengual Andrés, Santiago Implementación de una solución web y móvil para la gestión vehicular basada en Arquitectura de Aspectos y metodologías ágiles: Un enfoque educativo de la teoría a la práctica. RISTI, Dic 2017, no.25, p.98-111. ISSN 1646-9895

Quezada-Sarmiento, P. A., Suasnavas-Rodriguez, M. G., Chango-Canaveral, P. M., Gonzaga-Vallejo, C., Enciso, L., & Calderon-Cordova, C. A. (2018). Used of social networks and web application to design and promote the ecotourism route in the southern amazon of ecuador. Paper presented at the Iberian Conference on Information Systems and Technologies, CISTI, , 2018-June 1–7. doi:10.23919/CISTI.2018.8399399

Quezada-Sarmiento, P. A., Enciso-Quispe, L. E., Jumbo-Flores, L. A., & Hernandez, W. (2018). Knowledge representation model for bodies of knowledge based on design patterns and hierarchical graphs. Computing in Science and Engineering, doi:10.1109/MCSE.2018.2875370

Quezada-Sarmiento, P. A., Mayorga-Diaz, M. P., Paredes-Cruz, R. E., Guaigua-Vizcaino, M. I., Ulehlova, E., Chango-Cañaveral, P. M., & Enciso, L. (2018). Determination of leadership factors that influence the development of innovation in intensive software systems through literature mapping. Paper presented at the Proceedings of 2018 the 8th International Workshop on Computer Science and Engineering, WCSE 2018, 406-410.

Ruschmann, D. (1997) “Turismo e planejamento sustentável: a proteção do meio ambiente”. Campinas: Papirus, 1997

Soto-Guerrero, F. M., Quezada-Sarmiento, P. A., Condolo-Herrera, L. C., Mengual-Andrés, S., Moreno-León, D. I., & Rey-Mendoza, I. L. (2018). Development of a methodological framework for educational innovation in the context of the tutorial action based on agile methodologies and software engineering knowledge standards. [Desarrollo de un framework metodológico para la innovación educativa en el contexto de la acción tutorial basado en metodologías ágiles y estándares de conocimiento de ingeniería de software] RISTI - Revista Iberica De Sistemas e Tecnologias De Informacao, (E15), 233–242

Smith, Randy. (2003). Manual de ecoturismo para la amazonia ecuatoriana. Ecuador: Ediciones Abya - Yala.

Stone, L. S., & Stone, T. M. (2011). Community-based tourism enterprises: challenges and prospects for community participation; Khama Rhino Sanctuary Trust, Botswana. Journal of Sustainable Tourism, 19(1), 97–114.

Vila, T. D., & Vila, N. A. (2014). Gestión de las redes sociales turísticas en la web 2.0. Vivat Academia, (129), 57–78.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 15/09/2018
Aceitação/Acceptance: 05/11/2018

Impactos da implantação da Telemedicina no Tratamento e Prevenção do Câncer

Antônio Augusto Gonçalves[1], Carlos Henrique Fernandes Martins[1], Sandro Luís Freire de Castro Silva[1], Cezar Cheng[1], Roberto Luiz Silva dos Santos[1], Saulo Barbará de Oliveira[2]

augusto@inca.gov.br, cmartins@inca.gov.br, sandrofreire@gmail.com, cheng@inca.gov.br, roberto.santos@inca.gov.br, saulobarbara@gmail.com

[1] Instituto Nacional de Câncer José de Alecar Gomes da Silva, Rua do Rezende 195, 20230-026, Rio de Janeiro, Brasil.

[2] Universidade Federal Rural do Rio de Janeiro, BR-465 Km7, 23897-000, Seropédica – Rio de Janeiro, Brasil.

Pages: 222–230

Resumo: A Telemedicina é apresentada como uma solução inovadora para os desafios no setor da saúde proporcionando um melhor acesso aos pacientes em casa ou em organizações de cuidados primários, aumentando o acesso ao cuidado especializado, reduzindo os problemas devido à dispersão geográfica e melhorando a qualidade de vida dos cidadãos. O tratamento e prevenção do câncer é um processo longo, complexo e de alto risco. A implementação da Telemedicina pode ser considerada como uma inovação disruptiva que muda radicalmente a relação médico-paciente. Este estudo de caso pretende descrever a utilização da Telemedicina na redução de custos e divulgação das informações para o atendimento e tratamento de pacientes no Instituto Nacional do Câncer (INCA).

Palavras-chave: Telemedicina, Tecnologia da Informação e Saúde.

Impacts of the implementation of Telemedicine in the Treatment and Prevention of Cancer

Abstract: Telemedicine is presented as an innovative solution to the challenges in the health sector by providing better access to patients at home or in primary care organizations, increasing access to specialized care, reducing problems due to geographic dispersion and improving quality of life of citizens. Treatment and prevention of cancer is a long, complex and high-risk process. The implementation of Telemedicine can be considered as a disruptive innovation that radically changes the doctor-patient relationship. This case study intends to describe the use of Telemedicine in reducing costs and disseminating information for the care and treatment of patients at the National Cancer Institute (INCA).

Keywords*:* Telemedicine, It and Health.

1. Introdução

A maioria dos países está enfrentando graves problemas nos serviços de saúde com o aumento da demanda de cuidados de saúde devido ao crescimento na expectativa de vida da população aliado ao aumento de doenças crônicas; exigência de maior acessibilidade aos cuidados extra-hospitalares, necessidade de maior eficiência, equidade e individualização dos cuidados de saúde, recursos financeiros limitados, além das dificuldades de recrutamento e retenção de profissionais treinados.

Existe uma enorme pressão visando a melhoria dos recursos médicos que além de escassos são de alto custo, tais como pessoal especializado e disponibilidades de leitos hospitalares. Estes desafios ajudam a transformar os serviços de saúde remotos em uma das áreas de grande potencial de crescimento na área de assistência. A Telemedicina é apresentada como uma solução inovadora para os futuros desafios no setor da saúde causados por mudanças demográficas. A Telemedicina pode ser usada na prestação de serviços de saúde através do uso de tecnologias de informação e comunicação, em situações em que os participantes estão em locais diferentes (Heinzelmann et al., 2005).

A Telemedicina proporciona um melhor acesso aos pacientes em casa ou em organizações de cuidados primários, aumentando o acesso ao cuidado especializado, reduzindo os problemas devido à dispersão geográfica e melhorando a qualidade de vida dos cidadãos (Kamsu & Foguem, 2014).

A Telemedicina é recomendada como uma solução para facilitar a criação de um serviço de saúde mais integrado, com o objetivo de aumentar o acesso, qualidade, satisfação e eficiência do tratamento dos pacientes utilizando diferentes tecnologias de informação e telecomunicações inovadoras para compartilhar informações. O desenvolvimento da Telemedicina transformou radicalmente as atividades médicas e as interações do médico com grande impacto na relação entre médicos e pacientes. Além disso, as informações trocadas entre o médico e outros especialistas podem apoiar a tomada de decisão (Kamsu, Tiako, Fotso & Foguem, 2015).

Agha et al. (2009) observaram que a Telemedicina tem desempenhado um papel importante na redução de custos na medicina e, ao mesmo tempo, aumentado a acessibilidade aos cuidados de saúde, particularmente em áreas remotas e carentes. Além disso, os resultados de uma pesquisa realizada para comparar consulta de Telemedicina e consulta de internação sugerem que os pacientes estavam igualmente satisfeitos com os dois tipos de consulta, considerando a competência do médico e sua habilidade em um relacionamento interpessoal.

A implementação eficiente da Telemedicina depende do entendimento correto do impacto da prestação de um serviço de Telemedicina para diferentes áreas. Outras etapas do plano de implementação, como o desenvolvimento de um plano de serviços, o desenvolvimento de um plano de negócios e a tecnologia relacionada devem seguir a avaliação e a priorização da intervenção necessária, o que deve melhorar a eficácia, a eficiência e a qualidade dos cuidados de saúde. (Rasmussen, Jensen, Froekjaer, Kidholm, Kensing & Yderstraede, 2015).

Thakur et al. (2012) ressaltam que apesar da lenta aceitação, a adoção da Telemedicina nos serviços de saúde tem sido efetiva. A ineficiência ainda existe na maioria dos serviços de saúde mas é necessário entender como superar essas ineficiências usando inovação.

As percepções das diversas partes interessadas sobre os benefícios da Telemedicina, o consenso entre os participantes e suas interações podem influenciar positivamente o processo de implantação.

A Telemedicina pode ser projetada para facilitar a criação de um serviço de saúde mais integrado, com o objetivo de aumentar o acesso, a qualidade, a satisfação do paciente e a eficiência do tratamento em pacientes com câncer. A pesquisa e o tratamento relacionados ao câncer são um processo longo, complexo e de alto risco e envolvem uma infinidade de organizações. A implementação da Telemedicina pode ser considerada como um tipo de inovação disruptiva que muda radicalmente a relação médico-paciente (Gonçalves et al., 2018).

2. Telemedicina

A Organização Mundial de Saúde adota a seguinte descrição da Telemedicina: A prestação de serviços de saúde, onde a distância é um fator crítico, por todos os profissionais de saúde que usam tecnologias de informação e comunicação para a troca de informações válidas para diagnóstico, tratamento e prevenção de doenças e lesões, pesquisa e avaliação, e para a educação continuada dos profissionais de saúde, tudo no interesse de promover a saúde dos indivíduos e suas comunidades (Who, 1997).

A Telemedicina é um conceito que utiliza diversas tecnologias para oferecer treinamento de profissionais de saúde e atendimento domiciliar à pacientes, abrange um campo de investigação interdisciplinar realizada em diferentes áreas de pesquisa. Envolve, por exemplo, hardware como vídeo conferências, sistemas de colaboração e de imagens entre outros, incluindo serviços de consultoria e diagnóstico, além de facilitar o planejamento, coordenação, colaboração e educação (Singh et al., 2009). No entanto, apesar da promessa, a adoção de iniciativas de Telemedicina, tem sido lenta, desigual e de escopo limitado (Chandwani & Dwivedi, 2015, Sims, 2016).

Simon, et al. (2012) destacam o desenvolvimento da Telemedicina e a transformação radical das atividades médicas e interações médico-paciente. Além disso, as informações trocadas entre médico e outros especialistas devem apoiar a tomada de decisão e trazer benefícios às organizações e pacientes, permitindo o acesso adequado a serviços de saúde de alta qualidade e com boa relação custo-benefício. através das seguintes atividades médicas:

- Tele-assistência - permite que um profissional médico auxilie remotamente outro profissional de saúde durante a realização de um ato.
- Teleconsulta - visa permitir que um profissional médico faça uma consulta remota a um paciente e, se necessário, um profissional de saúde possa estar presente;
- Tele-expertise - visa permitir que um profissional médico solicite remotamente a opinião de um ou mais especialistas médicos que possuam formação ou competências especiais;
- Monitoramento Remoto - permite que um profissional médico interprete à distância os dados necessários para o acompanhamento médico de um paciente e, quando apropriado, tome decisões relacionadas ao cuidado do paciente. O registro e transmissão de dados podem ser automatizados ou realizados pelo próprio paciente ou por um profissional de saúde;

Saliba et al. (2012) ressaltam que nos países em desenvolvimento a Telemedicina pode aumentar o acesso ao atendimento especializado, reduzindo os impactos da dispersão geográfica e das desigualdades de oferta de serviços de saúde, melhorando a qualidade de vida dos cidadãos. A Telemedicina tem o potencial de trazer benefícios para os sistemas de saúde e pacientes, facilitando o acesso a serviços de saúde de alta qualidade e com boa relação custo-benefício, particularmente em países de baixa renda. A Telemedicina utiliza tecnologias avançadas de telecomunicações para o intercâmbio de informações e prestação de serviços de saúde superando barreiras geográficas, sociais e culturais, proporcionando um meio flexível para conectar profissionais de saúde e pacientes provenientes de regiões geograficamente afastadas, reduzindo a necessidade deslocamento.

Um grande número de aplicações de Telemedicina foram implementadas por organizações públicas, institutos de pesquisa e serviços de saúde privados. Apesar dos benefícios esperados e do estabelecimento de redes nacionais e de classe mundial, a adoção da Telemedicina e sua integração como um serviço de saúde rotineiro tem enfrentado grandes desafios. Poucos estudos foram realizados para apoiar a implementação de serviços de Telemedicina nas organizações de saúde. Os gestores de saúde devem focar na adoção e confiabilidade da tecnologia e em fatores organizacionais, como processo, estrutura, cultura e gestão (Rasmussen et al., 2015).

Wallace et al. (2012) observam que a implementação da Telemedicina introduz frequentemente mudanças radicais nos serviços de saúde que podem ser negligenciadas devido à falta de conhecimento e planejamento. Entretanto, essas modificações no processo clínico podem ter impacto severo sobre a eficácia dos resultados. Muitas vezes, a relação custo-benefício dessas aplicações não é demonstrada e questões relativas à análise qualitativa do nível de satisfação entre profissionais de saúde, pacientes e parentes não são explicitados.

Em suas pesquisas Vlmarlund et al. (2013) ressaltam que diferentes barreiras continuam a desafiar a adoção generalizada de Telemedicina pelas organizações de saúde. Essas barreiras incluem tecnologia, restrições financeiras, normas legais, estratégia de negócios e questões de recursos humanos. O primeiro passo para o desenvolvimento de uma implementação bem-sucedida é a avaliação das necessidades das partes interessadas e dos requisitos técnicos. A Telemedicina deve ser orientada pelos pedidos de pacientes e médicos, e não pela tecnologia disponível. Cada parte interessada tem os seus requisitos específicos que devem ser abordados

Para alcançar o potencial da Telemedicina em um cenário de recursos escassos, é necessário que os pesquisadores desenvolvam uma compreensão abrangente e profunda de como o processo de Telemedicina funciona nesses contextos, incluindo uma análise de processos, contingências, estruturas e mecanismos causais. Esse profundo conhecimento é altamente valioso para gerentes e profissionais envolvidos no projeto e implementação da Telemedicina.

Esta pesquisa foca nos mecanismos que fundamentam a Telemedicina ter uma implantação bem-sucedida em ambientes de poucos recursos, como no Brasil. Especificamente, examinando as interações médico-paciente e médico-especialista com uma iniciativa de Telemedicina buscando entender quais são os mecanismos relativos ao projeto e implementação de Telemedicina, especialmente em ambientes de poucos recursos.

3. Metodologia

Este estudo de caso qualitativo pretende descrever a viabilidade da utilização da Telemedicina na redução de custos e divulgação das informações para o atendimento e tratamento de pacientes no Instituto Nacional do Câncer (INCA). O INCA é um órgão do Ministério da Saúde, que tem como responsabilidade o desenvolvimento de ações e a coordenação das mesmas para a prevenção e controle do câncer no Brasil. As unidades hospitalares do INCA estão localizadas no município do Rio de Janeiro, integrando o Sistema Único de Saúde (SUS) e oferecem tratamento integral às pessoas que têm câncer.

Segundo Yin (2009), o estudo de caso é um método de pesquisa que investiga um fenômeno contemporâneo em seu ambiente natural, adotando múltiplas fontes de evidência sobre uma ou poucas entidades, sem o uso de manipulação ou controle. O estudo de caso pode ser classificado em três tipos quanto ao objetivo da pesquisa: descritivo - descreve o fenômeno dentro de seu contexto; exploratório – trata com problemas pouco conhecidos com o objetivo de definir hipóteses para futuras pesquisas e explanatório – investiga explicações para problemas levantados em estudos anteriores.

A intenção é contribuir com a divulgação de resultados descritivos voltados para a Telemedicina implantada no INCA, apresentando o que, como e onde se encontram os aspectos investigados (Yin, 2009).

4. O Caso do Instituto Nacional de Câncer

O projeto de Telemedicina do INCA nasceu do desejo da instituição de participar ativamente do desenvolvimento de políticas do Ministério da Saúde, especialmente no uso da Telemedicina na capacitação de profissionais de saúde, na gestão do conhecimento e nas ações de prevenção e controle do câncer na Rede de Atenção Oncológica.

Em 2009, o INCA inaugurou o seu Polo de Telemedicina passando a fazer parte da Rede Universitária de Telemedicina – RUTE.

As ações de Telemedicina do Instituto têm por objetivo a execução de atividades técnico-científicas voltadas ao controle e prevenção do câncer, que possibilitem a capacitação e o acompanhamento remoto de técnicos das secretarias de saúde e demais instituições parceiras, que ofereçam suporte aos profissionais de saúde, em reuniões virtuais, discussão de casos ou em educação continuada e no suporte ao paciente. Essas atividades visam fortalecer o processo de qualificação dos profissionais envolvidos nas ações nacionais de controle do câncer e aprimorar os mecanismos de gestão das ações nacionais de controle do câncer, além de buscar a melhora da saúde e bem-estar do paciente, utilizando tecnologias de informação e comunicação.

O objetivo deste estudo qualitativo foi examinar as mudanças nos serviços de saúde causadas pela implementação de um sistema de Telemedicina sob a perspectiva dos profissionais de saúde e participantes-chave. Foram identificadas experiências para os quatro eixos principais: processo, estrutura, cultura e gestão.

Na área de assistência médica hospitalar o processo é continuo e encadeado. As unidades hospitalares do INCA possuem serviços de saúde especializados com diversas clínicas

médicas, onde os pacientes, após as consultas com oncologistas e o diagnóstico do câncer, iniciam o processo de tratamento. Este pode ser ambulatorial ou via internação.

A implementação eficiente da Telemedicina depende do entendimento correto do impacto da prestação de um serviço de Telemedicina para diferentes áreas. Outras etapas do plano de implementação, como o desenvolvimento de um plano de serviços, o desenvolvimento de um plano de negócios e a tecnologia relacionada devem seguir a avaliação e a priorização da intervenção necessária, o que deve melhorar a eficácia, a eficiência e a qualidade dos cuidados de saúde.

A solução deve utilizar tecnologias avançadas de telecomunicações para o intercâmbio de informações e prestação de serviços de saúde superando barreiras geográficas e proporcionando uma alternativa de baixo custo para conectar profissionais de saúde e pacientes provenientes de regiões geograficamente afastadas da unidade hospitalar.

As aplicações de Telemedicina são muito amplas, pois podem abranger desde o primeiro contato entre o médico e o paciente, passando pelo apoio remoto a formulação diagnóstica, pelo tratamento clínico até intervenções cirúrgicas, com o auxílio de diferentes meios de comunicação. Assim, esse método é capaz de conectar centros de referência de tratamento de câncer com unidades de atendimento da periferia, proporcionando assistência em dúvidas diagnósticas, de condutas e ensino. Este projeto está em andamento com o objetivo de disponibilizar aplicativos desenvolvidos no INCA para a Rede de Atenção Oncológica (RAO).

Nos últimos anos, o Serviço de Tecnologia do INCA cresceu significativamente e passou a auxiliar na conduta, mudando fundamentalmente a capacidade de desenvolver sistemas de informação para apoiar o processo de cuidados à saúde. Os recentes avanços nas Tecnologias de Informação e Comunicação (TIC) têm tornado o uso da Telemedicina cada vez mais viável e acessível, podendo ser vista como um meio para proporcionar melhores cuidados médicos. Smartphones, computadores ou outros recursos eletrônicos podem ser usados para acessar sistemas de informação ou comunicação utilizando a Internet.

Mais especificamente, a utilização da TIC para o desenvolvimento da prática médica a distância pode abrir novas possibilidades na oferta de serviços de saúde, especialmente no atendimento domiciliar. Dentre as aplicações mais conhecidas estão à videoconferência médica, os trabalhos colaborativos de segunda opinião, consultas on line e telediagnósticos por imagem. Este amplo leque de possibilidades tem viabilizado o desenvolvimento de diversos aplicativos com o objetivo de aumentar a acessibilidade aos serviços de saúde por uma parcela cada vez maior da sociedade.

A vantagem de usar uma plataforma de comunicação padronizada vem progressivamente sendo adotada nos serviços de prevenção e tratamento de diversas patologias. A utilização da web para aplicações de Telemedicina parece ser um caminho definitivo. A Internet tem se tornado a infraestrutura padrão para o acesso às aplicações sofisticadas.

Este ambiente tecnológico garante vantagens de acessibilidade e usabilidade aos pacientes e médicos. Entretanto, há diversas questões que devem ser detalhadas e discutidas. Com referência particular a todas as aplicações relativas aos serviços de cuidado à distância, conhecidas por aplicações de telecare. Nestes aplicativos o papel do paciente torna-se central já que ele é envolvido ativamente no processo de tratamento.

Mesmo com a crescente utilização da internet no Brasil, ainda existe uma barreira à adoção de soluções baseadas na utilização de aplicativos que necessitem de tecnologias sem fio ou wireless, para acesso a rede. Diversas aplicações estão sendo desenvolvidas com funcionalidades que podem ser ativadas via palmtops e smartphones. O desafio é tornar tecnologias de comunicação como a 3G e 4G mais acessíveis a médicos e pacientes.

O Serviço de Tecnologia da Informação do INCA desenvolveu em conjunto com os profissionais de saúde, da área de atendimento domiciliar, uma aplicação que roda em smartphones com tecnologia wireless, 3G ou 4G. Esse aplicativo provê acesso às informações dos pacientes atendidos pelo serviço de Assistência Domiciliar do INCA e realizar coleta de dados das visitas, prescrições e pedidos de materiais através de smartphones. As principais funcionalidades do aplicativo são: prescrição de medicamentos; aprazamento de medicamentos; controle da dispensação e administração ao paciente; consulta as prescrições médicas; e controle de materiais enviados aos pacientes.

A telemedicina como uma ferramenta de interação paciente-médico remota está crescendo e consultas virtuais com especialistas são possíveis. Os benefícios potenciais da telemedicina incluem melhor acesso aos cuidados de saúde, menor tempo de espera para consultas e maior adesão aos protocolos de tratamento.

4.1. Resultados

A análise dos resultados indicou que a implantação desse aplicativo no INCA foi fundamental para diminuir a fragmentação do atendimento domiciliar e aumentar a agilidade no tratamento, permitindo um melhor aproveitamento dos recursos humanos com maior acesso às informações dos pacientes e redução dos custos de tratamento. O sistema proporciona um meio flexível para conectar profissionais de saúde que atendem pacientes geograficamente afastados com redução da necessidade de viagens. A implantação do sistema tem como principais objetivos:

- Melhorar as condições de vida dos pacientes com câncer sem chance de cura;
- Reduzir o tempo e custos de atendimento a pacientes;
- Permitir o acesso rápido a informações do paciente;
- Reduzir a pressão sobre hospitais, já comprometidos pela falta de leitos e recursos humanos;
- Permitir o uso mais efetivo do corpo clínico, através da centralização de especialistas e da descentralização da assistência, alcançando um número maior de pessoas;
- Permitir a cooperação e integração de pesquisadores com o compartilhamento de registros clínicos.

5. Conclusão

O objetivo desta pesquisa foi estudar o sistema de telemedicina implantado e sua adoção por médicos do INCA. A telemedicina no Instituto vem evoluindo, principalmente pela adoção de tecnologias inovadoras. A capacidade de inovação do INCA facilitou a adoção da Telemedicina, entretanto, dentro do instituto existem áreas mais propensas que outras, desta forma a adoção da Telemedicina não ocorreu de forma homogênea.

A telemedicina vem mudando a relação médico-paciente, por esta razão, médicos e pacientes devem confiar no uso da telemedicina. A confiança é um requisito importante para a adoção de inovação e pode, portanto, ser fundamental para os médicos na tomada de decisões sobre o uso da telemedicina.

O envolvimento dos profissionais de saúde desde a concepção do projeto até a implantação e uso efetivado do aplicativo foi outro fator fundamental para o sucesso desse projeto. O uso do smartphone, por ser uma tecnologia amplamente utilizada pelos profissionais também facilitou a adoção.

Outro fator importante está relaciona a organização do processo de compartilhamento de informações entre o médico solicitante e outros especialistas porque facilita a rastreabilidade e transparência do processo de telemedicina. Isso requer a definição de padrões adequados de intercâmbio de dados adaptado às necessidades de avaliação de desempenho e da gestão segura da informação. Portanto, a confidencialidade deve ser respeitada, e a proteção dos dados deve ser garantida através de princípios éticos e regras formais de segurança da informação relativas ao compartilhamento de dados dos pacientes.

A implantação da Telemedicina requer uma organização rigorosa sob a responsabilidade de médicos e cientistas da computação, com base em regras claras e apoiada por plataformas de gestão de conhecimento e processos de compartilhamento de informações para ambientes de tomada de decisão colaborativa e compartilhada.

Referências

Agha Z., Schapira R.M., Laud P.W., Mcnutt G., Roter D.L. (2010) Patient satisfaction with physician-patient communication during telemedicine. *Telemed. J. EHealth Off*. J. Am. Telemed. Assoc. 15 (9), 830e839.

Armfield, N.R., Edirippulige, S.K. Edirippulige, B. N. Bradford, Smith, C. (2014), Telemedicine is the cart being put before the horse? *Medical Journal of Australia*. 200 (9), 530–533.

Gonçalves A.A., Martins C.H.F., Barbosa J.G.P., de Castro Silva S.L.F., Cheng C. (2018) Telemedicine Collaboration in Cancer Treatment: A Case Study at Brazilian National Cancer Institute. *Lecture Notes in Computer Science*, vol 10898.

Heinzelmann, P.J., Lugn, N.E. Kvedar, J.C .(2005). Telemedicine in the future, J. *Telemed. Telecare* 11 (8) 384–390.

Kamsu, F.B, Foguem, C. (2014). Telemedicineand mobile health with integrative medicine in developing countries. *Health PolicyTechnol*;3(4):264–71.

Kamsu, F.B., Tiako, P., Tiako F., Fotso, L.P., Foguem, C. (2015). Modeling for effective collaboration in telemedicine, *Telematics and Informatics*, Volume 32, Issue 4, November Pages 776-786.

Rasmussen B. S. B., Jensen L. K., Froekjaer C., Kidholmb K., F. Kensing F. (2015). A qualitative study of the key factors in implementing telemedicalmonitor-ing of diabetic foot ulcer, *International Journal of Medical Informatics* 84 799–807

Rasmussen, B.S.B., Jensen L.K., Froekjaer, J., Kidholm, K., Kensing F., Yderstraede K.B. (2015). A qualitative study of the key factors in implementing telemedical monitoring of diabetic foot ulcer patients. *International Journal of Medical Informatics*;84:799–807.

Saliba, V., Legido-Quigley H., Hallik R., Aaviksoo A., Car J., McKe e M. (2012). Telemedicine across borders: a systematic review of factors that hinder or support implementation. *Int J Med Inform* ;81(12):793–809.

Simon, P., Pellitteri, Williatte. (2012). Le décret français de télémédecine: une garantie pour les médecins. *European Research in Telemedicine*. 1 (2), 70–75.

Singh R., Mathiassen L., Mishra, A.N (2009). A Theory of rural telehealth innovation-a paradoxical approach. *ICIS 2009 Proceedings*, p. 126.

Thakur, R., Hsu, S. H. Y., & Fontenot, G. (2012). Innovation in healthcare: Issues and future trends., *Journal of Business Research*, 65(4), 562–569.

Vlmarlund V, Le Rouge, C. (2013) Barriers and opportunities to the widespread adoption of telemedicine: a big-country evaluation. *Studies in Health Technology and Informatics*, 192:933, 2013

Wallace, D.L., Hussain A., Khan N., Wilson Y.T. (2012) A systematic review of the evidence for telemedicine in burn care: with a uk perspective, Burns: J. Int. Soc. Burn Injuries 38 (4) (2012) 465–480.

WHO Group Consultation on Health Telematics (1997: Geneva, Switzerland). (1998). A health telematics policy in support of WHO's Health-for-all strategy for global health development: report of the WHO Group Consultation on Health Telematics, 11-16 December, Geneva, 1997. Geneva: World Health Organization. http://www.who.int/iris/handle/10665/63857.

Yin, R. K. (2009). Case study research: Design and methods (4th Ed.). Thousand Oaks, CA: Sage.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 03/10/2018
Aceitação/Acceptance: 07/12/2018

A Governança de Tecnologia da Informação e Comunicação em Organizações de Saúde Pública: Estudo de Caso do Instituto Nacional de Câncer

Antônio Augusto Gonçalves[1], Carlos Henrique Fernandes Martins[1], Sandro Luís Freire de Castro Silva[1], Cezar Cheng[1], Roberto Luiz Silva dos Santos[1], Saulo Barbará de Oliveira[2]

augusto@inca.gov.br, cmartins@inca.gov.br, sandrofreire@gmail.com, cheng@inca.gov.br, roberto.santos@inca.gov.br, saulobarbara@gmail.com

[1] Instituto Nacional de Câncer José de Alecar Gomes da Silva, Rua do Rezende 195, 20230-026, Rio de Janeiro, Brasil.

[2] Universidade Federal Rural do Rio de Janeiro, BR-465 Km7, 23897-000, Seropédica – Rio de Janeiro, Brasil.

Pages: 231–243

Resumo: A administração pública brasileira é baseada nos modelos da nova gestão pública que buscam reduzir gastos, melhorar os serviços oferecidos e dar mais transparência na gestão de recursos públicos. Neste cenário, o setor público procura melhorar a governança corporativa auxiliando suas organizações no processo de tomada de decisões. O foco sobre governança de TIC se intensificou a medida que os governos se tornaram grandes consumidores de novas tecnologias utilizadas na modernização da administração pública. O objetivo deste artigo foi avaliar a implantação da governança de TIC em organizações de saúde pública através de um estudo de caso realizado no INCA. Os resultados da pesquisa qualitativa enfatizam que ainda existe um baixo nível de conhecimento com relação à governança de TIC.

Palavras-chave: Governança de TIC; Tecnologia da Informação e Comunicação; Saúde Pública.

The Governance of Information Technology and Communication in Public Health Organizations: Case Study of the National Cancer Institute

Abstract: The Brazilian public administration is based on the models of the new public management that search for reduce expenses, improve the services offered and give more transparency in the management of public resources. In this scenario, the public sector seeks to improve corporate governance by assisting its organizations in the decision-making process. The focus on ICT governance has intensified as governments have become major consumers of new technologies used in the modernization of public administration. The purpose of this article is to evaluate the implementation of ICT governance in public health organizations through a case study carried out at INCA. The results of qualitative research emphasize that there is still a low level of knowledge regarding ICT governance.

***Keywords**:* ICT Governance; Information and Communication Technology; Public Health.

1. Introdução

A administração pública brasileira é baseada nos modelos da nova gestão pública que buscam eficiência do estado, juntamente com a oferta de serviços públicos de melhor qualidade baseados no controle social e na garantia de direitos individuais (Imasato, Martins & Pieranti, 2011).

De acordo com Denhardt (2008), a nova gestão pública foi elaborada a partir de um seminário acadêmico realizado em 1968 na universidade de Syracuse, Nova York. Seus princípios estão baseados nas seguintes premissas: foco no cidadão; utilização de indicadores de desempenho; transparência; orientação para resultados; flexibilização da gestão e desenvolvimento dos servidores públicos.

Com a nova gestão pública, novas formas de gerenciamento empregadas essencialmente em organizações privadas foram adotadas na administração pública com o objetivo de reduzir os gastos e melhorar os serviços oferecidos à população com o aumento da participação da sociedade nas decisões além de dar mais transparência sobre a gestão dos recursos públicos e do processo de tomada de decisões. A utilização de conceitos de governança pública e de Tecnologia da Informação e Comunicação (TIC) contribuíram para esta mudança (Mendonça et al., 2013).

Neste cenário, o setor público busca melhorar a governança corporativa auxiliando as organizações públicas no processo de tomada de decisões e melhorando assim o funcionamento do governo. A TIC como parte integrante da estratégia organizacional é considerada um fator crítico de sucesso destas ações (Cumps et al., 2009). A TIC, presente em todas as grandes organizações, independente de porte e natureza, tem papel importante no suporte às regras impostas pelas leis e normas governamentais e setoriais, principalmente quando se trata dos processos que suportam as transações comerciais das empresas (Weill & Ross, 2006).

A governança de TIC visa atingir metas relacionadas a custo, crescimento, utilização de ativos e flexibilidade de negócios, além de cumprir os requisitos legais e regulamentares (Bowen et al., 2007). A governança de TIC pode ser definida como um conjunto de políticas, estruturas organizacionais, processos de trabalho e responsabilidades que são estabelecidas pela alta administração para orientar ações e exercer controle sobre o uso e gerenciamento dos recursos de TIC nas organizações (Mendonça et al., 2013).

A governança de TIC tornou-se fundamental para as organizações enfrentarem os desafios apresentados na sociedade atual. No mundo globalizado, as tecnologias da informação proporcionam vantagem competitiva às organizações que empregam novas tecnologias para aumentar a eficácia, agilizando os processos de tomada de decisões e reduzindo custos (Calder & Moir, 2009).

A governança de TIC busca assegurar o alinhamento entre os interesses da área de negócios e as responsabilidades da área de tecnologia por meio do controle e implantação da estratégia de TIC. Esta estratégia deverá ser pensada, planejada e amparada pela alta

direção, gestores de negócios e gestores de TIC, para dar o suporte necessário ao negócio da organização (Weill & Ross, 2006).

De acordo com o Instituto de Governança (2003), a Governança de TIC é parte integrante da governança corporativa e consiste na liderança e gestão das estruturas organizacionais e processos que assegurem que a TIC dê sustentação as estratégias e objetivos da organização. Percebendo a importância da implantação da governança de TIC nas organizações privadas, o setor público logo acompanhou as premissas e adotou estes conceitos.

A governança de TIC no setor público pode ser descrita como uma prestação de serviços com qualidade e transparência do governo para atender as necessidades dos cidadãos em instituições nacionais, municipais ou regionais. As organizações do setor público têm o objetivo de alcançar seu objetivo, melhorando a qualidade de vida para seus cidadãos, diferindo do setor privado, onde o lucro é a principal medida de sucesso. Segundo Wilkin e Campbell (2010) há diferentes obstáculos tais como a complexidade da demanda pública em serviços governamentais e as diferentes conexões políticas da organização. As diferenças entre as políticas, formas de gestão e práticas em organizações públicas também são grandes desafios.

No setor público, o foco sobre governança de TIC se intensificou a medida que os governos se tornaram grandes consumidores de novas tecnologias utilizadas na modernização da administração pública onde questões relacionadas a desempenho, transparência, qualidade do serviço e eficiência tornaram-se o foco central. Nesse cenário, garantir a transparência e a eficiência dos investimentos em TIC tornou-se essencial para fornecer melhores serviços públicos. No entanto, apesar dos esforços dos governos para fomentar o desenvolvimento da governança de TIC, pouco ainda se sabe sobre os efeitos destes investimentos no desempenho das organizações públicas (Ali & Green, 2007).

Atualmente, a tecnologia da informação é considerada um ativo estratégico tanto para organizações públicas como privadas, e as empresas estão cada vez mais dependentes. (Campbell et al., 2009) A TIC desempenha um papel fundamental e transformador nas organizações públicas, não sendo apenas uma ferramenta de gestão, mas fundamentalmente de governança. Assim, a TIC é considerada na área pública como um campo de estudo sendo um subconjunto da governança corporativa que define os direitos de decisão e o quadro de responsabilidades a fim de estimular comportamentos desejáveis no uso de recursos de TIC (Weill & Ross, 2006).

A revolução digital tem aperfeiçoado de forma radical a capacidade e qualidade de captura de grandes volumes de dados sobre os processos e atividades do sistema de saúde. O resultado imediato é a melhoria da prestação de cuidados de saúde assegurando um tratamento adequado e bons resultados para pacientes. De forma geral, a informação também é um recurso fundamental para garantir o controle gerencial, político e democrático do sistema de saúde. Em suma, para melhorar sua governança (Smith et al., 2012).

Campbell et al. (2009) ressaltam que as organizações públicas enfrentam grandes desafios para atender as diferentes prioridades governamentais influenciadas por vários grupos de interesses e intensamente afetadas pela conjuntura econômica. Além disso,

uma organização pública de saúde possui grande diversidade de operações, o que torna esta pesquisa ainda mais relevante.

A área de governança de TIC nas organizações do setor público ainda é limitada em número de pesquisas. Existem estudos sobre governança de TIC em organizações do setor público que procuraram entender políticas introduzidas pelo governo e seus efeitos no setor de TI de uma perspectiva de governança. Esse estudo é mencionado por Hope (2009) que apresentou um estudo feito pela USAID no fornecimento de assistência técnica a países em desenvolvimento e concluiu que a governança de TIC é uma ferramenta que pode fornecer suporte aos governos especialmente em áreas de alta tecnologia.

A maioria dos estudos sobre governança de TIC foram realizados em organizações privadas. Diferentemente das organizações do setor privado que visam o retorno sobre o investimento, o setor público busca atingir metas relacionadas com a oferta de serviços a sociedade e com foco no bem-estar social o que dificulta generalizações. Ao se analisar as organizações públicas no Brasil, há pouca pesquisa sobre o relacionamento da estrutura organizacional com o desempenho da governança de TIC. O objetivo deste artigo é avaliar os impactos da implantação da governança de TIC no desempenho de organizações de saúde pública através de um estudo de caso realizado no Instituto Nacional de Câncer (INCA).

2. Referencial Teórico

A Governança de TIC é influenciada pelas discussões que ocorrem no âmbito da Governança Corporativa e propõe mecanismos para aumentar a segurança, rastreabilidade e a transparência dos processos pertinentes a TIC (ITGI, 2007). Segundo Weill & Ross (2006), uma Governança de TIC eficaz deve tratar três questões: quais decisões devem ser tomadas para garantir a gestão e o uso eficaz da TIC, quem deve tomar estas decisões e como estas decisões serão tomadas e monitoradas?

A Governança de TIC consolidou-se como um tópico importante para que as organizações obtenham alinhamento estratégico, retorno sobre o investimento em TIC e entrega de valor. No setor público, essa realidade não é diferente. A governança de TIC aumentou a sua importância especialmente motivada por demandas para implantar as reformas necessárias para a modernização da gestão pública (Campbell, McDonald & Sethibe, 2009; Bermejo et al., 2014).

O alinhamento estratégico de TIC é um processo contínuo de ajustes que as organizações utilizam para se obter a interligação entre os objetivos e estratégias de negócios com os objetivos e estratégias da área de TIC (Affeldt & Vanti, 2009). A Governança Corporativa dita estratégias corporativas que guiarão as atividades da Governança de TIC. Esta por sua vez será guiada pelos comitês de TIC, que determinarão as ações e ferramentas a serem utilizadas em cada área de abrangência da TIC. No setor privado, há uma definição clara das funções e responsabilidades dos comitês criados com o objetivo de monitorar e direcionar a gestão de recursos de tecnologia de informação da organização. Os gestores executam atividades de acordo com as normas estabelecidas dentro do comitê para dar retorno à instituição (Ali & Green, 2007). No setor público, há diferentes níveis de responsabilidade e grupos de interesses dentro e fora das organizações, nos níveis local, estadual e federal, que aumenta os conflitos (Bermejo et al., 2014).

Campbell et al. (2009) identificam os seguintes desafios ao se implementar governança de TIC no setor público: i) dificuldade em medir o valor público e desempenho; ii) coprodução de arquitetura de TIC entre organizações públicas; iii) dificuldade de conciliar diferentes grupos de interesses. Os pesquisadores ressaltam ainda a presença de múltiplos objetivos, maior regulação e menos incentivos à produtividade.

Weill & Ross (2004) destacam formas eficientes pelas quais as organizações públicas e sem fins lucrativos fazem a gestão da Tecnologia de Informação, promovendo o alinhamento estratégico, criação de valor, estruturas de decisão híbridas e princípios de investimento em infraestrutura de TIC. Ao analisar diferentes mecanismos de governança nas organizações públicas australianas, Ali & Green (2007) destacam a presença de comitês estratégicos de TIC, sistemas de medição de resultados e comunicação corporativa como tendo impactos positivos no desempenho da governança.

Khalfan & Gough (2002) destacam as diferenças da Governança de TIC entre o setor privado e público. As organizações do setor público têm o objetivo de melhorar a qualidade de vida dos seus cidadãos, diferindo do setor privado, onde o lucro é a principal medida de sucesso. Nos países em desenvolvimento, onde as decisões são tomadas pelo governo através de um processo hierárquico é necessário investigar a eficiência de governança e do gerenciamento de investimentos de TIC.

Em todos os estudos relacionados, o envolvimento adequado dos diferentes grupos de interesses é um dos fatores mais importantes para a implantação bem-sucedida da Governança de TIC. As partes interessadas devem ser de todos os níveis hierárquicos e participarem de todos os estágios do processo, caso contrário, a implantação não terá sucesso (Nfuka & Rusu, 2010).

No contexto brasileiro, iniciativas governamentais foram empreendidas para melhorar a governança de TIC no setor público. Iniciativas como a criação da Instrução Normativa 4 e de vários Acórdãos do Tribunal de Contas da União (TCU) que exigem das organizações públicas brasileiras a implementação da governança de TIC com o objetivo de oferecer serviços públicos que agreguem valor ao cidadão através dos seguintes requisitos : i) Planejamento como ferramenta para promover o alinhamento da TIC com os objetivos corporativos; ii) Transparência na aquisição de serviços e recursos de TIC, iii) Segurança da informação, iv) Gerenciamento de pessoal de TIC e v) Monitoramento e controle do desempenho de TI (Bermejo et al., 2014).

A questão considera o desafio do alinhamento entre o desenvolvimento de diretrizes de Governança de TIC e características específicas das organizações do setor público. No contexto brasileiro, esse desafio pode ser evidenciado pelo baixo desempenho das organizações públicas de acordo com uma pesquisa realizada em 2010 pelo Tribunal de Contas da União, aonde a maioria das organizações públicas brasileiras (57%) estava em estágio inicial do desenvolvimento de governança de TIC (Brasil, 2010). Além disso, a maioria dos indicadores de TIC a governança piorou desde a pesquisa no ano de 2007 (Bermejo et al., 2014).

Na revisão teórica foi identificado o modelo elaborado por Van Grembergen & De Haes (2008) que fornece uma perspectiva holística da implementação da governança de TIC em uma organização. Portanto, nesta pesquisa, foi utilizado o framework destes autores que fornece a possibilidade de se avaliar a implementação da governança de TIC através de

três elementos destacados a seguir: Estruturas – envolvem a organização e a localização da função da TI, a existência de papeis e responsabilidades claramente definidos e uma diversidade de comitês de TI e negócios; Processos – referem-se a tomada de decisões estratégicas, ao sistema de informações e de planejamento estratégico, monitoramento e controle; e Mecanismos Relacionais – diz respeito ao relacionamentos, ao alinhamento estratégicos e sua comunicação adequada (Van Grembergen & De Haes, 2008).

3. Metodologia

A pesquisa retratada neste artigo é um estudo teórico empírico com abordagem qualitativa, de natureza descritiva e exploratória (Miles, Hubermam & Saldaña, 2014). Os dados foram coletados por meio de 5 entrevistas. Os dados foram coletados por meio de 5 entrevistas apoiadas em roteiros semiestruturados, sendo 02 gestores que exercem cargos executivos nas áreas de TIC, 2 gestores da área de Planejamento e 1 gestor da coordenação de assistência do Inca. A Tabela 1 apresenta informações relevantes do perfil e da experiência dos entrevistados.

Entrevista	Cargo	Atividades Desempenhadas
Entrevistado 1	Chefe	Serviço de Tecnologia da Informação Tomada de Decisões Técnicas (Hardware/Software). Apresentar Indicadores para Diretoria Treinamento de equipe Gestão de projetos de TIC Gestão pessoal de TIC
Entrevistado 2	Gerente de Governança e Inovação em TIC	Serviço de Tecnologia da Informação Coordenar e acompanhar as implementações da Governança de TIC e PDTI Relatórios de progresso para o chefe de TI.
Entrevistado 3	Chefe de Divisão de Planejamento	Divisão de Planejamento Coordena o Planejamento Estratégico Institucional
Entrevistado 4	Gerente de Projetos	Acompanhamento dos projetos institucionais Acompanhamento do Orçamento
Entrevistado 5	Coordenador de Assistência	Área de Negócio Acompanhamento do Projetos de TIC Relatório de necessidades de TI

Fonte: Elaboração própria.

Tabela 1 – Profissionais entrevistados

O estudo adota o método de estudo de caso único (Yin, 2005), tendo como principais formas de coleta de dados a observação participante, documentos internos e cinco entrevistas semiestruturadas.

A pesquisa qualitativa necessita do uso de várias técnicas de coleta de dados e evidências. Nesta pesquisa os dados foram coletados pela triangulação de: i) Entrevista semiestruturada: através da utilização na pesquisa de campo de entrevista com roteiro estruturado por assuntos, de forma a abordar todo o contexto da pesquisa,

gravadas integralmente em meio digital, com perguntas abertas que foram transcritas e interpretadas para dar suporte ao objetivo proposto da pesquisa; ii) Pesquisa documental: foram levantados registros, documentos, políticas, normas do INCA; iii) Observação direta: durante as visitas de campo foram detectados comportamentos, atitudes e informações visuais dos locais de trabalho.

4. Apresentação e análise dos resultados

Os resultados do estudo de caso, da implementação da GTIC no Inca, são apresentados abaixo, usando a estrutura, processos e mecanismos de relacionamento "como os elementos necessários de um quadro de governança de TIC" de acordo com Van Grembergen e De Haes (2008).

O framework destes autores avalia a implementação da governança de TIC através de três elementos fundamentais: Estrutura: aborda a organização e a função da área de TIC com papeis e responsabilidades claramente definidos e os comitês de TIC e de negócios; Processos: referem-se ao processo de tomada de decisões estratégicas, sistema de informações planejamento estratégico e a estrutura de monitoramento e controle e Mecanismos de Relacionamentos que trata dos relacionamentos, do alinhamento estratégico e da comunicação. Analisando documentos internos, verificamos que o Instituto implantou em 2004, através da portaria 015/2004, formalmente um sistema gerencial colegiado composto pelo:

- Conselho Deliberativo – constituído pelo Diretor Geral, Chefe de Gabinete, 12 Chefias das áreas de negócio, representante do voluntariado, representante da associação de funcionários, além da área de Tecnologia da Informação, responsável por formular e examinar as políticas institucionais, definir linha estratégicas de ação, deliberar sobre o orçamento anual, aprovar e acompanhar o planejamento estratégico. É a instância máxima para tomada de decisão no Instituto.
- Diretoria Executiva – constituída pelo Diretor Geral, Chefe de Gabinete e 8 Chefias das áreas de negócio, responsável por executar as políticas e estratégias aprovadas pelo Conselho Deliberativo.
- Câmaras Técnicas-Políticas – sua composição é multidisciplinar, não hierarquizada, e aberta à participação de funcionários de diversos setores, composta por funcionários das várias áreas de negócio do Instituto, são instâncias de debate institucional e apoio ao Conselho Deliberativo no processo de avaliação e acompanhamento do planejamento estratégico. São 4 as câmaras: a) atenção oncológica; b) informação, educação e comunicação; c) incorporação tecnológica; e d) desenvolvimento institucional. Esta última câmara, analisa e avalia os projetos de Tecnologia da Informação

A Governança de TIC é determinada pela forma como a TIC é organizada e aonde estrategicamente ela é posicionada na estrutura organizacional (Van Grembergen & De Haes 2008). O serviço de TIC é gerido por um CIO Sênior, que se reporta diretamente ao Chefe de Gabinete. Este serviço é composto pelas seguintes áreas: Recursos Tecnológicos, responsável pela gestão de ativos e atendimento ao cliente; Desenvolvimento, responsável pelo desenvolvimento e manutenção dos sistemas de informações; e a

área de Governança e Inovação em TIC recentemente criada com objetivo de alinhar as necessidades de TIC ao negócio. O chefe da TIC apesar de não fazer parte da Diretoria Executiva, se reporta diretamente ao um dos membros desse conselho e sempre que necessário é convidado para suas reuniões.

A Governança de TIC na administração pública federal vem sendo fomentada pela Secretaria de Tecnologia da Informação e Comunicação (SETIC) do Ministério do Planejamento, através de normas, guias e instruções normativas e principalmente pelos órgãos de controle, a exemplo do Tribunal de Contas da União (TCU), que vem nos últimos anos promovendo levantamentos e realizando auditorias a vista de avaliar a situação da governança nas instituições públicas federais.

A estruturação da Governança de TIC no Inca se deu principalmente pela necessidade de estar *compliance* com as recomendações desses órgãos reguladores. Em documento interno o chefe do Serviço de Tecnologia da Informação solicita ao CEO a criação de uma nova estrutura para atender as exigências dos órgãos de controle:

"A Governança de TIC está se tornando cada vez mais importante na nossa sociedade complexa e baseada em conhecimento. Desta forma, existe a necessidade da criação de uma estrutura adequada que direcione a TI aos objetivos institucionais. A adoção desta estrutura é estimulada pela SLTI e também ressaltada pelos órgãos de controle, a exemplo do TCU, o qual recomenda, à SLTI, a normatização da obrigatoriedade de que os entes sob sua jurisdição, na qual o Inca se enquadra, estabeleçam comitês de tecnologia da informação".

Em 2013, através da portaria Inca no. 600 de 30 de julho de 2013 foi criado o Comitê de Tecnologia da Informação, composto por 3 representantes da Divisão de Tecnologia da Informação e 7 representantes das diversas áreas de negócio da instituição. Neste período a Divisão de Tecnologia da Informação, hoje Serviço de Tecnologia da Informação, era subordinada a Coordenação de Administração (CFO) dentro da estrutura organizacional da empresa. Em levantamento realizado pelo TCU em 2012, o Inca, teve o índice de maturidade governança (iGovTI) avaliado em 0,41, considerado intermediário e posteriormente em 2014, 0,43 intermediário.

Em 2016 o Comitê de Tecnologia da Informação, foi reestruturado passando a se chamar Comitê Estratégico e Gestor de Tecnologia da Informação e Comunicação (CEGTIC), vinculado a Direção Geral do Instituto, com natureza deliberativa, de caráter permanente, do tipo estratégico/executivo e constituído com a finalidade de direcionar, monitorar e avaliar o uso estratégico da TIC (portaria Inca no. 417 de 02/05/2016). Este comitê é composto por 9 Coordenadores e Chefes das áreas de negócio, que também fazem parte da Diretoria Executiva, mais 2 representantes da área de TIC. O comitê é responsável pela aprovação do Plano Diretor de Informática, do orçamento necessário para execução, priorização, monitoração dos projetos de TIC e seus resultados.

Nos documentos de gestão do INCA, foram levantados a visão, missão e os objetivos estratégicos da organização. A estratégia da organização está fortemente baseada na introdução das tecnologias inovadoras em conectividade, compartilhamento de informações e prestação de serviços através de soluções baseadas na Intranet e em

plataformas web. Para cumprir estes objetivos, foram realizadas mudanças na estrutura organizacional.

Através das entrevistas e análise dos documentos do INCA observou-se que a estratégia da instituição tem foco na automação dos seus processos administrativos e na produção, disseminação e aplicação do conhecimento, dando ênfase na introdução das tecnologias mais recentes de conectividade e de distribuição de informações para prestação de serviços de saúde aos seus pacientes. A instituição passou recentemente por um processo de reestruturação de sua organização estratégica. Os principais atores desta reestruturação foram o Serviço de TI, a Divisão de Planejamento, a Coordenação de Administração Geral e a Coordenação de Assistência. O objetivo principal foi garantir a aplicação e o uso de TIC para obter eficácia e eficiência dos processos relacionados à gestão de suprimentos das unidades hospitalares do instituto e as atividades clínicas da assistência aos pacientes.

Nas entrevistas e documentos analisados fica claro o papel estratégico do serviço de TI de: desenvolver e melhorar a infraestrutura técnica da organização; constituir e gerenciar regras e procedimentos operacionais para utilização dos recursos de tecnologia da Informação; atender as coordenações da organização no desenvolvimento e manutenção de sistemas de informações; desenvolver programas de treinamento dos usuários; estabelecer políticas de segurança de informação; e liderar projetos de aquisição e implementação de novas soluções de Tecnologia de Informação e Comunicação.

A maioria dos entrevistados destacou que o alinhamento, comunicação e relacionamento entre o serviço de TIC e os demais setores da organização são aspectos importantes que devem ser considerados na implementação da governança de TIC. Na verdade, o

Práticas de Governança de TIC	
Serviço de TIC	O serviço de TIC é encabeçado por um chefe de TI Sênior que reporta a chefia de gabinete do diretor executivo (CEO) do INCA. O chefe de TI não se reporta diretamente ao CEO, mas possui as mesmas responsabilidades que outros chefes de departamentos. O chefe de TI tem papel e responsabilidades claras nos planos de organização interna, e a área participa na elaboração do relatório de acompanhamento anual a ser enviado aos órgãos de controle
Comitê Gestor e Estratégico de TIC	O Comitê Gestor e Estratégico de TIC foi estabelecido para discutir questões relativas ao planejamento estratégico de TIC, ao PDTI, a Finanças, Recursos Humanos e Segurança da Informação. O comitê foi estabelecido como iniciativa da área de TI e está formalmente estabelecido com deveres e responsabilidades formais. No Comitê Estratégico e Gestor de TIC, o chefe de serviço de TI é seu presidente formalmente instituído e seu papel é avaliar a estratégia atual e sugerir melhorias. “Comitê l relacionado a serviços de apoio discute as questões relacionadas à TI, esta comissão liderada pelo diretor da instituição pública estudada
Estrutura de TIC	O serviço de TIC é composto por 3 gerentes das áreas de Recursos Tecnológicos, Desenvolvimento de Sistemas e Governança e Inovação em TIC. Os gerentes acompanham todos os projetos e planos para o serviço; Comitês e grupos são formados para acompanhar os projetos, mas devem informar o chefe do serviço e gerentes no final das discussões, e, na maioria dos casos, os gerentes participam do trabalho dos comitês e dos grupos de trabalho.

Fonte: Elaboração Própria

Tabela 2 – Práticas de Governança de TIC observadas na estrutura do INCA

alinhamento entre negócios e TI e uma forte relação entre ambos é necessária para alcançar os objetivos estratégicos da instituição. Uma barreira importante levantada nas entrevistas é a falta de conhecimento de ferramentas de gestão por parte dos profissionais de TI e demais gestores na área de saúde que impede a organização de perceber o valor e o retorno dos investimentos em tecnologia. Apesar dos esforços contínuos, encontrar equipes de TI qualificadas e mantê-las é um grande desafio da área pública. Assim, a formação de uma equipe de TI qualificada e sua manutenção são aspectos importantes que devem ser abordados. Na tabela 2 são destacadas as práticas de Governança de TIC que foram observadas na estrutura do INCA através da realização de entrevistas e análise de documentos usando o framework de Van Grembergen & De Haes (2008).

4.2. Processos

O Serviço de Tecnologia da Informação e Comunicação do Ministério do Planejamento (SETIC) normatiza os processos de governança de TIC que devem ser seguidos pelos órgãos da Administração Pública Federal, entre eles o INCA. Verificamos em documentação interna que o Instituto segue a Instrução Normativa no. 04 de 2014 (IN4) nos projetos de TIC. A IN4 define os processos que devem ser executados antes, durante e depois de uma contratação de bens e serviços de TIC.

O Comitê Estratégico e Gestor de TIC é responsável no Inca, pela indicação dos membros da Equipe de Elaboração do Plano Diretor/Estratégico de TI (PDTI), que é bianual. Essa equipe é composta por um representante de cada área de negócio do Instituto. Após todas as etapas de elaboração, o PDTI é submetido ao Comitê para aprovação e posterior acompanhamento da execução.

Como o orçamento no Inca é único, os projetos de TI aprovados dependem da existência de recursos para serem executados. Os projetos aprovados no PDTI seguem a IN4, que envolve 3 fases: planejamento da contração; seleção de fornecedores; e gestão de contrato. O COBIT está inserido dentro da organização, que utiliza o "Levantamento do Perfil de Governança de TI" do Tribunal de Contas da União, que é realizado a cada dois anos, para verificar o nível de maturidade dos órgãos da APF. Nesse levantamento são acompanhados vários processos do COBIT. O Inca em sua última avaliação teve aferido o nível de maturidade como sendo intermediário. A área de Governança e Inovação em TIC foi criada em 2016 com o objetivo de elevar o nível de maturidade da Governança de TI para avançada.

O ITIL é utilizado pela STI, porém nem todos seus processos foram implementados. Nas entrevistas foi identificado que existem restrições devido a falta de recursos humanos capacitados. Entrevistado da STI: "Temos dificuldade de implantar o ITIL completamente, principalmente por termos poucos funcionários capacitados e que conheçam todos os seus processos".

4.3. Mecanismos de Relacionamento

Há alguns anos o Inca vem trabalhando fortemente na comunicação, principalmente com seu público interno. Nas entrevistas foi reportado que apesar da cultura do órgão ser de uma comunicação pessoal, vários avanços na comunicação institucional foram observados. Fala do CIO: "O Informe Inca ajuda a divulgar os projetos de TIC mais relevantes da Instituição, além do Informe, temos o Postmaster, correio eletrônico

direcionado a todos os funcionários, além da Intranet que apresenta diariamente a notícias de interesse institucional".

A área de Comunicação Social do Inca faz parte do Comitê Estratégico e Gestor da TIC, o que facilita organizar as ações de comunicação das ações/projetos de TIC para área de negócio. Uma das maiores dificuldades encontradas pela área da TIC é traduzir seu conhecimento técnico em uma linguagem que possa ser entendida pela área de negócio. O CIO do Inca enfatizou que o Serviço Tecnologia da Informação tem um bom relacionamento com todas as áreas de negócio do Inca e que apesar de não fazer parte da Diretoria Executiva é frequentemente convidado para participar de suas reuniões. A Direção Geral do Inca sabe da importância da comunicação relacionada a governança de TIC chegar a todos os níveis da organização. A área de Governança e Inovação em TIC vem fazendo um trabalho em conjunto com a área de Comunicação Social do Inca, visando melhor essa comunicação, para que a mesma chegue a todos os níveis da instituição. Como a Governança de TIC se iniciou de dentro da área de TIC esse processo se encontra em evolução.

5. Conclusões

A análise do cenário através do framework de Van Grembergen e De Haes identificaram que os maiores desafios enfrentados pelo departamento de TI atualmente estão relacionados a restrições de orçamento e a fragmentação do processo de tomada de decisões. Além disso, a análise das práticas da governança de TI na organização de saúde pública estudada confirmou que apesar dos avanços identificados, a governança de TI não está implementada integralmente.

A implementação da governança de TI é um esforço colaborativo entre os diferentes níveis de gerenciamento e grupos de funcionários da organização. O princípio do COBIT, "atender às necessidades das partes interessadas" e os princípios da ISO 38500, "responsabilidade" e "estratégia", não podem ser praticados sem o envolvimento das partes interessadas. Além disso, em todos os estudos relacionados, o envolvimento efetivo de todos os envolvidos é um dos fatores mais importantes para uma implementação bem-sucedida.

O estudo demonstra a complexidade da estrutura organizacional de uma instituição pública de saúde como o INCA e o grande número de processos a serem controlados pelos poucos mecanismos de controle atualmente disponíveis. Além disso, estas instituições sofrem forte influência por diferentes grupos de interesse. Isso mostra uma grande diferença entre organizações públicas e privadas onde fatores externos podem influenciar fortemente a alta administração.

Os resultados da pesquisa enfatizam que ainda existe um baixo nível de conhecimento com relação à governança de TI nas organizações governamentais. As estruturas, processos e mecanismos que promovem a atribuição de responsabilidades na gestão de projetos e de contratos de TI devem contribuir para uma efetiva implementação da governança nas organizações públicas. Neste sentido, a criação de uma área especifica de governança de TI foi fundamental para o INCA. A pesquisa ressalta que o desempenho do comitê de TI e o apoio da de direção do INCA tiveram uma influência positiva e afetaram consideravelmente o desempenho da gestão de recursos e investimentos de TI. Outros fatores críticos de sucesso estão relacionados à comunicação entre a TI e demais

setores organizacionais, o desenvolvimento de habilidades necessárias para a tomada de decisão dos gestores de TI e a criação e execução de processos de controle gerencial sob responsabilidade da área de governança de TI.

Este trabalho oferece uma direção importante quando destaca que as organizações públicas de saúde começam a considerar a TI como peça chave para a governança corporativa. Ao contemplar esse conceito, as iniciativas de implementação da governança de TI desempenham um papel importante para alcançar uma melhor gestão dos recursos de tecnologia da informação para o atendimento das demandas da sociedade relacionadas a saúde pública que é responsabilidade do Estado.

Os resultados deste trabalho poderão ser utilizados como insumos no processo de tomada de decisão dos líderes das organizações de saúde pública em questões relativas à implementação dos mecanismos de governança analisados. Eles também podem apoiar os gestores de TI e da equipe de governança na identificação de necessidades de treinamento e no planejamento de ações de comunicação e educação, com base nas conclusões apresentadas.

Referências

Affeldt, F. S.; Vanti, A. A. (2009). Alinhamento estratégico de tecnologia da informação: análise e modelos de propostas para pesquisas futuras. *Revista de Gestão da Tecnologia e Sistemas de informação*, v.6, n.2, p. 203-226.

Ali, S.; Green, P. (2007). IT governance mechanisms in public sector organizations: An Australian context. *Journal of Global Information Management*, 15(4), p. 41–63.

Bermejo, P. H. S.; Tonelli, A. O. & Zambalde, A. L. (2014). Developing IT governance in Brazilian public organizations. *International Business Research*,7(3), 101–114.

Board Briefing on IT Governance, 2nd Edition.(2003). Illinois: ITGI, 2003, 66 p. Disponível em: www.itgi.org.

Bowen, P. L.; Cheung, M. D.; & Rohde, F. H. (2007). Enhancing IT governance practices: A model and case study of an organization's efforts. *International Journal of Accounting Information Systems*, 8(3), 191–221.

Calder, A.; Moir, S. (2009). IT governance, implementing frameworks and standardsfor the corporate governance of IT. IT Governance Publishing.

Campbell, J.; McDonald, C. & Sethibe, T. (2009). Public and private sector IT Governance: Identifying contextual differences. *Australian Journal of Information Systems*, 16(2), 5–18.

COBIT 4.1. (2007). Rolling Meadows, IL, 212 p.

Cumps, B.; Martens, D.; De Backer, M.; Haesen, R., Viane, S.; Dedene, G., et al. (2009). Inferring comprehensible business/ICT alignment rules. *Information and Management*, 46(2), 116–124.

Creswell, J. W. (2010). Projeto de pesquisa: Métodos qualitativo, quantitativo e misto. 3ª ed. Porto Alegre: Ed. Artmed.

Denhardt, R. B. (2008). Theories of public organization (5th ed.). Belmont, CA: Thomson/Wadsworth.

Fernandes, A. A.; Abreu, Vladimir F. (2012). Implantando a Governança de TI: da estratégia à gestão dos processos e serviços. 3ª ed. Rio de Janeiro: Ed. Brasport.

Ferreira, E. F. C. (2011). Governança corporativa na saúde suplementar: Estudo de caso em operadora de plano de saúde. Dissertação (Mestrado em Administração de Empresas) – Universidade de Fortaleza – UNIFOR, CMA, Fortaleza, 93 p.

Hope, K.R.S. (2009). "Capacity development for good governance in developing countries: some lessons from the field", *Development in Practice*, 19 (1):79–86, 2009.

Imasato, T.; Martins P. E. M., & PIERANTI O. P. (2011). Administrative reforms and global managerialism: A critical analysis of three Brazilian state reforms. *Canadian Journal of Administrative Sciences*, 28(2), 174–187.

IT Governance Institute. "Board briefing on IT governance", 2nd ed., 2003.

Instituto Brasileiro de Governança Corporativa (IBCG). (Acesso em setembro de 2016). Disponível em http://www.ibgc.org.br/.

Khalfan, A.; Gough, T.G. (2002). Comparative analysis between the public and private sectors on the IS/IT outsourcing practices in a developing country: a field study, *Logistics Information Management*, vol. 15, no.3, pp. 212-223.

Mendonça, C. M. C.; Guerra L.C.B. & Neto M. V.S. (2013). Governança de tecnologia da informação: um estudo do processo decisório em organizações públicas e privadas. *Revista. Administração Pública*, Rio de Janeiro, v. 47, n. 2.

Miles, M.B.; Huberman. M.A. & Saldaña, J. (2014). Qualitative data analysis. A methods sourcebook. London: SAGE Publications.

Nfuka, E. N.; Rusu, L. (2010). Critical success factors for effective IT governance inthe public sector organizations in a developing country: the case of Tanzania.AMCIS 2010 proceedings, 1–15.

Smith P.C.; Anell A.; Busse R.; Crivelli L.; Healy J.; Lindahl A.K.; Westert G; Kene T. (2012). Leadership and governance in seven developed health systems. *Health Policy*. p. 37-49.

Van Grembergen, W.; De Haes, S.(2008). "Implementating Information Technology Governance: Models, Practices, and Cases", IGI Global, p.25.

Weill, P.; Ross, J W. (2006). Governança de TI - Tecnologia da Informação, São Paulo: Ed. Makron Books.

Wilkin, C. L.; Campbell, J. (2010). "Corporate Governance of IT: A Case Study in an Australian Government Department", In *Proceedings of the Pacific Asia Conference of Information Systems*.

Yin, R. K. (2015). Estudo de caso: planejamento e métodos. Porto Alegre: Bookm.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 20/09/2018
Aceitação/Acceptance: 21/11/2018

Modelo Ontológico de los Ciberdelitos: Caso de estudio Colombia

Jhon Marin[1], Yuri Nieto[2], Freddy Huertas[3], Carlos Montenegro[4]

Jhon.marin@cun.edu.co, yuri_nieto@cun.edu.co, fa.huertas@uniandes.edu.co, cemontenegrom@udistrital.edu.co

[1] Perito en Informática Forense Ciber GAULA Policía Nacional, Bogotá Colombia.

[2] Docente Corporación Unificada Nacional de Educación Superior CUN, Bogotá , Colombia

[3] Estudiante de Maestría de Ingeniería de Sistemas y Computación de Universidad de los Andes, Bogotá, Colombia

[4] Decano Facultad de Ingeniería Universidad Distrital Francisco José de Caldas, Bogotá, Colombia

Pages: 244–257

Resumen: Este articulo presenta un modelo Ontológico de los ciberdelitos tomando como caso de estudio Colombia. Se tienen en cuenta diferentes aspectos como son clasificación, jurisprudencia, nivel de impacto y se determinan las nuevas modalidades que se están presentando. A través del modelo ontológico se coleccionan conceptos clave y su interrelación colectiva con el fin de apoyar la toma decisiones en el ámbito de la ciberseguridad a través de la tipificación de los ciberdelitos. Este documento busca brindar herramientas que permitan reconocer el ciberdelito en Colombia, de igual forma, establecer patrones de conducta que permitan prevenir ser víctima de estos delitos y apoyar la toma decisiones respecto a la seguridad en la web.

Palabras-clave: Ciberdelitos; hackers; malware; ontología.

Ontological Model of Cybercrimes: Case study Colombia

Abstract: An Ontological model of cybercrimes is presented through the paper taking Colombia as a case study. Different aspects are taken into account such as classification, jurisprudence, level of impact and the new modalities that are currently in used. Through the ontological model, key concepts and their collective interrelation are collected in order to support decision-making on the cybersecurity field supported by typifying cybercrimes. This document seeks to provide tools that allow recognizing cybercrime in Colombia, in the same way, establish patterns of behavior that allow to prevent being a victim of these crimes and support decision-making in regard of web security.

Keywords*:* Cybercrimes; hackers; malware; Ontology.

1. Introducción

El Cibercrimen y su constante evolución, ha promovido que los delincuentes que anteriormente actuaban de manera aislada, sin coordinación, con un alcance local y corriendo riesgos al realizar sus actividades de forma presencial, se constituyan en empresas criminales, con roles específicos dentro de una organización, adicionalmente cuentan con recursos económicos y de software, así como alcances ilimitados y total anonimato.

El mercado ilegal de datos y el ciber crimen como servicio han facilitado su auge y transformación, así como, el fácil acceso al mercado ilegal de Malware (programas maliciosos), las monedas virtuales y la dificultad de rastreo de actividades ilícitas en la internet profunda o Dark Net.

Aunque desde 1983 comenzó a hablarse de la criminalidad informática y en 1992 se acuñó el término Cibercrimen (Fernando Miró Llinares. 2012), el fenómeno de la criminalidad relacionada con el uso de las Tecnologías de la Información y la Comunicación sigue siendo novedoso y por ello, parcialmente incomprendido por la sociedad en general y en particular, por las instituciones encargadas de la prevención de esta amenaza.

Según Microsoft, Colombia es el tercer país en Latinoamérica con mayor índice de Cibercrimen. De acuerdo con cifras avaladas por el IDC Colombia (empresa dedicada al análisis del mercado de Tecnología Informática y Telecomunicaciones), los ciberataques han aumentado entre el 30% y el 40% en América Latina, una de las zonas con mayor actividad en el mundo. Durante 2015, se registraron más de 20 violaciones a la seguridad por segundo en la región, lo cual equivale a 400 mil vulneraciones a causa de virus. Por su parte, en el mismo año, Colombia ha sido el tercer país más afectado por el Cibercrimen con 5 millones de ataques informáticos, seguido de Brasil y México con 27 millones y 16 millones de incidentes de este tipo respectivamente **(**http://www.microsoftcolombia.com/dia-de-cibercrimen-colombia**).**

La ontología es una base importante para el conocimiento que representa el mundo real (Yang 2009). La filosofía acuño el término de ontología y ha ido creciendo en el campo de la informática y el sistema de información (Koutero, Fujita y Sugawara 2010) (Maedche y Staab 2002) (Busagala et al., 2008). Hoy en día, el término ontología se usa en numerosos campos de Ingeniería, la ontología que se presenta en el artículo busca de forma clara, clasificar los ciberdelitos según su tipo de ataque y objetivo, definiendo su jurisprudencia en Colombia, su impacto y las nuevas modalidades que se presentan. Brindar a la sociedad un conocimiento claro de las malas conductas en la web y como estas son tipificadas y reconocidas por el gobierno como un delito, les brindará así mismo el apoyo necesario para tomar decisiones referentes a su seguridad en la web.

El presente artículo tiene como objetivo definir un modelo Ontológico del ciberdelito tomando como caso de estudio a Colombia. En la segunda sección se define el modelo ontológico del Cibercrimen tomando como caso de estudio Colombia. Se muestra en la tercera sección la definición del ciber crimen. En la cuarta sección ciber guerra y ciber defensa, por último, en la quinta sección se presentan las conclusiones y trabajo a futuro.

2. Definición ontológica del ciber crimen

En la figura 1 se plantea el Modelo Ontológico del Ciber crimen, definiendo los tipos de ataques que se pueden realizar y su tipicidad en el código Penal.

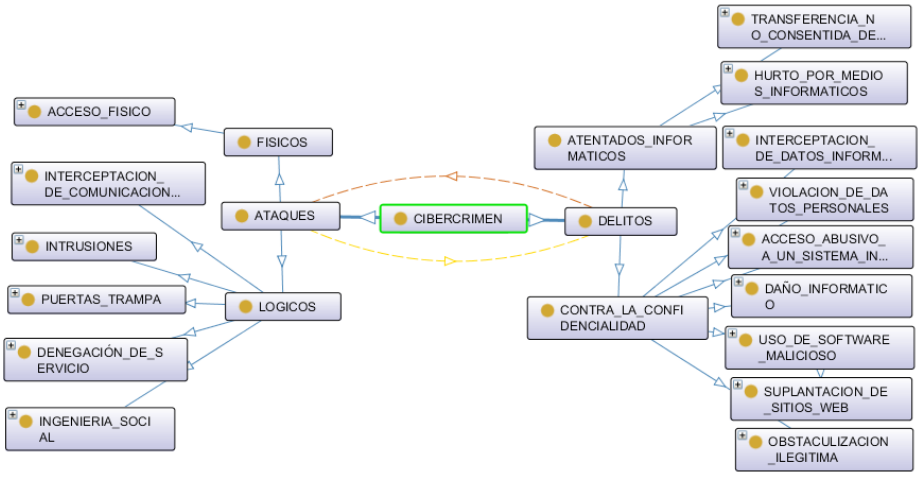


Figura 1 – Modelo Ontológico del ciber crimen

Para llevar a cabo la definición ontológica, se tomó como referente las leyes colombianas y en específico el código penal colombiano. El Cibercrimen se divide en ciberataques. Estos son la forma como se realiza la conducta y en cómo se tipifican estas conductas en el código penal colombiano.

2.1. Ciber ataques

Los sistemas informáticos están compuestos por cierta cantidad de componentes que permiten su correcto funcionamiento, entre ellos, electricidad, hardware, sistema operativo, aplicaciones, datos, red, usuarios, entre otros. Los ataques se pueden realizar en cada eslabón de esta cadena, aprovechando las vulnerabilidades que el atacante explote, es así, que se pueden presentar los siguientes ataques:

2.1.1. Acceso físico

El atacante tiene acceso a las instalaciones e incluso a los equipos:

- Interrupción del suministro eléctrico, su objetivo es generar daño físico a los equipos, causando fallas en el sistema eléctrico.
- Apagado manual del equipo, se realiza de forma presencial y busca la perdida de información.

- Vandalismo, a través de ataques físico causar daños a la infraestructura.
- Apertura de la carcasa del equipo y robo del disco duro.
- Monitoreo del tráfico de red, con el fin de detectar vulnerabilidades que puedan ser explotadas en futuros ataques.

2.1.2. Interceptación de comunicaciones

El atacante busca recopilar la información de acceso y de esta forma tener acceso restringido al sistema.

- Secuestro de sesión, robo de perfiles o de identidad del usuario.
- Falsificación de identidad, a través de phishing, simular la identidad de un usuario o de una página.
- Re direccionamiento o alteración de mensajes, se usa para realizar ataques tipo troyano, simulando ser un correo de confianza.

2.1.3. Denegaciones de servicio

Es un ataque a un sistema de computadoras o red que causa que un servicio o recurso sea inaccesible a los usuarios legítimos. Su objetivo es interrumpir el funcionamiento normal de un servicio.

- Explotación de las debilidades del protocolo TCP/IP.
- Explotación de las vulnerabilidades del software del servidor.

2.1.4. Intrusiones

Es un ataque el cual busca dar acceso a personal no autorizado a un sistema informático, otorgándole privilegios de administrador o incluso acceder a información reservada.

- Análisis de puertos, Una técnica que los hackers usan para encontrar las debilidades de un equipo o de una red. Si bien esta técnica no es un ataque en realidad, los hackers la usan para detectar qué puertos están abiertos en un equipo. De acuerdo con la información de los puertos abiertos, puede obtenerse acceso no autorizado.
- Elevación de privilegios: este tipo de ataque consiste en aprovechar una vulnerabilidad en una aplicación al enviar una solicitud específica (no planeada por su diseñador). En ciertos casos, esto genera comportamientos atípicos que permiten acceder al sistema con derechos de aplicación. Los ataques de desbordamiento de la memoria intermedia (búfer) usan este principio.
- Ataques malintencionados (virus, gusanos, troyanos).

2.1.5. Ingeniería Social

En la mayoría de los casos, el eslabón más débil es el mismo usuario. Muchas veces es él quien, por ignorancia o a causa de un engaño, genera una vulnerabilidad en el sistema al brindar información (la contraseña, por ejemplo) al pirata informático o al abrir un archivo adjunto.

2.1.6. Puertas trampa

Son puertas traseras ocultas en un programa de software que brindan acceso a su diseñador en todo momento.

Es por ello que los errores de programación de los programas son corregidos con bastante rapidez por su diseñador apenas se publica la vulnerabilidad.

2.2. Delitos informáticos

En Colombia se tipificaron en el Código Penal los delitos informáticos, los cuales se dividen en: delitos contra la confidencialidad y atentados informáticos.

2.2.1. Delitos contra la confidencialidad

Son aquellos atentados contra la confidencialidad personal, la integridad y la disponibilidad de los datos y de los sistemas informáticos.

- Acceso abusivo a un sistema informático: se refiere al que sin autorización, acceda a un sistema informático protegido en contra de la voluntad de quien tenga el legítimo derecho a excluirlo.
- Obstaculización ilegítima de sistema informático o red de telecomunicación: incurre en esta conducta el que, sin estar facultado para ello, impida u obstaculice el funcionamiento o el acceso normal a un sistema informático, a los datos informáticos o a una red de telecomunicaciones.
- Interceptación de datos informáticos: se refiere al que, sin orden judicial previa intercepte datos informáticos en su origen, destino o en el interior de un sistema informático.
- Daño Informático: lo realiza quien, sin estar facultado para ello, destruya, dañe, borre, deteriore, altere o suprima datos informáticos, o un sistema de tratamiento de información o sus partes o componentes lógicos.
- Uso de software malicioso: se refiere al que, produzca, trafique, adquiera, distribuya, venda, envíe, introduzca o extraiga del territorio nacional software malicioso u otros programas de computación de efectos dañinos.
- Violación de datos personales: lo realiza quien con provecho propio o de un tercero, obtenga, compile, sustraiga, ofrezca, venda, intercambie, envíe, compre, intercepte, divulgue, modifique o emplee códigos personales, datos personales contenidos en ficheros, archivos, bases de datos o medios semejantes.
- Suplantación de sitios web para capturar datos personales: se refiere a la persona que con objeto ilícito diseñe, desarrolle, trafique, venda, ejecute, programe o envíe páginas electrónicas, enlaces o ventanas emergentes.

2.2.2. Atentados informáticos

Son aquellos ataques que no transgreden la información personal, pero que producen un daño o un detrimento económico a través del entorno digital.

- Hurto por medios informáticos y semejantes: hace referencia al que, superando medidas de seguridad informáticas, hurte manipulando un sistema informático,

una red de sistema electrónico, telemático u otro medio semejante, o suplantando a un usuario ante los sistemas de autenticación y de autorización establecidos.

- Transferencia no consentida de activos: la comete la persona que, con ánimo de lucro y valiéndose de alguna manipulación informática o artificio semejante, consiga la transferencia no consentida de cualquier activo en perjuicio de un tercero.

3. Ciber crimen

Pocos años atrás, los programas maliciosos se limitaban al "ciber-vandalismo", una forma de expresión antisocial que irrumpía en los PCs causando diversos daños. Pocos de ellos, no obstante, estaban diseñados con este fin, e inevitablemente producían daños colaterales en los archivos o dejaban el equipo inservible. La mayoría de las amenazas en esta época consistían en virus y gusanos.

Hoy en día, por el contrario, la amenaza más grave proviene del llamado Cibercrimen. (Miro 2015) Los criminales se sirven del anonimato que la red otorga para, mediante códigos maliciosos, acceder a los equipos y robar dinero o datos confidenciales como contraseñas, logins, códigos PIN, etc.

Las amenazas del Cibercrimen incluyen: virus, gusanos, troyanos, ataques de hackers, phishing demás. Estas amenazas son cada día más sofisticadas y su número crece exponencialmente.

Cuando se hace referencia a cibercrimen se entiende que se trata de delitos cometidos a través de internet por medio del uso de un computador o mecanismo análogo (por ejemplo: smartphone, pendrive, tablet, etc.). También se han empleado términos tales como computer crime, computer related crime, digital/electronic/virtual, IT, high tech-crime, delitos informáticos, entre otros (Miró 2011).

Todos los usuarios de la web, de manera directa o indirecta están expuestos a ser víctimas de fraude, extorsión y espionaje informático. Este delito, mejor conocido como Cibercrimen, no sólo pretende conseguir un beneficio principalmente económico, sino que también abarca el dominio de Internet como ataques con fines políticos, programas informáticos maliciosos, etc.

Cuando se hace referencia a cibercrimen se entiende que se trata de delitos cometidos a través de internet por medio del uso de un computador o mecanismo análogo (por ejemplo: smartphone, pendrive, tablet, etc.). También se han empleado términos tales como computer crime, computer related crime, digital/electronic/virtual, IT, high tech-crime, delitos informáticos, entre otros.

3.1. Ciber crimen en Colombia

Según el Ministerio de Tecnologías de la Información y las Comunicaciones de Colombia (MINTIC) 8 de cada 10 colombianos usan actualmente internet, el Ministerio TIC e Ipsos Napoleón Franco presentaron el estudio de Consumo Digital, donde se revela cómo, dónde, cuándo y cada cuánto los colombianos usan Internet (http://www.mintic.gov.co/portal/604/w3-article-1629.html, s.f.).

El estudio reveló que el 80% de los encuestados usa Internet y que el mayor incremento del uso de la red se dio en los estratos 1 y 2, con un crecimiento del 17% en comparación al uso que le daban en 2010. También se observa que el 54% de los colombianos que usan Internet, lo hacen todos los días y pasan en promedio 2,6 horas navegando.

Si bien es cierto que Colombia es pionera en América Latina en temas de conectividad y acceso a internet, esto ha provocado que más personas sean propensas a ser víctimas del ciber crimen, debido a que aunque se logró mayor conectividad y mayor acceso a las nuevas tecnologías; no se realizó el mismo esfuerzo para capacitar a los ciudadanos sobre el uso adecuado de internet.

3.1.1. Estrategias del gobierno colombiano

El 14 de julio de 2011, se realizó en Colombia el CONPES 3701 (Consejo Nacional de Política Económica y Social) LINEAMIENTOS DE POLÍTICA PARA CIBERSEGURIDAD Y CIBERDEFENSA, el cual buscaba generar lineamientos de política en Ciberseguridad y Ciberdefensa orientados a desarrollar una estrategia nacional que contrarreste el incremento de las amenazas informáticas. Adicionalmente, recoge los antecedentes nacionales e internacionales, así como la normatividad del país en torno al tema.

La problemática central se fundamentó en que la capacidad actual del Estado para enfrentar las amenazas cibernéticas presenta debilidades y no existe una estrategia nacional al respecto. A partir de ello se establecen las causas y efectos que permitirán desarrollar políticas de prevención y control, ante el incremento de amenazas informáticas. Para la aplicabilidad de la estrategia se definen recomendaciones específicas a desarrollar por entidades involucradas directa e indirectamente en esta materia. Así lo ha entendido el Gobierno Nacional al incluir este tema en el Plan Nacional de Desarrollo 2010-2014 "Prosperidad para Todos", como parte del Plan Vive Digital.

Dicho CONPES se realizó teniendo en cuenta los ciberataques perpetuados en varias partes del mundo y los ocurridos en Colombia, debido a esto se creó la Ley 1273 de 2009, con la cual se creó un nuevo bien jurídico tutelado "de la protección de la información y de los datos".

El gobierno Colombiano no podía ser indiferente ante los cambios que se estaban presentando en el mundo, donde aparecieron grupos hacktivistas como anonimous que demostraron tener la capacidad de burlar y bloquear cualquier sistema. Es por ello que, al no existir legislación vigente para investigar estos delitos, se creó esta Ley mediante la cual se tipificaron los ciberdelitos en el código penal colombiano.

3.1.2. Impacto del ciber crimen en Colombia

Según la compañía de ciber seguridad Digiware, Colombia participó con el 8,05% del total de los delitos informáticos de América Latina, lo que equivale a pérdidas por más de US$6.179 millones. Con estas cifras, Colombia es quinto en la clasificación latinoamericana en materia de ataques informáticos.

Digiware mostró que Brasil es primero y representa el 25,13% de los ataques a nivel regional, seguido por México con 15,53%, Venezuela 11,91% y Argentina con 9,63%.

Diariamente, en Colombia se producen en promedio 542.465 ataques informáticos. Así se distribuyen los ataques por sectores económicos (sectores-mas-afectados-por-cibercrimen-en-colombia/250321):

El sector financiero: 214.600 ataques por día (39,56%).

Telecomunicaciones: 138.329 ataques por día (25,5%).

El sector Gobierno: 83.756 ataques por día (15,44%)

Sector energético: 19.583 ataques por día (3,61%)

Industria: 51.263 ataques por día (9,45%).

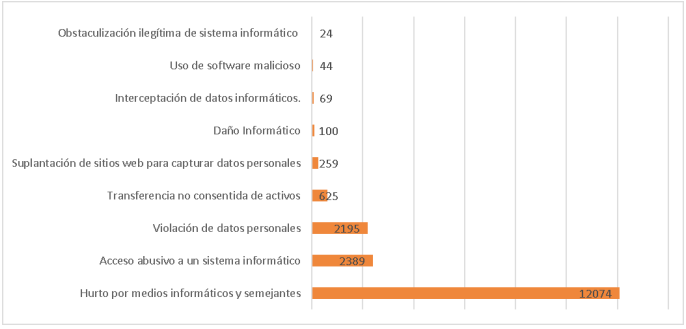


Figura 2 – Denuncias de delitos informáticos en Colombia

En la figura 2 se observa que entre el año 2014 y 2017, se recibieron 17.779 denuncias por violación a la ley 1273 de 2009, dando un panorama de los delitos que más se denuncian en el país. En cuanto a las tipologías criminales denunciadas ante la Policía Nacional en

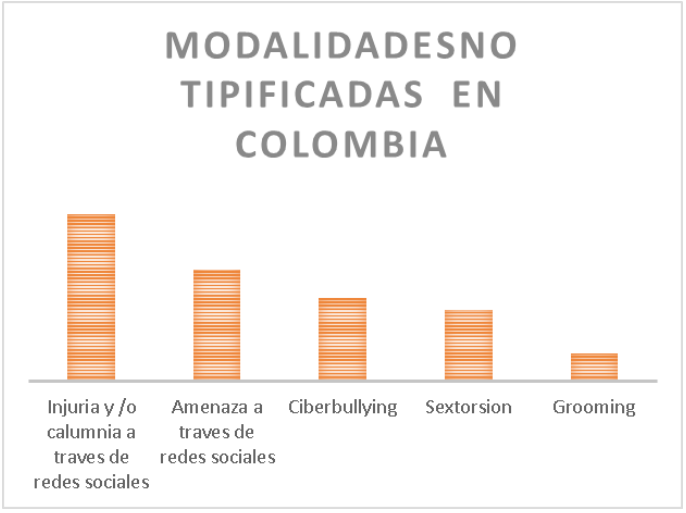


Figura 3 – Modalidades no tipificadas en Colombia

el citado periodo de tiempo, se evidencia un aumento significativo en el número de estas por conductas delictivas que vulneraron la integridad personal, patrimonio económico de entidades público - privadas, así como la integridad, disponibilidad y confidencialidad de la información que circula a través del ciberespacio.

El hurto por medios informáticos y semejantes es la tipología criminal de mayor frecuencia, equivalente al 68%, seguido del acceso abusivo a un sistema informático con el 13% y violación de datos personales con 12% de la muestra.

Se refleja con esto, que las intenciones cibercriminales desde y hacia Colombia están ligadas principalmente a los campos comerciales y financieros, afectando tres aspectos primordiales que soportan el comercio electrónico que son: confianza, sistemas de pago y seguridad.

En la figura 3, se puede observar las modalidades más frecuentes de los ciberdelitos en Colombia. Este tipo de modalidades afectan en su gran mayoría a los niños, niñas y adolescentes con un 75%, teniendo en cuenta que son más susceptibles al engaño y vulnerables en el ciberespacio.

Figura 4 – Mapa de calor ciber ataques en Colombia

En la figura 4, se evidencia que la mayor cantidad de ataques en Colombia se presentan en las ciudades capitales, las principales ciudades con más reportes de incidentes informáticos se observan en la tabla 1.

4. Ciber Guerra y Ciber Defensa

El término de ciberguerra se refiere a aquella lucha desarrollada en el espacio virtual, en donde actores estatales y no estatales se enfrentan por beneficios particulares e incluso altruistas, dejando de paso pérdida de información, propiedad intelectual y hasta económica.

La ciberguerra en un nivel profundo se trata de conocimiento, acerca de quién sabe, cuándo, dónde, qué, y por qué, así como cuán segura es una sociedad o estamento militar con respecto al conocimiento de sí mismo y de su adversario (Riascos 2015).

REPORTE DE INCIDENTES INFORMATICOS	
CIUDAD	CANTIDAD
BOGOTA	9709
MEDELLIN	691
CALI	475
BARRANQUILLA	240
BUCARAMANGA	1273
CIUDADES CON MAS DENUNCIAS LEY 1273	
BOGOTA	2607
CALI	1607
MEDELLIN	998
BUCARAMANGA	594
IBAGUE	448
BARRANQUILLA	398

Tabla 1 – Reporte de incidentes y denuncias en Colombia.

Similarmente el concepto de ciberpoder, al cual se refiere como un "conjunto de recursos relacionados con la creación, control y comunicación de información basada en sistemas electrónicos y computacionales, esto incluye no solo la internet, computadores conectados en red, sino también intranets, tecnología celular y comunicaciones satelitales" (universidad externado de colombia 2017).

Ciberguerra y ciberpoder guardan una relación absoluta, la cual está orientada al control de un recurso clave que alimenta el poder de un actor sobre otro, esto es, la información como elemento decisivo en el ciberespacio o en un campo de batalla real (Vasquez 2015).

Es evidente que el Cibercrimen y los fenómenos acordes que se ciernen en un espacio virtual que carece de soberanías, representan sobre el mundo tangible, caracterizado por territorialidades e intereses, una amenaza de difícil control por parte de un Estado, sin alianzas estratégicas y grupos colaborativos tanto a nivel nacional como internacional.

En este sentido, las consecuencias en términos distributivos dada la interdependencia económica en el sistema internacional y la existencia de jurisdicciones nacionales, crean un interés e incentivo común que, según Keohane (1984), es el principio para el desarrollo de la cooperación entre Estados. En un sentido similar, la dificultad de persecución sobre actores atomizados y dispersos en el mundo hace necesaria la cooperación policiva entre agencias de seguridad nacionales y transnacionales, a la hora de perseguir en el mundo físico crímenes y acciones que hallan su base de operación en un mundo informático.

Las Fuerzas Armadas alrededor del mundo se preparan desde hace tiempo para las batallas y guerras en un nuevo escenario cyber. Toda actividad bélica en nuestros días trae consigo más allá de los medios y métodos tradicionales, una estrategia de ciberataques para debilitar, confundir o suprimir al enemigo. Las guerras en el siglo XXI tienen un mayor componente digital y de la información que nunca. También las técnicas de la guerra experimentan transformación digital.

Cabe recordar que la segunda confrontación global fue una victoria para los Aliados en gran medida por las Tecnologías de la Información, con el Profesor Alain Turing y los sabios -precursores de los científicos informáticos- que lo acompañaban descifrando las ordenes y mensajes de los comandantes Nazis. (Nye 2011) Luego con el uso de la tecnología nuclear por los EEUU en Hiroshima y Nagasaki se ratificó que estábamos como género humano cada vez más lejos de la confrontación física y las TIC a gran escala se tomaban el escenario de la guerra.

Hoy en día la profundización de esa tendencia con la combinación de tecnologías como robótica, drones, inteligencia artificial y big data para las nuevas confrontaciones. Los sistemas de información públicos y privados y las infraestructuras esenciales son el blanco de las nuevas tecnologías de la guerra. La ciberdefensa exige un blindaje, barreras de entrada, políticas, prácticas, estrategias para evitar las consecuencias de ataques devastadores.

4.1. Ciber ataque y Ciber defensa

La estrategia ciberbélica busca generar poderosas armas que permitan de ser necesario una ofensiva y agredir a los enemigos causando los mayores daños y mayores beneficios. También se requiere la defensa adecuada de los bienes digitales y de los sistemas de información. Las barreras y escudos que blinden a recursos críticos e instalaciones esenciales.

Las armas ofensivas y las defensivas deben ser disuasivas para que los países enemigos o con intereses opuestos no se atrevan a atacar. Los nuevos arsenales de los países están enfocados en el ciberespacio. La vigilancia, el robo de información o la destrucción son los principales objetivos de las cyberweapons.

En el reciente ataque internacional Wanna Cry, el arma desarrollada por una agencia de inteligencia aparentemente de EEUU y el escenario global de muchos equipos

Figura 5 – Ciber ataques en el mundo (http://www.norse-corp.com/)

informáticos con vulnerabilidades en Código no actualizado, conectados ampliamente a internet y a redes internas.

La guerra digital utiliza medios a distancia, no presenciales. Las armas no tienen control mediante tratados internacionales como sí ocurre por ejemplo con la proliferación de armas nucleares.

La ética del soldado, el límite del ser humano en la confrontación con el otro ahora encuentra un esguince, la inteligencia artificial como arma bélica.

Cada país en ejercicio de soberanía y con la mira en el cumplimiento del deber general de defensa de la vida, honra y bienes debe llevar a cabo una transformación digital de sus estrategias, tácticas, medios de defensa y ataque en el ecosistema digital.

Como se observa en la figura 5, los ciber ataques en el mundo se dan a cada segundo y son cada vez más frecuentes entre potencias mundiales, que al parecer, ya se encuentran en una guerra no declarada a través del ciber espacio.

4.2. Política de ciberdefensa en Colombia

En Colombia, la política pública que concierne a los temas de seguridad digital se estableció inicialmente en el documento CONPES 3701 de 2011, que incluye los lineamientos de política para ciberseguridad y ciberdefensa. Esta política busca mitigar los riesgos derivados de las amenazas informáticas y desarrollar políticas de prevención y control. La Política crea el sistema de seguridad digital coordinado por el Grupo de Respuesta a Emergencias Cibernéticas de Colombia (ColCERT), junto con el Centro Cibernético Policial (CCP) y el Comando Conjunto Cibernético (CCOC). EL ColCERT coordina los Equipos de Respuesta ante Incidencias de Seguridad (CSIRT). Las funciones compartidas por estas instituciones están encaminadas a: 1) responder a los incidentes informáticos, 2) responder a la amenaza de los delitos informáticos y 3) velar por la seguridad digital nacional.

En 2016, mediante el documento Conpes 3854 se determinó una nueva política de seguridad digital. El documento parte de una visión más enfocada a restablecer la confianza de los usuarios y lograr que el ecosistema digital sea una plataforma segura para los negocios electrónicos también desarrolla el principio de la responsabilidad compartida entre las múltiples partes interesadas en el cual cada sujeto tiene un grado de responsabilidad en la gestión de los riesgos de seguridad y en proteger el entorno digital. La nueva visión comprende la gobernabilidad participativa de múltiples partes interesadas, y definir los niveles de escalamiento para el reporte de incidentes digitales.

En Colombia se han creado laboratorios de ciber, quienes se encargan de combatir el ciber crimen, recolectar evidencia digital y analizarla, perseguir ciberdelincuentes a través de la red, lo cual no es tarea fácil, más aun cuando un ataque se puede realizar desde cualquier computadora conectada a la red

Los talentos en desarrollos digitales, en las novísimas tecnologías deben hacer parte de las nuevas Fuerzas Armadas, consolidar nuevos “batallones cyber” y consolidar estrategias de defensa.

A escala global crece la discusión sobre la aplicación de las reglas tradicionales sobre el derecho de la guerra entre las naciones, Convenios de Ginebra de 1949 y Protocolos Internacionales pueden aplicarse al nuevo escenario de la guerra digital.

En suma, a la tierra, mar y aire como elementos distintivos de las distintas Fuerzas Armadas se debe incorporar el ciberespacio. Otro reto para varias generaciones en el Siglo XXI.

5. Conclusiones y Trabajo Futuro

Al realizar el modelo ontológico de los ciberdelitos tomando como caso de estudio Colombia, se logró definir que los ciberdelitos están compuestos por la forma en que se realiza el ataque y su definición punitiva en el código penal, permitiendo establecer la relación entre la forma en que se realiza la conducta y su tipicidad.

Se tienen en cuenta diferentes aspectos como son clasificación, jurisprudencia, nivel de impacto y se determinan las nuevas modalidades que se están presentando. A través del modelo ontológico se coleccionan conceptos clave y su interrelación colectiva con el fin de proporcionar una vista abstracta y tipificar así mismo los ciberdelitos. Una ontología es una colección de conceptos clave y su interrelación colectiva con el fin de proporcionar una vista abstracta de cualquier dominio de aplicación, con el fin de tipificar acertadamente los ciberdelitos. Este documento busca brindar herramientas que permitan reconocer el ciberdelito en Colombia, de igual forma, establecer patrones de conducta que permitan prevenir ser víctima de estos delitos.

El cambio social ha determinado el surgimiento de nuevos riesgos que se materializan en nuevas fenomenologías criminales. Los cibercrímenes representan, justamente, el nacimiento de comportamientos que, en su sentido puro, como delitos informáticos en sentido estricto, tienen lugar en realidades virtuales o simuladas como el ciberespacio, protegen objetos inmateriales como los datos y la información y tutelan nuevos bienes jurídicos intermedios. Si a ello se suma el hecho de que estos delitos son realizados mediante técnicas particulares (que en muchos casos consisten en la manipulación de energía o bytes basados en sistemas binarios, que pueden ser traducidos a los humanos mediante infraestructuras informáticas o Hardware), se aprecian delitos que distan mucho de ser asimilables a los delitos tradicionales.

Parece necesario y relevante complementar las categorías tradicionales del delito en la tipicidad, con una perspectiva digital que, por cierto, ya no es la excepción a la regla, sino que comienza a ser la regla general en la criminalidad moderna. Esto incluso en ámbitos hasta ahora reservados a la criminalidad física, como la criminalidad organizada y transnacional, que comienza a actuar mediante organizaciones virtuales transnacionales (OVT) que dificultan aún más combatir este tipo de delincuencia sin fronteras.

En Colombia por ser un país con un alto índice de conectividad a internet, está expuesto a los diferentes tipos de ciber ataques, sumado a esto, el nivel de capacitación acerca de la navegación segura y la prevención hacia los delitos informáticos es muy baja, siendo así que muchos colombianos navegan, pero pocos saben hacerlo de manera responsable y segura. De igual forma, la denuncia de estos delitos es muy baja, ya que las personas ignoran que muchas de las conductas de las cuales son víctimas, se consideran un delito y están tipificadas en el Código Penal, lo cual dificulta que se ataque el fenómeno delictivo de forma certera y efectiva.

A futuro se investigará sobre como nuevas tendencias del ciber crimen en el mundo impactan a Colombia.

Referencias

Ontological Knowledge Model to Engineering Project Integration Based on PMS. MSc(c). Yuri Nieto 1, Dr. Roberto Ferro 2, Dr. Carlos Montenegro3

El cibercrimen: Fenomenología y criminología de la delincuencia en el ciberespacio. Fernando Miró Llinares. 2012 http://www.microsoftcolombia.com/dia-de-cibercrimen-colombia.

Design of an Assisting Agent Using a Dynamic Ontology, Alexis Koutero, Shigeru Fujita, Kenji Sugawara, 2010

Yang, D.: Product Configuration Knowledge Modeling Using Ontology Web Language. Expert Systems With Applications 36(3), 4399–4411 (2009)

L. S. P. Busagala, W. Ohyama, T. Wakabayashi, and F. Kimura, "Improving automatic text classification by integrated feature analysis," IEICE Trans. Inf. Syst., vol. E91-D, no. 4, pp. 1101–1109, Nov. 2008.

Y. Sure, J. Angele, and S. Staab, "OntoEdit: Guiding ontology development by methodology and inferencing," Lect. Notes Comput. Sci., vol. 2519, pp. 1205–1222, 2002. (http://www.mintic.gov.co/portal/604/w3-article-1629.html, s.f.) (http://www.mintic.gov.co/portal/604/articles-3510_documento.pdf, s.f.)https://www.dinero.com/empresas/articulo/sectores-mas-afectados-por-cibercrimen-en-colombia/250321 https://caivirtual.policia.gov.co/sites/default/files/informe_amenazas_de_cibercrimen_en_colombia_2016_-_2017.pdf https://caivirtual.policia.gov.co/ciberincidentes/tiempo-real

Arquilla y Ronfeldt (1993, p. 27) Nye (2011, p. 123)

Nada volverá a ser igual: ciberguerra y ciberpoder. Autor Javier Alberto Castrillón-Riascos, 2015.

Nos llegó la Ciberguerra. Universidad externado de Colombia. 2017

Propuesta de buenas prácticas para fortalecer los controles de prevención y detección temprana del cibercrimen en las empresas colombianas. Autores Vásquez Zarate, Katy Alejandra Cárdenas Rodríguez, María Paula, 2015.

Ciber crimen y vida diaria en el mundo 2.0, Dr. Fernando Miro Linares, 2015.

Miró, La oportunidad criminal en el ciber espacio, revista electrónica de ciencia penal y criminología, 2011. http://www.norse-corp.com/

Miró, el ciber crimen, fenomenología y criminología de la delincuencia en el ciber espacio, 2012.

Miro, La oportunidad criminal en el ciber espacio, 2011.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 07/10/2018
Aceitação/Acceptance: 19/12/2018

Modelo para predecir el rendimiento académico basado en redes neuronales y analítica de aprendizaje

Nelson Salgado Reyes[1], Jéfferson Beltrán Morales[2], Javier Guaña Moya[3], Charles Escobar Teran[4], Damian Nicolalde Rodriguez[5], Gustavo Chafla Altamirano[6]

nesalgado@puce.edu.ec, jtbeltran@uce.edu.ec, eguana953@puce.edu.ec, cescobar637@puce.edu.ec, danicolalde@puce.edu.ec, gxchafla@puce.edu.ec

[1,3,4,5,6] Docente Facultad de Ingeniería, Av. 12 de octubre 1076 y Roca, 17012184, Quito, Ecuador.

[2] Docente Facultad de Ingeniería Ciencias Físicas y Matemáticas, Av. Universitaria y Bolivia, 170129, Quito, Ecuador.

Pages: 258–266

Resumen: Un elemento de gran importancia para las instituciones educativas universitarias, educadores y estudiantes lo constituye el rendimiento académico de los mismos en el tránsito de su formación profesional. La minería de datos educativos desarrolla modelos y métodos para explorar los datos recopilados de los entornos de aprendizajes educativos mediante analíticas de aprendizajes con el fin de detectar patrones que permitan predecir variables de interés. La presente investigación describe un modelo predicativo de rendimiento académico usando técnicas de redes neuronales sobre un conjunto de datos reales de 300 estudiantes de la carrera de Sistemas de la Universidad Central del Ecuador. Este registro fue suministrado por el entorno virtual de aprendizaje https://uvirtual.uce.edu.ec/ desarrollado en Moodle y usado en dicha Universidad.

Palabras-clave: rendimiento académico; redes neuronales, learning analytics; big data

Model to predict academic performance based on neural networks and learning analytics

Abstract: An element of great importance for university educational institutions, educators and students is the academic performance of them in the transition of their professional training. The mining of educational data develops models and methods to explore the data collected from the educational learning environments through learning analytics in order to detect patterns that allow predicting variables of interest. The present research describes a predictive model of academic performance using neural network techniques on a set of real data of 300 students of the Systems career of the Central University of Ecuador. This registration was provided by the virtual learning environment https://uvirtual.uce.edu.ec/ developed in Moodle and used in said University.

Keywords*:* academic performance; neural networks, learning analytics; big data.

1. Introducción

Durante el tránsito de los estudiantes en la adquisición de competencias que les permita un buen desarrollo futuro de su profesión, afrontan el desafío constante de superar materias académicas. Según la teoría de aprendizaje, la probabilidad de éxito del durante sus estudios es un problema multifactorial, siendo la interacción aprendizaje-enseñanza un elemento transcendental. (Acevedo & Rocha, 2011)

A pesar de que el rendimiento académico es una variable multifactorial, muchos de los estudios alrededor de la misma incluyen solo factores personales y socioeconómicos; sin embargo, el surgimiento y la aplicación de las nuevas tecnologías de enseñanza sobre todo el uso de las plataformas virtuales, permiten a las universidades recolectar una gran cantidad de información en tiempo real. Estos cuantiosos datos electrónicos generados proporcionan un abordaje multivariante en el estudio del rendimiento académico. (Ye & Biswas, 2014)

La analítica de aprendizaje se puede definir como el proceso de determinar, evaluar e interpretar el significado de grandes volúmenes de datos educacionales; utilizando para ello algoritmos matemáticos. (Baker, 2014; Jaramillo & Paz, 2015) Aunque existen diversas técnicas para el análisis de este tipo de información, existen cada vez más trabajos enfocados al uso de algoritmos basados en redes neuronales artificiales. (Isljamovic & Suknovic, 2014)

Las redes neuronales artificiales (RNA) representan esquemas simplificados de la estructura física y del funcionamiento de una neurona biológica; por lo que sus algoritmos poseen capacidad de predecir resultados mediante la clasificación de patrones.

En los últimos años se han desarrollado múltiples tipologías de redes neuronales en función del problema a resolver. Una de las arquitecturas de mayor uso en problemas de clasificación y predicción, responde a redes perceptrón con múltiples capas a las que se asocian una cantidad de neuronas. De esta forma, el aprendizaje se obtiene mediante "retro propagación del error", que se basa en obtener el mínimo de una función de error calculada a través de la salida del sistema y el modelo obtenido por el entrenamiento de la red. (Del Brío, Sanz, & Garcia, 2006)

El uso de analítica de aprendizaje y de estos modelos novedosos, aplicados a la educación no son todavía suficientemente utilizados en Ecuador. A pesar de que muchas Universidades en el país utilizan plataformas con entornos virtuales en los procesos de formación, los datos obtenidos no son aprovechados correctamente para realizar predicciones que permitan tomar decisiones oportunas. De aquí que el objetivo del presente trabajo se centra en elaborar un modelo de predictivo del rendimiento académico usando técnicas de redes neuronales sobre la información obtenida mediante analíticas de aprendizajes.

Dado la problemática, los resultados predictivos del modelo sirven a la Universidad para evitar casos de deserción de estudiantes, porque en general, la eficiencia y efectividad del sistema educativo en formar profesionales está actualmente poco trabajada por las universidades, y organizaciones gubernamentales competentes, lo que genera efectos negativos importantes para el país, sobre todo si se considera que el gobierno proyecta una fuerte inyección de inversión por la reforma educacional y la intención de utilizar este indicador como parte de las acreditaciones de las instituciones educativas a nivel superior.

Las variables motivacionales contribuyen a mejorar el rendimiento académico, pero no son las de más peso en la Universidad Central del Ecuador. Se destaca la motivación docente ligada a la heterogeneidad por facultad, la incidencia de la presión familiar, y la necesidad de ambientación al sistema previo al ingreso.

2. Revisión de la literatura

La mayoría de los estudios sobre "Learning Analytics" se centran en predecir el rendimiento académico del alumno cuantificando el grado final o por la aprobación o no de un determinado curso, dificultando la comparación de las diferentes investigaciones respecto a los predictores más estables de rendimiento académico. (Altujjar, Altamimi, Al-Turaik, & Al-Razgan, 2016; Il-Hyun, Yeonjeong, Jeonghyun, & Jongwoo, 2014)

A continuación, se detallan las que se han considerado más importantes:

Romero, et al (2014) recolectaron datos de estudiantes de un foro de discusión en línea, los clasificaron en varios subconjuntos de datos y luego evaluaron la precisión predictiva de cada conjunto a través de métodos de minería de datos. Para ello utilizaron un algoritmo de clasificación y datos de interacción de los alumnos antes de un examen para predecir el rendimiento en el aprendizaje de los mismos.

Huang&Fang (2013) aplicaron cuatro modelos matemáticos para predecir el rendimiento académico de estudiantes de un curso de ingeniería utilizando para ello las calificaciones finales de los mismos. Los resultados finales mostraron que los puntajes de los exámenes finales de los estudiantes eran predecibles con un 88% de precisión en base a ocho variables recopiladas de un sistema de gestión de aprendizaje (LMS).

Usman&Adenub (2013), desarrollaron un modelo basado en redes neuronales artificiales para predecir la calificación final de estudiantes universitarios antes degraduarse. Se trabajó con 30 estudiantes del Departamento de Ciencias de la Computación de la Universidad de Educación "Tai Solarin". Los resultados mostraron la capacidad de las redes neuronales para predecir correctamente la calificación final de los estudiantes con un 92.7% de precisión.

Hu&Shih (2014) desarrollaron un sistema de alerta temprana basado en árboles de decisión para predecir si los estudiantes aprobarían o no. El modelo fue construido utilizando datos de 300 estudiantes y 13 variables recogidos mediante analítica en línea. Los resultados revelaron un 95% de precisión.

Isljamovic&Suknovic (2014), utilizan diferentes algoritmos de redes neuronales artificiales para encontrar el modelo más adecuado para la predicción de éxito académico de estudiantes graduados en la facultad de ciencias de la organización de la Universidad de Belgrado. El modelo permite identificar aquellos factores de mayoy influencia en el rendimiento académico de los estudiantes.

Zacharis (2016), construyó un modelo basado en redes neuronales multi capas para predecir el rendimiento estudiantil en un entorno de aprendizaje mixto. Para ello utilizo datos de los estudiantes almacenados en un servidor Moodle y predijo el éxito del estudiante en curso, basado en cuatro actividades de aprendizaje: comunicación por correo electrónico, creación de contenido colaborativo con wiki, interacción de contenido

medida por archivos vistos y autoevaluación mediante cuestionarios en línea. El modelo predijo el rendimiento de los estudiantes con una tasa de clasificación correcta, del 98,3%.

3. Marco teórico

3.1. Analítica de aprendizaje en la educación

La disponibilidad de grandes conjuntos de datos obtenidos por la interacción de los estudiantes en los "Entornos Virtuales de Aprendizaje" en las instituciones educativas proporcionan grandes oportunidades para mejorar el proceso de aprendizaje de los mismoss; dando lugar a un enfoque novedoso en el ámbito educativo conocido como "analítica de aprendizaje" (Learning Analytics). (Abu, 2016; Rojas, 2017)

El "learning analytics" aprovecha la gran cantidad de información que generan los estudiantes y sus contextos durante el proceso educativo con la finalidad de entender, mejorar y optimizar el aprendizaje y el rendimiento académico. (Conijn, Snijders, Kleingeld, & Matzat, 2017). Según Daud (2017), los datos obtenidos pueden ser útiles en diferentes contextos: en el caso de los docentes estos pueden monitorear el progreso de los estudiantes, los mismos que pueden retroalimentarse. De igual forma los directores y administradores esta información les permite tomar decisiones respecto a incrementos de aulas, tasas de graduación, ect.

3.2. Redes neuronales

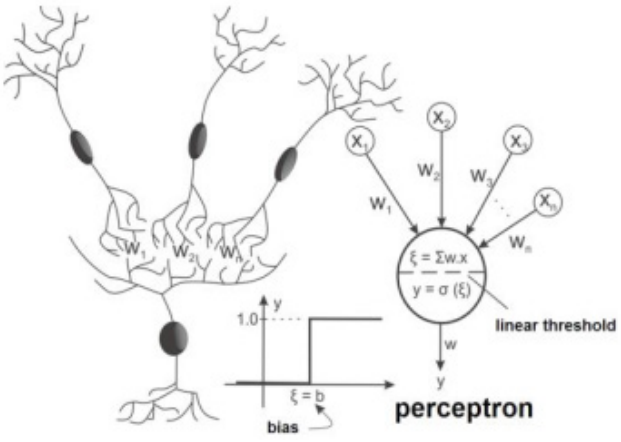


Figura 1 – Esquema de un modelo neuronal perceptrón
Fuente: (Del Brio, Sanz, & Garcia, 2006)

Las redes neuronales artificiales (RNA) constituyen algoritmos computacionales inspirados en las redes neuronales biológicas, comportándose de forma similar a las mismas en sus funciones más comunes. Las RNA son entrenadas para aprender, generalizando resultados y extrayendo patrones de un conjunto de datos. (Del Brio, Sanz, & Garcia, 2006)

Las redes neuronales artificiales reciben los datos de entrada del entorno y los agregan de una manera especial para formar una entrada de red (ξ), que pasan a través de un umbral lineal y la salida, "y" se transfiere a la neurona vecina o al entorno (Figura 1). El disparo neural se produce cuando ξ supera el umbral del borde denominado sesgo, "b". La dinámica de la neurona lineal se asume comúnmente durante el cálculo de ξ y la entrada de la red se calcula como un producto escalar de las señales de entrada "x" independientemente de la neurona y su peso "w".

Para n señales, la operación de la neurona perceptrón se expresa mediante:

$$y = \begin{Bmatrix} 1, \sum_{i=1}^{n} w_i x_i \geq b \\ 0, \sum_{i=1}^{n} w_i x_i < b \end{Bmatrix} \quad (1)$$

Un peso positivo ($w_i > 1$) aumenta la neurona, mientras que un peso negativo reduce e inhibe la actividad neural. Este tipo de neuronas artificiales se denomina perceptrón porque establece un mapeo entre la actividad de entrada (estímulo) y la señal de salida. El umbral neuronal puede considerarse como un nodo de entrada adicional cuyo valor es siempre 1, y su peso es igual a "b".

En problemas no lineales las capas adicionales de neuronas se insertan entre la capa de entrada y la neurona de salida, por lo que se obtiene una arquitectura de red neuronal perceptrón multicapa (MLP). Cómo esta capa no puede interactuar con el entorno, se denomina capa oculta y los nodos de esta capa se denominan nodos ocultos. Las capas ocultas adicionales aumentan la capacidad del perceptrón para resolver un problema de clasificación no lineal.

Las redes neuronales con propagación hacia atrás representa es un tipo de MLP entrenado. Sin embargo, el proceso de aprendizaje es una actualización del algoritmo de perceptrón simple durante el cual los pesos se están ajustando en las capas ocultas.

La función de activación de las neuronas en la capa oculta desempeña un papel para proporcionar el aprendizaje de la red neuronal. Las neuronas de las capas ocultas no tienen una capacidad superior a la red de perceptrón ordinaria en el caso de problemas lineales. Las funciones de activación no lineales a menudo se usan para las neuronas en la capa de salida, lo que influye en que la red con múltiples capas sea particularmente poderosa. Las funciones de activación más frecuentemente usadas en el algoritmo de propagación hacia atrás son:

- Linear: x; $(-\infty, +\infty)$
- Sigmoidea logística: $\frac{1}{1+e^{-x}}$; $(0, +1)$
- Tangente hiperbólica: $\frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}$; $(-1, +1)$
- Exponencial: e^{-x}; $(0, +\infty)$
- Escalón: $\begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$; $[0, +1]$

4. Metodología

A continuación, se describe la metodología utilizada y aplicada a un conjunto de 300 estudiantes de la "Carrera de Ingeniería Informática" de la Universidad Central del Ecuador.

4.1. Recolección de datos

La información de los estudiantes se obtiene a través del intercambio de los mismos con el "Entorno Virtual de Aprendizaje" (https://uvirtual.uce.edu.ec/), desarrollado en Moodle. Los registros de datos de suministrados por el sistema fueron extraídos y sometidos a una limpieza utilizando los paquetes "DataClean" y "DataCombine" del software de uso libre "R". De esta forma detectamos las variables disponibles que formaron parte de la investigación y que mostramos en la tabla1

Variable predictora	
Número total de clicks	*Número de consultas al profesor*
Cantidad de sesiones en linea	*Participación en los foros*
Tiempo total en linea (min)	*Participación en chat*
Número de visitas a la página del curso	*Participación en video-colaboración*
Número de comentarios	*Número de recursos vistos*
Irregularidad del tiempo de estudio	*Número de enlaces vistos*
Irregularidad del intervalo de estudio	*Número de páginas de contenido vistas*
Mayor periodo de inactividad	*Número de publicaciones de discusión vistas*
Tiempo hasta la primera actividad (min)	*Número total de publicaciones de discusión*
Tiempo promedio por sesión (min)	*Numero de cuestionarios iniciados*
Número de intentos por prueba	*Número de asignaciones enviadas*
Número de cuestionarios vistos	*Número de presentación vistas*
Número de cuestionarios pasados	*Número de ediciones en la wiki*
Número de wiki vistas	*Grado promedio de evaluación*

Tabla 1 – Variables predictoras usadas

4.2. Análisis de datos

Las variables obtenidas fueron procesadas, utilizando el paquete "*nnet*" de R, que permite la creación, aplicación, visualización e implementación de redes neuronales de alimentación directa con una sola capa oculta, para modelos multinomiales log-lineales.

El diseño de la red neuronal utilizada se basó en una topología de red neuronal multicapa (MLP), la cual fue entrenada de forma independiente para cada asignatura obteniendo predicciones únicas de forma correcta.

La arquitectura de la red estuvo determinada por una capa de entrada y una capa intermedia u oculta con funciones de activación sigmoidal tangente hiperbólica y una capa de salida con función de activación lineal.

5. Resultados

La tabla 2, describe la precisión de clasificación y recuperación para cada curso y las diferencias en el rendimiento predictivo de la red neuronal entrenada individualmente en los datos de cada curso.

Num. curso	Curso	Exact.	Exact. base	Recordado	Grados asignación
1	Programación I	0.621	0.533	0.521	-
2	Programación II	0.772	0.685	0.570	-
3	Estructura de Datos	0.588	0.610	0.621	-
4	Programación Distribuida	0.866	0.810	0.857	Si
5	Aplicaciones de Minería de Datos	0.917	0.895	0.906	Si
	Promedio	0.7528		0.6950	-

Tabla 2 – Rendimiento de la red neuronal por cursos

Según los resultados obtenidos, el curso de *Aplicaciones de Minería de Datos* presenta el mayor valor de precisión y recuerdo con un 91,7% y 89,5% respectivamente. Al observar los resultados restantes, las precisiones de predicción van del 58.8% para el curso de *Estructura de Datos* al 86,6% para el curso de *Programación Distribuida*. Al promediar los puntajes obtenidos, encontramos una precisión del 75,28%. De igual forma podemos encontrar que los puntajes de recuerdo tienen una alta variabilidad que va del 52,1% en *Programación I* hasta el 90,6% para el curso de Aplicaciones de Minería de Datos; con valor promedio de recuperación de 69,50%.

Junto a los resultados anteriores se extrajeron las características de importancia para cada uno de los cursos, los cuales se exponen en la tabla 3. En el caso del curso de *Minería de Datos*, que como habíamos mencionado presenta el mayor valor de precisión, tenemos que el grado del cuestionario fue la característica más relevante con un 12,4%; la regularidad del estudio y el promedio de clics por sesión fueron los siguientes mejores predictores con un 12,4% y 10,9% respectivamente. En contraste a estos resultados tenemos que el grado promedio de las evaluaciones representa una característica de poca importancia en este curso, pues solo obtuvo un puntaje de 2,4%; resultado que puede explicarse ya que algunas calificaciones son un registro de una determinada tarea mas no refleja que tan bien se realizó la misma.

El otro curso que, según los resultados, tiene calificaciones y asignación como predictor es el de *Estructura de Datos*. Para el mismo, el promedio de las evaluaciones fue la característica más importante representando un 25,3%; seguido del promedio de clics por sesión con un 15,6%. Las características menos relevantes fueron: los mensajes enviados, publicaciones en foros/chat/videos, el grado promedio del examen y el número de cuestionarios realizados con un puntaje de 0.0%.

5. Conclusiones

El objetivo de esta investigación es determinar la efectividad de un modelo basado en redes neuronales para predecir el rendimiento del estudiante. La revisión de la literatura indicó que estos enfoques superan a todos los otros clasificadores, con respecto a la precisión de la predicción. Una red neuronal de perceptrón multicapa fue entrenada por un algoritmo de propagación inversa, para predecir la capacidad de pasar exitosamente la carrera.

Características	Curso 4	Curso 5
Regularidad de estudio	12.4	11.6
Grado promedio de examen	19.5	0.0
Número de cuestionarios realizados	9.4	0.0
Máximo periodo de inactividad	9.7	10.0
Longitud total de sesión	10.0	10.3
Total número de sesiones	8.3	9.6
Total número de clics	7.4	11.9
Promedio de clics por sesión	10.9	13.5
Promedio de longitud de sesión	8.1	8.9
Número de asignaciones realizadas	1.9	1.9
Mensajes enviados	0.0	0.0
Publicaciones en foro, chat, video	0.0	0.0
Grado promedio de evaluación	2.4	22.4

Tabla 3 – Características importantes por cada curso

La tasa de precisión de la clasificación es alta, con un promedio de 75,28% para todos los cursos; lo que muestran la efectividad de los predictores obtenidos para la predicción del rendimiento académico.

Referencias

Abu, A. (2016). Educational Data Mining & Students' Performance Prediction. *International Journal of Advanced Computer Science and Applications(IJACSA)*, 212–220.

Acevedo, C., & Rocha, F. (2011). Estilos de aprendizaje, género y rendimiento. *Journal of Learning Styles*, 71-84.

Altujjar, Y., Altamimi, W., Al-Turaik, & Al-Razgan, M. (2016). Predicting Critical Courses Affecting Students Performance: A Case Study. *Procedia Computer Science*, 65–71.

Baepler, P., & Murdoch, C. (2010). Academic analytics and data mining in higher education. *International Journal for the Scholarship of Teaching and Learning*, 1–9.

Baker, R. (2014). Educational data mining: An advance for intelligent systems in education. *IEEE Intelligent Systems*, 78–82.

Conijn, R., Snijders, C., Kleingeld, A., & Matzat, U. (2017). Predicting student performance from LMS data: A comparison of 17 blended courses using Moodle LMS. *IEEE Transactions on Learning Technologies*, 1–15.

Daud, A., Radi, N., Ayaz, R., Lytras, M., & Abbas, F. (2017). Predicting Student Performance using Advanced Learning Analytics. *Proceedings of the 26th International Conference on World Wide Web Companion* (págs. 415-421). Australia: WWW '17 Companion.

Del Brio, B. M., Sanz, A., & Garcia, A. (2006). *Redes neuronales y sistemas borrosos*. EUU: RA-MA EDITORIAL.

González, J. C., Ramos, S., & Hernández, S. (2017). Modelo Difuso del Rendimiento Académico Bi-explicado. *Revista de Sistemas y Gestión Educativa*, 55–64.

Hamasa, H., Indiradevi, S., & Kizhakkethottam, J. (2016). Student academic perfomance prediction model using decision tree and fuzzy genetic algorithm. *Procedia Technology* , 326–332.

Hu, Y., Lo, C., & Shih, S. (2014). Developing early warning systems to predict students' online learning performance. *Computers in Human Behavior*, 469–478.

Huang, S., & Fang, N. (2013). Predicting student academic performance in an engineering dynamics course: A comparison of four types of predictive mathematical models. *Computers & Education* , 133–145.

Il-Hyun, J., Yeonjeong, P., Jeonghyun, K., & Jongwoo, S. (2014). Analysis of Online Behavior and Prediction of Learning Performance in Blended Learning Environments. *Educational Technology International*, 71–88.

Isljamovic, S., & Suknovic, M. (2014). Predicting students academic perfomance using artificial neural network : A case study. *The Eurasia Proceedings of Educational & Social Sciences (EPESS)*, 68–72.

Jaramillo, A., & Paz, H. (2015). Aplicación de Técnicas de Minería de Datos para Determinar las Interacciones de los Estudiantes en un Entorno Virtual de Aprendizaje. *Revista Tecnológica ESPOL*, 64–90.

Lu, O., Huang, A. Y., Huang, J. C., Lin, A. J., Ogata, H., & Yang, S. J. (2018). Applying Learning Analytics for the Early Prediction of Students' Academic Performance in Blended Learning. *Journal of Educational Technology & Society*, 220–232.

Rojas, P. (2017). Learning Analytics: A Literature Review. *Educ.Educ*, 106–128.

Schalk, P., Wick, D., Turner, P., & Ramsdell, M. (2011). Predictive assessment of student performance for early strategic guidance. *In Frontiers in Education Conference (FIE)*. Rapid City, Estados Unidos de América.

Usman, O., & Adenubi, A. (2013). Artificial Neural Network (ANN) Model for Predicting Students' Academic Performance. *Journal of Science and Information Technology*, 23–37.

Ye, C., & Biswas, G. (2014). Early prediction of student dropout and performance in moocs using higher granularity temporal information. *Journal of Learning Analytics*, 169–172.

Zacharis, N. Z. (2016). Predicting student academic performance in blended learning using artificial neural networks. *International Journal of Artificial Intelligence and Applications (IJAIA)*, 17–29.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 15/09/2018
Aceitação/Acceptance: 05/11/2018

Jardín botánico: Prototipo desoftware para la gestión y divulgación de plantas nativas basado en código QR y realidad aumentada

Edwin Barrientos-Avendaño[1], Dewar Rico-Bautista[2], Luis Anderson Coronel-Rojas[3], Fabian Ranulfo Cuesta-Quintero[4]

ebarrientosa@ufpso.edu.co, dwricob@ufpso.edu.co, lacoronelr@ufpso.edu.co, fcuestaq@ufpso.edu.co.

[1] Departamento Sistemas e Informática. Grupo de ingeniería en innovación, tecnología y emprendimiento (GRIITEM). Universidad Francisco de Paula Santander Ocaña, Sede Algodonal Vía Acolsure, 546551, Ocaña, Colombia.

[2] Departamento Sistemas e Informática. Grupo (GRIITEM). Universidad Francisco de Paula Santander Ocaña, Sede Algodonal Vía Acolsure, 546551, Ocaña, Colombia.

[3] Departamento Sistemas e Informática. Grupo (GRIITEM). Universidad Francisco de Paula Santander Ocaña, Sede Algodonal Vía Acolsure, 546551, Ocaña, Colombia.

[4] Departamento Sistemas e Informática. Grupo (GRIITEM). Universidad Francisco de Paula Santander Ocaña, Sede Algodonal Vía Acolsure, 546551, Ocaña, Colombia.

Pages: 267–282

Resumen: En la universidad Francisco de Paula Santander Ocaña (UFPS Ocaña) se encuentra un jardín botánico. El Jardín Botánico Universitario Jorge Enrique Quintero Arenas fue creado con la misión de conservar el ecosistema de bosque seco y, los tipos de vegetación y flora presentes del nororiente de Colombia, propiciando espacios adecuados para la investigación, y la educación ambiental. Este artículo da a conocer los resultados de un prototipo de software.En la primera fase, se da a conocer el estado de arte de los modelos, tecnologías, análisis de términos y conceptos. En la segunda, se visualiza la definición de los requerimientos funcionales y no funcionales que permitan la virtualización del jardín botánico. La tercera fase, es el diseño del esquema interactivo (álbum digital) basado en herramientas tecnológicas como realidad aumentada y realidad virtual, de las diferentes plantas y animales. En la fase final, se muestra el prototipo de la aplicación móvil para tener un control y seguimiento de las plantas.

Palabras-clave: Internet de las cosas; Código QR; Álbum Digital; Realidad Aumentada; Universidad inteligente.

Botanical garden: Software-prototype for management and divulgation of native plants based on QR code and augmented reality.

Abstract:At the Francisco de Paula Santander Ocaña University (UFPS Ocaña) there is a botanical garden. The “Jorge Enrique Quintero Arenas” University

Botanical Garden was created with the mission of conserving the dry forest ecosystem and the types of vegetation and flora present in the northeast of Colombia, providing adequate spaces for research and environmental education. This article presents the results of a software prototype. In the first phase.In the first phase, the state of the art of the models, technologies, analysis of terms and concepts is presented. In the second phase, it visualizes the definition of the functional and non-functional requirements that allow the virtualization of the botanical garden. The third phase is the design of the interactive scheme (digital album) based on technological tools such as augmented reality and virtual reality of the different plants and animals. In the final phase, the prototype of the mobile application is shown to have a control and monitoring of the plants.

Keywords: Internet of Things; QR Code; Digital Album; Augmented Reality; Smart University.

1. Introducción

Hoy en día, los jardines botánicos se consideran los espacios verdes públicos más importantes e influyentes en la mayoría de los países del mundo. Soportan aplicaciones científicas y de investigación en los campos de la botánica, la jardinería, la conservación y la ecología vegetal(Barkoosaraei & Moshiri, 2017). Varían ampliamente en diseño y propósito, pero la mayoría están típicamente asociados con la conservación del medio ambiente, la educación o la interpretación histórica. Sin embargo, los estudios han demostrado que los visitantes del jardín botánico a menudo están motivados por intereses recreativos y de ocio (Wassenberg, Goldenberg, & Soule, 2015). Una comunidad puede volverse inteligente y ecológica a través del despliegue estratégico de la infraestructura y los servicios de Tecnología de la Información y las Comunicaciones para alcanzar los objetivos de la política de sostenibilidad (Luvisi & Lorenzini, 2014). Uno de los principales atributos que distingue a un jardín botánico de un jardín puramente expositivo es la práctica de acceder a las plantas cuando ingresan a la colección y luego llevar registros de este material a lo largo de su vida (Walter, 1991), (Agostini et al., 2017). las plantas también pueden utilizarse como bioindicadores y su participación en las redes de comunicación puede representar una contribución significativa para construir una comunidad inteligente y verde (Barkoosaraei & Moshiri, 2017).

La infraestructura de sensores para el jardín reúne y analizan datos sobre cambios de clima y condiciones de humedad del suelo y luego se conecta al teléfono del usuario con alertas oportunas. Monitorea continuamente condiciones y alerta al usuario de los cambios que requieren acción inmediata. A diferencia de los aspersores preestablecidos, un buen sistema podrá controlar automáticamente el sistema de riego existente basado en los datos recopilados por el sensor de jardín (Biswal, Singh, Jeberson, & Dhar, 2015). Los parámetros se almacenan en estación de trabajo local y se comparte con la nube para ser mostrado en un sistema web, el objetivo principal del diseño es resolver la gestión de problemas, mantenimiento, detección de fallas y tiempo medio de reparación (Basuvaiyan & Rathinasabapathy, 2017). Siendo coherentes la infraestructura debe estar soportada por energías solares.

Las universidades tienen un papel que desempeñar para abordar temas en relación con la disponibilidad de un personal debidamente cualificado, a través de la provisión de graduados en áreas de conocimiento soporte de la agricultura. Los jardineros

humanos con experiencia son muy buenos para tratar las plantas de forma individual. Desafortunadamente, contratar a un equipo de jardineros para llevar a cabo esta tarea en grandes invernaderos no es rentable. Desde el 2013 la Universidad Francisco De Paula Santander Seccional Ocaña (UFPS Ocaña), a través de la facultad de ciencias agrarias y del ambiente, ha respondido a la necesidad de la comunidad científica universitaria de crear un espacio destinado para la conservación, la investigación y la educación ambiental dentro de la institución. Este espacio se conoce hoy como el JARDÍN BOTÁNICO UNIVERSITARIO JORGE ENRIQUE QUINTERO ARENAS, creado con la misión de utilizar la apreciación botánica para inspirar, educar y servir como foco de desarrollo científico y comprometido con la diversidad biológica del nororiente del país. El jardín botánico universitario es un gran complejo para el desarrollo científico de la región en un área tan sensible como la botánica y servirá de apoyo para la creación de programas de conservación del Catatumbo y el desarrollo de trabajos de grado de los programas de pregrado y posgrado en la institución.

Durante los años siguientes se ha consolidado como uno de los proyectos más importantes dentro del plan de desarrollo 2015 – 2019 de la universidad. Mediante este tiempo, cientos de estudiantes han disfrutado de los servicios del jardín, visitas guiadas a las colecciones vivas y al área de conservación, charlas, conferencias, celebraciones, orquideario, camping, avistamiento de aves, observatorio astronómico, entre muchas otras incluidas en un plan operativo, ver apéndice A. Siempre con la intención de generar dentro de los estudiantes y la comunidad universitaria en general, el sentido de pertenencia y la sensibilidad ambiental tan necesaria para hacer frente a los retos emergentes en el escenario global actual.Es por eso que el jardín botánico de la Universidad Francisco de Paula Santander Ocaña será muy importante para el desarrollo socio ambiental de la universidad, así como también en mantener una función ecológica de los bosques que se encuentran en los predios, convirtiéndose en promotor de la conservación del bosque seco de la región, de educación ambiental y de Investigación de la flora nativa del departamento; además por tener en sus programas académicos la ingeniería ambiental, se convierte en un compromiso del alma máter desarrollar proyectos de esta índole. Contribuyendo con el mejoramiento de la institución e incentivando la participación de la comunidad estudiantil, especialmente los miembros de la Facultad de las ciencias agrarias y del ambiente.

En un proceso de articulación e interdisciplinariedad, el programa de ingeniería de sistemas con el objetivo de impactar en el sector agropecuario y ambiental de la región viene desarrollando en el jardín botánico que permitan dar soporte a todos los procesos llevados a cabo en los laboratorios basados en internet de las cosas (IoT) y procesos de virtualización y así acercarlos al concepto de inteligente (smart). Se proyecta desarrollar infraestructuras bajo el protocolo IPV6 (Rico-Bautista, Medina-Cárdenas, & Santos Jaimes, 2008; Santos & Rico-bautista, 2007), que permitan implementar dispositivos IoT generando dentro de la universidad entornos inteligentes de forma segura logrando integrarlas a la red actual de la universidad, buscando la interoperabilidad de los dos protocolos IPV4 e IPV6 (Lobo & Rico-Bautista, 2012; Rico-Bautista & Alvernia-Acevedo, 2017; Rico-Bautista, Medina-Cárdenas, & Rojas-Osorio, 2016), y dando soporte a todos los procesos de seguridad bajo el esquema de redes definidas por software.

El artículo se presenta en cuatro (4) partes que son: (i) Metodología, (ii) Conceptualización, (iii)Estructuración y discusión acerca de los resultados obtenidos hasta el momento, (iv) Definición de conclusiones acerca del objeto de estudio.

2. Metodología

El tipo de investigación utilizada para el desarrollo es la descriptiva, ya que esta consiste en llegar a conocer las situaciones, costumbres y actitudes predominantes a través de la descripción exacta de las actividades, objetos, procesos y personas. Su meta no se limita a la recolección de datos, sino a la predicción e identificación de las relaciones que existen entre dos o más variables. Los investigadores no son solos tabuladores, sino que recogen los datos sobre la base de una hipótesis o teoría, exponen y resumen la información de manera cuidadosa y luego analizan minuciosamente los resultados, a fin de extraer generalizaciones significativas que contribuyan al conocimiento. (Morales, s.f.). En cuanto al enfoque se toma al cuantitativo, ya que con este se tiene una idea, luego se explora y se plantea el problema haciendo una revisión de la literatura, diseñando la investigación, luego se hace un trabajo de campo y por último se genera un reporte, el cual permite generar estadísticas a nivel del software para el jardín botánico.

Para abarcar diferentes perspectivas el desarrollo del proyecto implementará dos metodologías ágiles, Scrum y basada en prototipos (Sirkin& Ju, 2014), que permitirán desarrollarlo, para determinar actividades que tendrán asignados tiempos de cumplimiento y grupos de trabajo que estén a cargo de desarrollarlas (Rising&Janoff, 2000) (Sievers, Pires, &Brusius, 2015). La asignación de tareas permitirá, a través de la puesta en común en el transcurso del desarrollo de las actividades para generar nuevas ideas (Schwaber, 1997), que el grupo de trabajo entregue mejoras del prototipo, cada revisión generará un producto potencialmente entregable.

Las fases que se estipulan en este proyecto son:

- Construcción del estado de arte de los modelos, tecnologías, análisis de términos y conceptos desde el mundo occidental y oriental sobre jardines botánicos inteligentes, nuevas tendencias de los jardines botánicos a nivel nacional e internacional.
- Definición de los requerimientos funcionales y no funcionales que permitan la virtualización del jardín botánico Jorge Enrique Quintero Arenas. (Recopilación de toda la información de biodiversidad con que cuenta el jardín botánico: fotografías, documentos, información relevante).
- Diseño del esquema interactivo (álbum digital) del jardín botánico basado en herramientas tecnológicas como realidad aumentada y realidad virtual, de las diferentes plantas y animales que se encuentren en el jardín botánico.
- Desarrollo de una aplicación móvil para tener un control y seguimiento de las plantas ubicadas en el jardín botánico.

3. Conceptualización

3.1. Internet de las cosas base de un entorno *Smart*.

Internet de las cosas o IoT (Internet of Things), se refiere a la conexión de objetos tecnológicos o que sean electrónicos a Internet, este concepto se deriva del avance de la tecnología y a la necesidad de compartir y controlar las cosas que nos rodean, incluyendo las necesidades de árboles y plantas (Li, Xu, & Zhao, 2015), (Luvisi &

Lorenzini, 2014). Algunos de los campos en los que IoT se presenta fuertemente son: la adopción generalizada de redes basadas en el protocolo IP (Airehrour, Gutierrez, & Ray, 2016), la economía en la capacidad de cómputo, la miniaturización, los avances en el análisis de datos (Danieletto, Bui, & Zorzi, 2013) y el surgimiento de la computación en la nube, salud, herramientas de aprendizaje, seguridad (Flauzac, Gonzalez, & Nolot, 2015), optimización de procesos (Aziz, 2016), agricultura (Stočes, Vaněk, Masner, & Pavlík, 2016), entre otros.

IoT posee una arquitectura macro de tres niveles o capas (Hurlburt, Voas, & Miller, 2012). En el nivel básico despliega diferentes sensores, cada uno de ellos es una fuente de información, y diferentes tipos de sensores capturan diversos contenidos y formatos de información. El nivel intermedio integra varias redes cableadas e inalámbricas para transferir la información de las cosas con precisión. El nivel más superior es de la aplicación que consta de tres capas (capa de tecnología, capa de middleware, capa de aplicación) y se compone de aplicaciones que exportan toda la funcionalidad de los sistemas para el usuario final (Rodríguez Molano et al., 2015).Internet de las cosas es un tema emergente de importancia técnica, social y económica (Whitmore, Agarwal, & Da Xu, 2015). Tener sensores u objetos dispersos que generan información desde cualquier lugar, requiere la interconexión de estos objetos heterogéneos a través de Internet (Jara, Moreno-Sanchez, Skarmeta, Varakliotis, & Kirstein, 2013). Cobran importancia los Objetos inteligentes (*Smart objects*), objetos físicos con un sistema embebido que le permite procesar información y comunicarse con otros dispositivos y realizar acciones con base en una acción o evento determinado (Ghaleb, Subramaniam, Zukarnain, & Muhammed, 2016).

No obstante, todos estos sistemas complejos presentan retos tales como: 1) la interoperabilidad y los estándares debido a las diferencias entre software y hardware utilizado por cada uno en los diferentes procesos; 2) el análisis avanzado de enormes cantidades de datos producidos por la red de sensores, que requieren de un tratamiento que permita su posterior análisis (Danieletto et al., 2013); 3) la seguridad entendido como los ataques a dispositivos conectados a Internet, el temor a la vigilancia y las preocupaciones relacionadas con la privacidad (Jeon & Jeong, 2016) y 4) aspectos relacionadas con las economías emergentes (Bleda, Jara, Maestre, Santa, & Gómez Skarmeta, 2012; Byun et al., 2016).Estas cifras ponen en manifiesto una preocupación a nivel de seguridad ya que muchos fabricantes de hardware y software de dispositivos tales como teléfonos móviles, sensores, cámaras de vigilancia, GPS, consolas de videojuegos, neveras, televisores, relojes; cada uno de ellos corriendo aplicaciones como redes sociales, georeferenciación, sistemas transaccionales, alertas, y demás; cada una de dichas aplicaciones genera datos en una amplia gama de formatos (texto, gráficos, imágenes, videos, sonidos, etc) los cuales pueden llegar a ser interceptados o robados por agentes maliciosos quienes explotan las vulnerabilidades valiéndose de las debilidades de fabricación tanto en hardware como en software y en la forma como estos dispositivos se comunican con ellos mismos y con el exterior (Katz. Matías David, 2013).

Un informe de la *Hewlett Packart Enterprise* sobre IoT revela cifras importantes a tener en cuenta ya que en la implementación de un sistema IoT muchas organizaciones pueden estar expuestas a pérdida de información o crear agujeros de seguridad por las características débiles que presentan los dispositivos donde el 70% no cifran sus

comunicaciones, las cuentas de usuarios válidas pueden ser identificadas en un 70%, los dispositivos que por facilidad para los usuarios finales en su administración se les agrega interfaz gráfica, el 60% presentan vulnerabilidades a distintos ataques como secuencias de comandos en sitios cruzados (XSS) (Rico-Bautista, Parra-Valencia, & Guerrero, 2017), (Sánchez-Torres, Rodríguez-Rodríguez, Rico-Bautista, & Guerrero, 2018).

3.2. Modelo universidad inteligente: Eje *Smart environment*.

El concepto de *Smart university*(Rico-Bautista et al., 2017), al igual que el de *smart city*, está fuertemente anclado al de la potencia de las TI, haciendo especial énfasis en el concepto del "internet de las cosas" (*Internet of things*) y de las actuales *smartnetworks*, las cuales interconectan sistemas y personas y, además, estimulan la innovación para facilitar un conjunto de objetivos en beneficio de todos.

Partiendo de lo expresado por Maestre (Góngora, 2016) que:

> *"Una ciudad inteligente es un territorio caracterizado por el uso intensivo de las tecnologías, principalmente de información y comunicación, para promover la colaboración, la innovación y la eficiencia para lograr el mejoramiento de la calidad de vida de los ciudadanos y la sostenibilidad de las ciudades a través del mejoramiento de los servicios hacia los ciudadanos" (p.6).*

Ahora desde este concepto de *Smart city*, trasladado a toda su comunidad y a su entorno, Maciá, Berná, Sánchez, Lozano, & Fuster (MACIÁ, 2017) generan como resultado el concepto de *Smart university*, el cual mantiene el mismo objetivo principal de mejorar la calidad de vida de su comunidad aplicando de forma global, intensiva y sostenible las TI bajo el principio de servicio a la ciudadanía. Es por eso por lo que se trata de un entorno controlado, en el cual las políticas económicas, de sostenibilidad, construcción o desarrollo deciden de forma local y no dependen de factores externos, aunque siguen criterios similares a las políticas y estrategias nacionales o internacionales. Para lograr todo es necesario apoyarse en el poder de las TI que existen actualmente, ya que su correcto uso será el pilar fundamental para el éxito del proyecto. La infraestructura actual permite disponer de sistemas digitales conectados de manera global, capaces de soportar aplicaciones, servicios (Y. C. Medina-Cárdenas, Areniz-Arévalo, & Rico-Bautista, 2016; Y. C. Medina-Cárdenas & Rico-Bautista, 2009; Y. Medina-Cárdenas & Rico-Bautista, 2012; Medina & Rico-Bautista, 2008), plataformas y sensores, todos ellos funcionando en conjunto bajo un ecosistema o marco estandarizado de normas y protocolos capaces de actuar de manera integral y a su vez desacoplada, desempeñando un papel imprescindible en la generación correcta de conocimiento e interconexión.

Las principales ventajas de emplear TI son:

- Fácil acceso a los datos
- Adquisición de información
- Procesamiento rápido de la información
- Transmisión rápida
- Almacenamiento y análisis de la información
- Interactividad

Todos los paradigmas de las tecnologías emergentes proporcionarán una serie de características a cada una de las capas definidas en un modelo arquitectural genérico *Smart network* y proporcionando servicios a distintos niveles. Se puede dividir estos modelos en dos capas y a su vez, en una serie de niveles que proporcionan una división de responsabilidades. De los ejes básicos de una *Smart university*, anteriormente citados, nos centramos en *Smart* environment (Entorno y eficiencia). Este eje se centra en el uso de las tecnologías verdes *Green IT*, para desarrollar entornos inteligentes, donde sean capaces de optimizar los recursos naturales, preservar y proteger el medio ambiente.

4. Resultados y discusión

4.1. Entorno inteligente jardín botánico

El proceso de "smartificación", tomado no como un objetivo, sino como una forma de vida, un medio, un proceso continuo de mejora, supone un cambio de paradigma o modelo, mediante el cual se pretende afianzar el concepto de una universidad más abierta, y como consecuencia adaptar su modelo de gestión a los nuevos tiempos rediseñando sus relaciones entre los sectores público y privado, la relación con la comunidad universitaria, sus sinergias y sus ejes transversales donde todos los actores y sus infraestructuras deben coordinarse para un fin común: la sostenibilidad y calidad de vida (Ali & Majeed, 2018; Green, 2014; Hipwell, 2014; MACIÁ, 2017; Staskeviciute & Neverauskas, 2008). Por tal razón es de vital importancia que cualquier universidad pueda y deba constituir, un punto de encuentro entre los distintos agentes económicos públicos y privados, propiciando un entorno donde las empresas(Arévalo-Ascanio, Bayona-Trillos, & Rico-Bautista, 2015), puedan debatir sobre distintas opciones, recibir apoyo científico y participar en proyectos novedosos

La visión de Internet de las Cosas (IoT) prevé un futuro Internet que incorpore objetos físicos que ofrecen funcionalidad alojada como servicios IoT. Para facilitar el acceso sin fisuras y la gestión del ciclo de vida útil de los recursos IoT distribuidos y heterogéneos,

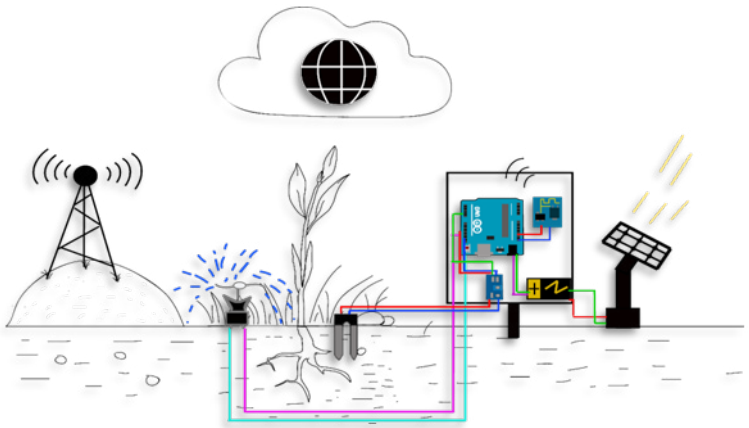


Figura 1 – Control de humedad de las plantas.

la informática orientada al servicio y los enfoques orientados a los recursos han sido ampliamente utilizados como tecnologías prometedoras. Sin embargo, una arquitectura de referencia que integra los servicios de IoT en cualquiera de estas dos tecnologías sigue siendo un desafío de investigación abierto (Dar, Taherkordi, Baraki, Eliassen, & Geihs, 2015). Se contará con sensores de humedad para el monitoreo constante del área donde se encuentren las plantas intervenidas quienes estarán enviando información constante sobre el estado del terreno a un dispositivo central soportado por tecnología Arduino quien será el encargado de enviar información a la nube alertando sobre la necesidad de riego al administrador del jardín botánico. La alimentación eléctrica para estos dispositivos se realizará con la implementación de paneles solares (Basuvaiyan & Rathinasabapathy, 2017) quienes alimentarán los sensores y los Arduinos generando esquema ambiental ecológico. (ver figura 1)

En el jardín botánico, existen muchos factores que comprometen el crecimiento de las plantas, uno de ellos es la humedad, el cual afecta de gran manera el proceso de transpiración y la planta podría marchitarse si la humedad es demasiada alta o demasiada baja, dependiendo del tipo de planta. Lo que se desea es automatizar el proceso de obtención de datos debido a que la lejanía de la ubicación de las plantas impide un control presencial sobre las mismas. Nos permitirá monitorear a través de una aplicación móvil todo lo relacionado con la hoja de vida de las plantas que son objeto de estudio por parte del administrador del jardín. La aplicación tendrá como función lograr en tiempo real la ubicación de las plantas bajo el sistema de Georreferenciación, (Albarracin Caro, 2017) adicionalmente se permitirá el registro de las diferentes intervenciones realizadas a cada una de ellas, por ejemplo: control de enfermedades, progreso en crecimiento, abonado, etc. Cada planta estará identificada con un código de barras bidimensional cuadrada conocido como “QR”, el cual permitirá al administrador realizar la lectura del código mostrándole una interfaz amigable donde se podrá observar la información de la planta registrada en el sistema y realizar registros en línea sobre características detectadas en campo logrando el almacenamiento en la nube para su respectiva gestión, permitiéndole a los interesados tener disponible información sobre un sistema web para la toma de decisiones, (ver Figura 2).

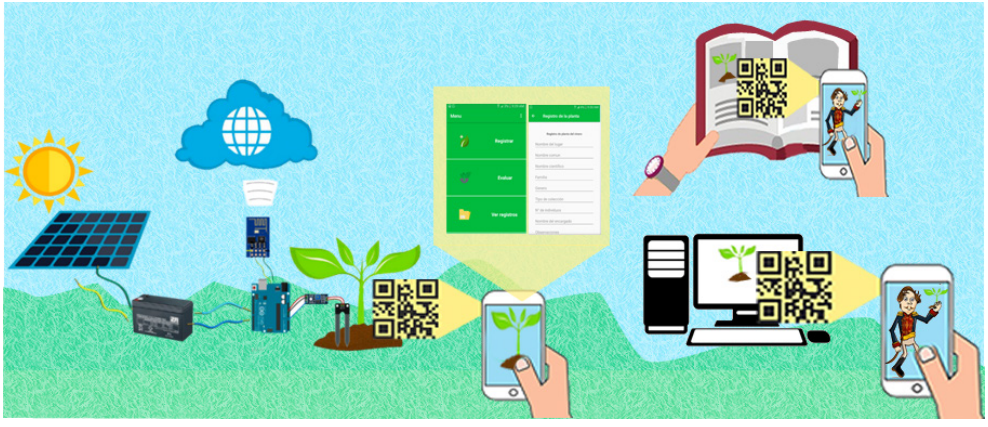


Figura 2 – Utilización de código QR, en la planta y álbum Digital.

El código QR es el marcador que ayudara en el proceso de dar soporte a la realidad aumentada, donde los visitantes del jardín podrán acercar su dispositivo y enseguida parecerá un video con nuestro personaje *santanderito* explicando las características de cada planta de una manera interactiva, así mismo las personas interesadas que no puedan estar físicamente en el jardín, podrán acceder al álbum digital desde cualquier lugar donde se encuentre con acceso a internet y podrá obtener toda la información relevante.

4.2. Fase III

Teniendo en cuenta que el jardín botánico, está en un proceso de socialización y divulgación de la biodiversidad con que cuenta, a estudiantes, docentes, administrativos y visitantes, se busca que esta información no sólo se quede en aquellos que físicamente pueden hacer el recorrido por el mismo, sino que desde cualquier lugar así sea de forma remota, se puedan dar a conocer todas esas maravillas de fauna y flora que lo caracterizan, de una manera didáctica, digital e implementando la realidad aumentada como una estrategia visual más llamativa, nace el proyecto Plant AR.

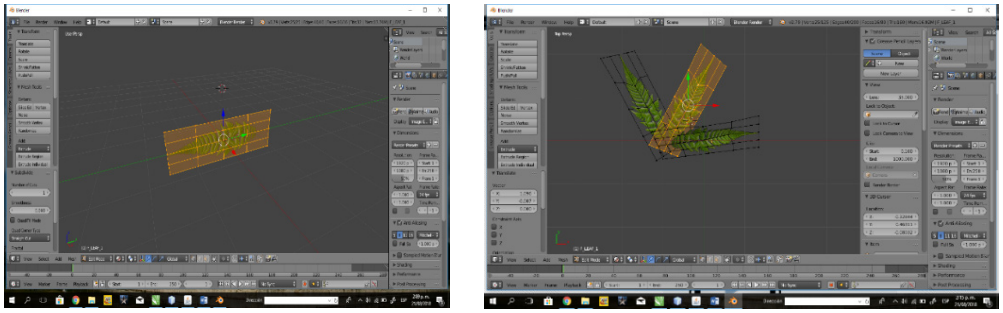


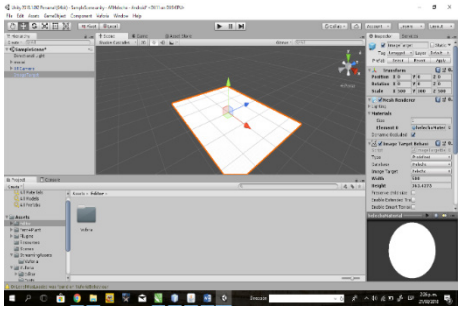
Figura 3 – Construcción hoja de helecho, izquierda. Duplicación hoja de helecho, derecha.

Los programas usados para la realización del álbum digital con realidad aumentada, son los siguientes: Blender, Vuforia, Unity y Android Studio. Proceso de creación: a continuación, se da una breve explicación de la forma como se crea cada una de la especies a través de los software mencionados y que le dan vida al álbum digital. Se utiliza una imagen 4k de una hoja de helecho para convertirse en un gráfico 3D, ajustando el alto, ancho y la longitud de la misma. Luego se duplica la imagen para dar forma al helecho, y de igual manera seguir ajustando la planta, (ver Figura 3).

Luego, se incluye Vuforia al documento Unity, instalando el archivo de Vuforia desde el documento Ajustando las propiedades y organizando el orden en que serán reconocidas las imágenes. Posteriormente, sobre la imagen, se posiciona el gráfico, en medio de la imagen, y se ajusta el ancho, alto y la longitud que tendrá, (ver Figura 4).

Se debe instalar Android Studio, además instalar correctamente SDK y JDK para el buen funcionamiento, luego, se procede a organizar las propiedades y el nombre que tendrá la App, finalmente, se presiona en correr y comienza la creación de la App. Cuando se disponga del archivo APK, se instala en cualquier dispositivo, que tenga una versión 5.0

de Android. Es así como se van diseñando cada una de las diferentes especies de fauna y flora con que contará el álbum digital, (ver Figura 5).



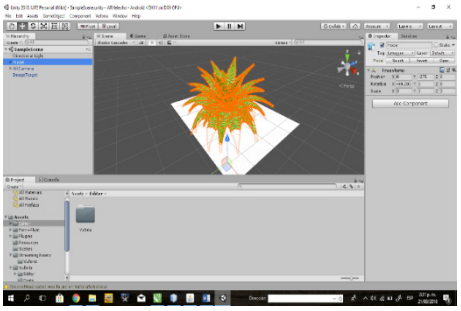
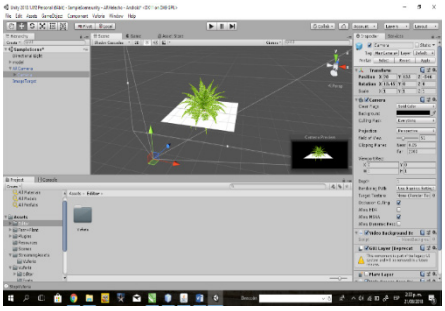
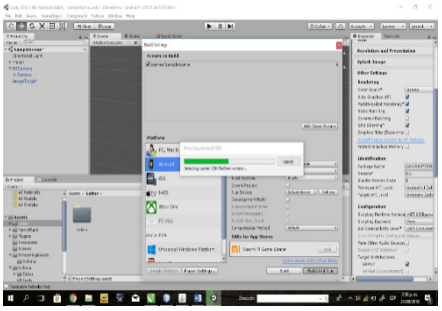


Figura 4 – Ajuste del ancho, alto y longitud, izquierda. Ubicación del helecho en plano 3D, derecha. Fuente: Autores





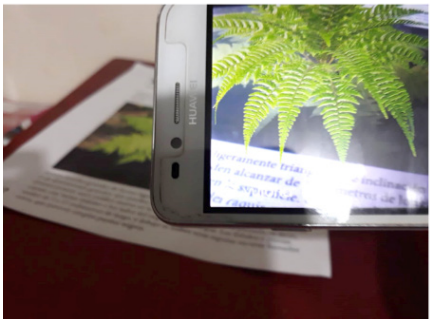


Figura 5 – Vista cámara previa, izquierda. Instalación Android Studio, centro. Prueba piloto realidad aumentada helecho con dispositivo móvil, derecha.

4.3. Fase IV

En el proceso de desarrollo de la App se analizó los diferentes formatos que se llevan de forma manual por parte del administrador del jardín botánico con el objetivo de identificar las características relevantes de cada una de las plantas, objeto de estudio por los investigadores, (ver Figura 6).

Una vez se tiene el análisis se procede a la elaboración de la aplicación móvil logrando desarrollar una aplicación con características de usabilidad y ergonomía bastante simple de manejar que permitiera satisfacer las necesidades de los investigadores. El usuario debe estar registrado para poder adquirir las credenciales necesarias para ingresar al sistema, una vez se autentica se le da acceso al menú opciones donde podrá realizar funcionalidades entre las que se encuentran, registrar una planta, evaluar características detectadas por los investigadores y realizar consultas inteligentes con el objetivo de tomar decisiones con base a información de interés. En la interfaz de registro de plantas los investigadores se encontrarán con un formulario donde se podrá capturar información de las plantas objeto de estudio presente en los viveros del jardín botánico, (ver Figura 7).

FORMATO DE ACCESIONES
JARDIN BOTANICO JORGE ENRIQUE QUINTERO ARENAS"
UNIVERSIDAD FRANCISCO DE PAULA SANTANDER OCAÑA

N° 0001	Fecha:	Hora:
Nombre del lugar		
Coordenadas de lugar		
Latitud	Longitud	
N° de accesion		
Nombre Comun:		
Nombre cientifico:		
Familia		
Genero		
Tipo de colección:		
N° de individuos:		
Nombre del encargado		
Observaciones:		

Figura 6 – Formato de Accesiones.

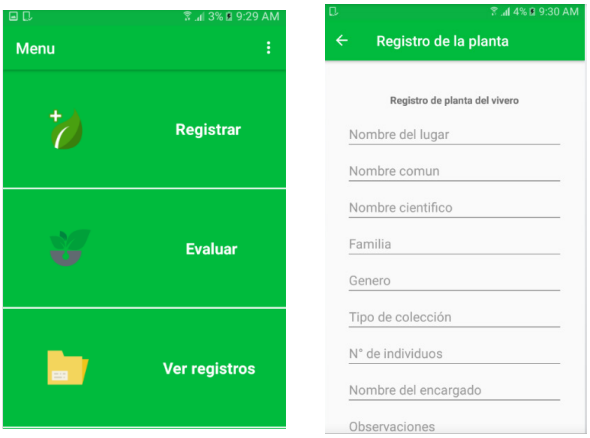


Figura 7 – Menú de opciones, izquierda. Registro de planta, derecha.

Dentro del registro de cada planta se podrá identificar cada una de ellas con un código QR que permitirá agilizar el proceso de búsqueda y trazabilidad de las diferentes intervenciones que se estén realizando, objeto de investigaciones como control de plagas, formas, crecimiento, etc. Una vez las plantas están identificadas en su

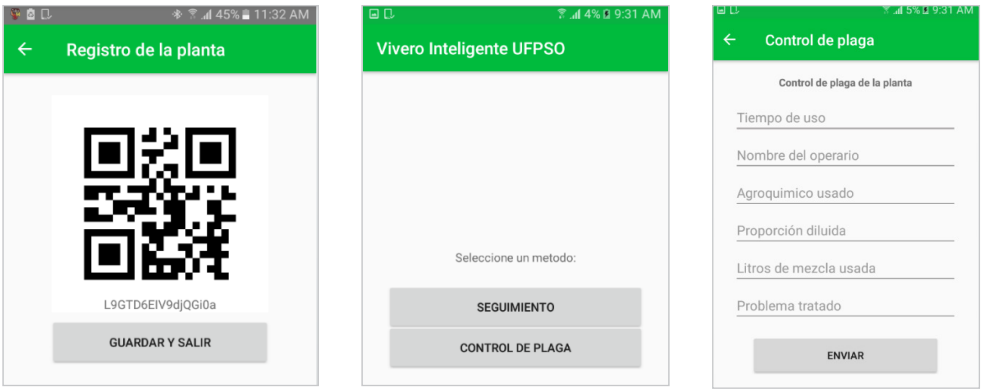


Figura 8 – Identificación código QR, izquierda. Opciones, centro. Control de plagas, derecha.

ubicación y características, el administrador del jardín botánico inteligente podrá de forma muy rápida realizar trabajos del día a día como control de plagas, seguimientos a muertes, trasplantes, donaciones que se pueden presentar. Todas estas interfaces permitirán al administrador e investigadores del jardín botánico inteligente obtener información en tiempo real bajo reportes de interés para la toma de decisiones, (ver Figura 8).

5. Conclusiones

Esta aplicación es una herramienta que permite dar soporte de forma rápida y ágil a los diferentes procesos que actualmente se llevan a cabo dentro del jardín botánico, tanto a nivel investigativo como de administración del mismo.

La implementación del código QR para identificar cada planta objeto de estudio es un plus que permite dar soporte a otros proyectos en curso como lo es el software de realidad aumentada para el recorrido del jardín Botánico por parte de visitantes ya que los QR servirán de marcadores para dar a conocer las características de plantas nativas como orquídeas, heliconias, árboles que embellecen el ambiente natural.

La toma de decisiones en tiempo real es fundamental para el jardín botánico, por tal razón dicha aplicación será de gran ayuda por la facilidad de operación de administradores e investigadores.

Con el desarrollo del álbum digital con realidad aumentada se propician espacios para la educación ambiental, así mismo se podrá identificar y conocer la riqueza a través de colecciones presentes de fauna y flora, a fin de generar conocimiento y valoración de la biodiversidad en todos sus componentes a nivel eco sistémico.

Mediante el álbum digital los estudiantes de diferentes carreras, visitantes y comunidad general podrán apreciar y educarse, a su vez conocer temas importantes del jardín botánico adquiriendo un sentido de pertenencia y obteniendo una sensibilización ambiental necesaria para hacer frente a los retos emergentes en el escenario global actual.

Agradecimientos

A la UFPS Ocaña, mediante la División de Investigación y Extensión (DIE) vincula a docentes, administrativos y estudiantes para que participen en la ejecución y desarrollo de proyectos de investigación.

Referencias

Agostini, A., Alenyà, G., Fischbach, A., Scharr, H., Wörgötter, F., & Torras, C. (2017). A cognitive architecture for automatic gardening. Computers and Electronics in Agriculture, 138, 69–79. https://doi.org/10.1016/j.compag.2017.04.015

Airehrour, D., Gutierrez, J., & Ray, S. K. (2016). Secure routing for internet of things: A survey. Journal of Network and Computer Applications, 66. https://doi.org/10.1016/j.jnca.2016.03.006

Albarracin Caro, R. (2017). Revisión y georreferenciación de registros Biológicos de Plantas Medicinales y Alimenticias Nativas.

Ali, M., & Majeed, A. (2018). How Internet-of-Things (IoT) Making the University Campuses Smart ?, 646–648. https://doi.org/10.1109/CCWC.2018.8301774

Arévalo-Ascanio, J. G., Bayona-Trillos, R. A., & Rico-Bautista, D. W. (2015). Responsabilidad social empresarial e innovación: Una mirada desde las tecnologías de la información y comunicación en organizaciones. Clío América, 9(18), 180–189.

Aziz, B. (2016). A formal model and analysis of an IoT protocol. Ad Hoc Networks, 36, 49–57. https://doi.org/10.1016/J.ADHOC.2015.05.013

Barkoosaraei, M. H. D. A., & Moshiri, S. (2017). Designing of Genow Botanical Garden with Sustainable Architecture Approach. TARIH KULTUR VE SANAT ARASTIRMALARI DERGISI-JOURNAL OF HISTORY CULTURE AND ART RESEARCH, 6(4), 1211–1226. https://doi.org/10.7596/taksad.v6i4.1140

Basuvaiyan, S., & Rathinasabapathy, V. (2017). IoT based Solar Photo Voltaic Monitoring System.

Biswal, E. V., Singh, E. H. M., Jeberson, W., & Dhar, E. A. S. (2015). Greeves : A Smart Houseplant Watering and Monitoring System, 4(7), 2499–2507.

Bleda, A. L., Jara, A. J., Maestre, R., Santa, G., & Gómez Skarmeta, A. F. (2012). Evaluation of the impact of furniture on communications performance for ubiquitous deployment of wireless sensor networks in smart homes. Sensors (Switzerland), 12(5), 6463–6496. https://doi.org/10.3390/s120506463

Byun, J., Kim, S., Sa, J., Kim, S., Shin, Y. T., & Kim, J. B. (2016). IoT(internet of things) based smart city services for the creative economy. International Journal of Smart Home, 10(7), 185–192. https://doi.org/10.14257/ijsh.2016.10.7.19

Danieletto, M., Bui, N., & Zorzi, M. (2013). RAZOR: A compression and classification solution for the internet of things. Sensors (Switzerland), 14(1), 68–94. https://doi.org/10.3390/s140100068

Dar, K., Taherkordi, A., Baraki, H., Eliassen, F., & Geihs, K. (2015). A resource oriented integration architecture for the Internet of Things: A business process perspective. Pervasive and Mobile Computing. https://doi.org/10.1016/j.pmcj.2014.11.005

Flauzac, O., Gonzalez, C., & Nolot, F. (2015). New security architecture for IoT network. In Procedia Computer Science (Vol. 52, pp. 1028–1033). Elsevier Masson SAS. https://doi.org/10.1016/j.procs.2015.05.099

Ghaleb, S. M., Subramaniam, S., Zukarnain, Z. A., & Muhammed, A. (2016). Mobility management for IoT: a survey. Eurasip Journal on Wireless Communications and Networking. https://doi.org/10.1186/s13638-016-0659-4

Góngora, G. P. M. (2016). Revisión de literatura sobre ciudades inteligentes: una perspectiva centrada en las TIC. Ingeniare, 19(19), 137–149. Retrieved from http://www.unilibrebaq.edu.co/ojsinvestigacion/index.php/ingeniare/article/viewFile/705/639%0Ahttp://www.unilibrebaq.edu.co/ojsinvestigacion/index.php/ingeniare/article/view/705

Green, J. (2014). The Internet of Things Reference Model. Internet of Things World Forum, 1–12.

Hipwell, S. (2014). Developing smart campuses #x2014; A working model. 2014 International Conference on Intelligent Green Building and Smart Grid (IGBSG), 1–6. https://doi.org/10.1109/IGBSG.2014.6835169

Hurlburt, G. F., Voas, J., & Miller, K. W. (2012). The Internet of Things: A Reality Check. IT Professional, 14(June), 56–59. https://doi.org/10.1109/MITP.2012.60

Jara, A. J., Moreno-Sanchez, P., Skarmeta, A. F., Varakliotis, S., & Kirstein, P. (2013). IPv6 addressing proxy: Mapping native addressing from legacy technologies and devices to the internet of things (IPv6). Sensors (Switzerland), 13(5), 6687–6712. https://doi.org/10.3390/s130506687

Jeon, J., & Jeong, S.-R. (2016). Designing a Crime-Prevention System by Converging Big Data and IoT. Journal of Korean Society for Internet Information, 17(3), 115–128. https://doi.org/10.7472/jksii.2016.17.3.115

Katz. Matías David. (2013). Redes y seguridad. Alfaomega Grupo Editor, (Mexico), 87.

Li, S., Xu, L. Da, & Zhao, S. (2015). The internet of things: a survey. Information Systems Frontiers, 17(2), 243–259. https://doi.org/10.1007/s10796-014-9492-7

Lobo, J., & Rico-Bautista, D. (2012). IMPLEMENTACIÓN DE LA SEGURIDAD DEL PROTOCOLO DE INTERNET VERSIÓN 6. Gerencia Tecnológica Informática, 11(29), 35–46. Retrieved from https://biblat.unam.mx/es/revista/gerencia-tecnologica-informatica/articulo/implementacion-de-la-seguridad-del-protocolo-de-internet-version-6

Luvisi, A., & Lorenzini, G. (2014). RFID-plants in the smart city: Applications and outlook for urban green management. Urban Forestry and Urban Greening, 13(4), 630–637. https://doi.org/10.1016/j.ufug.2014.07.003

MACIÁ, F. (2017). SMART UNIVERSITY. Hacia una universidad más abierta. (Alfaomega & Marcombo, Eds.) (Primera).

Medina-Cárdenas, Y. C., Areniz-Arévalo, Y., & Rico-Bautista, D. (2016). Modelo estratégico para la gestión tecnológica en la organización: plan táctico de la calidad (ITIL & ISO 20000). (Instituto Tecnológico Metropolitano, Ed.). Fondo Editorial ITM. https://doi.org/https://doi.org/10.22430/9789585414006

Medina-Cárdenas, Y. C., & Rico-Bautista, D. (2009). Modelo de gestión basado en el ciclo de vida del servicio de la Biblioteca de Infraestructura de Tecnologías de Información (ITIL). Revista Virtual Universidad Católica Del Norte, (27), 1–21.

Medina-Cárdenas, Y., & Rico-Bautista, D. (2012). Mejores prácticas de gestión para la calidad de los servicios en tecnologías de información. Gerencia Tecnológica Informática.

Medina, Y., & Rico-Bautista, D. (2008). Modelo De Gestión De Servicios Para La Universidad De Pamplona: Itil. Scientia Et Technica, XIV(39), 314–319. Retrieved from http://www.redalyc.org/articulo.oa?id=84920503056

Rico-Bautista, D., & Alvernia-Acevedo, S. A. (2017). ANÁLISIS DE UNA RED EN UN ENTORNO IPV6: UNA MIRADA DESDE LAS INTRUSIONES DE RED Y EL MODELO TCP/IP. Revista Colombiana Tecnologías de Avanzada, 1(29), 81–91. https://doi.org/https://doi.org/10.24054/16927257.v29.n29.2017.2490

Rico-Bautista, D., Medina-Cárdenas, Y., & Rojas-Osorio, J. A. (2016). PENTESTING EMPLEANDO TECNICAS DE ETHICAL HACKING EN REDES IPv6. Revista Ingenio UFPSO, 11(70–96), 1. Retrieved from http://revistas.ufpso.edu.co/index.php/ringenio/article/view/391/223

Rico-Bautista, D., Medina-Cárdenas, Y., & Santos Jaimes, L. M. (2008). Ipsec De Ipv6 En La Universidad De Pamplona. Scientia Et Technica, 2(39), 320–325. https://doi.org/http://dx.doi.org/10.22517/23447214.3239

Rico-Bautista, D., Parra-Valencia, J. A., & Guerrero, C. D. (2017). IOT: UNA APROXIMACION DESDE CIUDAD INTELIGENTE A UNIVERSIDAD INTELIGENTE. Revista Ingenio UFPSO, 13(1), 9–20.

Rodríguez Molano, J. I., Montenegro marín, C. E., Cueva Lovelle, J. M., Molano, J., Marin, C., & Cueva, J. (2015). Introducción al Internet de las Cosas. Redes de Ingeniería, 6(7), 53–59. https://doi.org/10.14483/udistrital.jour.redes.2016.1.a04

Sánchez-Torres, B., Rodríguez-Rodríguez, J. A., Rico-Bautista, D., & Guerrero, C. D. (2018). Smart Campus: Trends in cybersecurity and future development. Revista Facultad de Ingeniería, 27(47). https://doi.org/10.19053/01211129.v27.n47.2018.7808

Santos, L. M., & Rico-bautista, D. (2007). IPV6 en la Universidad de Pamplona: Estado del arte. Scientia Et Technica, XIII, 415–420. Retrieved from http://www.redalyc.org/articulo.oa?id=84903771

Staskeviciute, I., & Neverauskas, B. (2008). The Intelligent University's Conceptual Model. Inzinerine Ekonomika-Engineering Economics, (4), 53–58.

Stočes, M., Vaněk, J., Masner, J., & Pavlík, J. (2016). Internet of Things (IoT) in Agriculture - Selected Aspects. Agris On-Line Papers in Economics and Informatics, VIII(1), 83–88. https://doi.org/10.7160/aol.2016.080108

Walter, K. S. (1991). Computerized plant record systems for botanic gardens. Tropical Botanic Gardens. ACADEMIC PRESS LIMITED. https://doi.org/10.1016/B978-0-12-346850-5.50035-1

Wassenberg, C. L., Goldenberg, M. A., & Soule, K. E. (2015). Benefits of botanical garden visitation: A means-end study. Urban Forestry and Urban Greening, 14(1), 148–155. https://doi.org/10.1016/j.ufug.2015.01.002

Whitmore, A., Agarwal, A., & Da Xu, L. (2015). The Internet of Things—A survey of topics and trends. Information Systems Frontiers, 17(2), 261–274. https://doi.org/10.1007/s10796-014-9489-2

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 12/10/2018
Aceitação/Acceptance: 10/12/2018

Relación entre Elementos y Ventajas de la Cartera Estratégica de Proyectos de TI. Caso de Estudio: Universidades Públicas Españolas

Francisco Valverde-Alulema[1], Faraón Llorens-Largo[2]

fvalverde@uce.edu.ec, faraon.llorens@ua.es.

[1] Universidad Central del Ecuador, Quito, Ecuador.

[2] Universidad de Alicante, Alicante, España.

Pages: 283–298

Resumen: La Cartera de Proyectos de TI es una buena práctica alineada con la gobernanza de las TI en las organizaciones, y en particular en las Universidades. El objetivo de esta investigación es determinar un conjunto de elementos en la Cartera de Proyectos de TI para que esta pueda ser considerada estratégica, y las ventajas ocasionadas por la implantación de tales elementos. La propuesta inicial, producto de la revisión de la literatura científica existente, ha sido posteriormente puesta a consideración de un grupo de discusión de expertos en gobernanza de las TI. Finalmente se puso a consideración de los responsables de TI de 19 universidades públicas del sistema universitario español, una encuesta con 21 ventajas causadas por la implementación de 16 elementos de una Cartera Estratégica de Proyectos de TI en las universidades. Además, se determinó la relación causa efecto de los elementos y las ventajas.

Palabras-clave: Gobernanza de las TI, Cartera Estratégica de Proyectos de TI, Ventajas de la Cartera de Proyectos de TI, Elementos de la Cartera de Proyectos de TI, Mejores Prácticas

Relation between Elements and Advantages of IT Projects Strategic Portfolio. Study Case: Spanish Public Universities

Abstract: The IT Project Portfolio is a good practice aligned with governance of IT in organizations and in universities specifically too. Research objective is determine a set of elements the IT Project Portfolio in order to can be considered strategic, and the advantages caused by that it implementation. The initial proposal is result of the literature scientific review and a discussion group of IT Governance experts. Finally, a survey with questions related with the 21 advantages caused for 16 elements implementation of a Strategic Portfolio of IT Projects in universities was put to the consideration of the IT managers of 19 universities of the Spanish university system. In addition, the relation (cause / effect) between elements and advantages were determined

***Keywords**:* IT Governance, Strategic Project Portfolio of IT Projects, Advantages of the IT Project Portfolio, Elements of the IT Project Portfolio, Best Practices

1. Introducción

1.1. La Cartera de Proyectos de TI

De Haes y Van Grembergen (2009) encontraron 7 prácticas que deben ser implantadas para una adecuada aceptación de una cultura de gobernanza de TI: existencia de un comité directivo de TI, gestión de la cartera, control del presupuesto de TI, existencia de un comité directivo del proyecto de TI, el CIO (Chief Information Officer) que informa al CEO (Chief Executive Officer), y metodologías de gestión de proyectos y liderazgo de TI. Rosselet y Wentland (2009) incorporan interdependencias de proyectos relevantes para la selección y priorización. Además, actividades de identificación, evaluación de proyectos y balanceo de cartera surgieron como las fases más importantes del proceso de la gestión de la cartera de proyectos de TI. En qué y cómo se gastan los recursos indica las prioridades estratégicas de una organización. La cartera de proyectos de TI es una poderosa herramienta para la gobernanza de las TI, requiere conexiones estrechas entre principios, procesos, personas y desempeño (Cameron, 2005).

Existen numerosos trabajos con distintos modelos de cartera de proyectos (Gleisberg, Zondag y Chaudron, 2008; Bonham, 2005; Maizlish y Handler, 2005; Medellín, 2006; PMI, 2013). Algunos diseñados específicamente para las universidades (Correa y Benavides, 2013; Correa y Benavides, 2016), que es el ámbito de nuestra investigación. Para Medellín (2006), los objetivos que persigue la cartera de proyectos de TI son: lograr que la cartera esté alineada estratégicamente; obtener el máximo valor; lograr el balance correcto entre los proyectos que se ejecutan; desarrollar el número adecuado de proyectos; lograr satisfacción de los clientes en la medida que los resultados que se obtienen de los proyectos respondan mejor a sus necesidades; y comunicar las prioridades de forma clara y precisa al personal de todas las áreas de la organización.

La gobernanza de las TI debe generar un conjunto priorizado de proyectos de TI bien definidos y que puedan ejecutarse a corto plazo exitosamente. Medellín (2006) plantea un procedimiento para que las inversiones en proyectos de TI se alineen con los objetivos de la empresa. Además, demuestra que la cartera de proyectos de TI es fundamental en la estrategia de negocio y su combinación refleja las prioridades estratégicas. Jairak y Praneetpolgrang (2013) plantean el establecimiento de un procedimiento para priorizar los proyectos e involucrar a los responsables de TI y del negocio desde el inicio del proyecto. Cubeles (2007) afirma que la cartera de proyectos de TI es una práctica que permite aumentar el valor que genera TI y ayuda a la implantación de la estrategia a través de los proyectos.

Con respecto al proceso de selección de los proyectos de TI se plantean varias soluciones. Zhijie (2012) sugiere que la gobernanza de TI desempeña un papel importante en la selección de proyectos de TI y la priorización de la cartera. Gleisberg y otros (2008) recomiendan para la selección de la cartera de proyectos de TI crear uniformidad en la aplicación de criterios, evaluar toda la cartera de proyectos, realizar un análisis de cada proyecto individualmente y procurar un enfoque más sistemático de selección. Karhade

y Shaw (2007) proponen un modelo para la selección de proyectos basado en la inversión total requerida y la fecha en que se espera que el proyecto se inicie, entre otros aspectos. Chiang y Nunez (2009) proponen modelos que pueden ser utilizados por los tomadores de decisiones para proporcionar un soporte integral para la optimización de la cartera.

Con respecto a las propuestas encontradas sobre universidades, Fernández (2010) afirma que una de las mejores prácticas que se pueden aplicar en relación al principio de adquisición es la puesta en marcha de una cartera de proyectos de TI. Sierra (2012) determina que se debe diseñar una cartera de inversiones de TI basada en un conjunto priorizado de proyectos de TI bien definidos, es decir, se ejecute una cartera de proyectos de TI. Finalmente, diversos trabajos afirman que una de las bases del buen gobierno TI consiste en desarrollar una cartera de proyectos alineados con los objetivos de la universidad (Fernández y Llorens, 2011; Fernández, Gumbau y Llorens, 2012; Fernández, Hontoria y Llorens, 2014, y Gómez y otros, 2017). El modelo CPTI4U (Cartera de Proyectos TI para Universidades) (Fernández, Hontoria y Llorens, 2012) tiene como referencia varias metodologías, se alinea a la norma ISO/IEC 38500 y está adaptado a las universidades, de tal manera que usa un lenguaje dirigido tanto a la parte técnica como a la dirección. Este modelo ha sido implantado en 10 universidades españolas con muy buenos resultados (Fernández et al., 2012).

1.2. Objetivo de la Investigación

Como se ha visto en el apartado anterior, la implementación de la cartera de proyectos de TI es una buena práctica de gobernanza, y una herramienta de apoyo estratégico y fundamental en la alineación de las tecnologías de la información con los objetivos institucionales. Para esta investigación, nos hemos centrado en el sector de las universidades públicas españolas, adaptando los trabajos existentes a esta realidad. El objetivo de la investigación es determinar en primera instancia, cuales son las ventajas causadas por la implementación de una cartera de proyectos de TI que pueda ser considerada estratégica, además de determinar qué elementos debe poseer dicha cartera para generar estas ventajas. A partir de la base de elementos (parte tangible de la cartera de proyectos de TI) y ventajas (condiciones favorables y de mejoramiento para las instituciones), se procede a determinar la posible relación causa-efecto entre los elementos y ventajas. De esta forma se logra construir un instrumento que permite determinar la relación entre elementos y ventajas de una Cartera Estratégica de Proyectos de TI (CEPTIU). Posteriormente se establece el estado situacional de la CEPTIU en las universidades, tomando como caso de estudio al Sistema Universitario Español (SUE), además, servirá de línea base sólida, que permita iniciar su implementación o mejorar la existente.

2. Diseño de la investigación

Según afirman Hernández, Fernández y Baptista (2014) los problemas a los que se enfrentan actualmente las organizaciones son muy complejos y diversos. Por ello, el uso de un enfoque de investigación único, ya sea cuantitativo o cualitativo, es insuficiente para recoger esta complejidad. Así, plantean la introducción de métodos mixtos (Hernández et al., 2014). La combinación mixta ocurre cuando se integra el análisis cualitativo de los datos y la recolección de datos cuantitativos.

Esta investigación sigue el Diseño Exploratorio Secuencial (DEXPLOS), con una primera fase de recopilación y análisis de datos cualitativos, seguida de otra en donde se recaban y analizan datos cuantitativos. En concreto se ha utilizado una de las modalidades de este tipo de diseño denominada Derivativa, ya que la recopilación y el análisis de los datos cuantitativos son construidos sobre la base de los resultados previos cualitativos. La interpretación final es producto de la integración y comparación de ambos resultados. Esta modalidad de investigación plantea tres etapas (Hernández et al., 2014) (recopilar y analizar datos cualitativos, usar estos resultados para construir un instrumento cuantitativo y aplicar el instrumento a una muestra de una población para validarlo), que son los que vamos a seguir en esta investigación.

Para la recolección de datos cualitativos se ha utilizado la técnica de grupos de discusión (*focus groups* o grupo de expertos). Son reuniones de grupos pequeños o medianos, en las cuales los participantes conversan en torno a uno o varios temas en un ambiente relajado e informal (blogs, correos, foros), bajo la conducción de un especialista en dinámicas grupales (en nuestro caso los investigadores). Hernández et al. (2014) recomiendan que el grupo este compuesto entre tres y cinco personas cuando se tratan temas complejos y entre seis y diez participantes si las cuestiones a tratar son sobre asuntos más cotidianos, pero no excediéndose en ningún caso de un número manejable de individuos. El formato y naturaleza de la sesión o sesiones depende del objetivo, de las características de los participantes y del planteamiento del problema.

Basándonos en estos planteamientos, en la presente investigación se han seguido tres etapas:

1. Primeramente, se abordó la etapa de análisis cualitativo. A partir de la lectura de las publicaciones científicas sobre el tema, se procedió a construir una propuesta de 14 elementos y 17 ventajas que deben ser consideradas en una CEPTIU. Estos dos listados se pusieron a consideración del grupo de discusión formado por 10 expertos en gobernanza de las TI para su análisis, discusión y validación final. Fruto del debate y discusión se obtuvo un listado final de 16 elementos y 21 ventajas.
2. Seguidamente con los resultados obtenidos por la colaboración del grupo de expertos, se construyó un instrumento cuantitativo en forma de encuesta semiestructurada, con algunas preguntas cuantitativas y otras cualitativas.
3. Como tercer y último paso se aplicó el instrumento a la población objetivo: universidades públicas españolas que tienen implantada una cartera de proyectos de TI. Como sugiere el método DEXPLOS, los participantes de la primera parte cualitativa no fueron los mismos de la fase de validación cuantitativa.

3. Resultados y Discusión

3.1. Elementos y ventajas de la CEPTIU

Con el fin de validar los 14 elementos y 17 ventajas propuestos inicialmente por los investigadores, determinados a partir de una lectura y revisión sistemática de la literatura científica, se constituye un grupo de diez expertos en el área de la gobernanza de las TI en las universidades, todos españoles, pero con reconocimiento internacional.

El perfil de los 10 expertos que discutieron y validaron la propuesta es variado, intentando recoger los distintos roles directivos universitarios relacionados con la TI: vicerrectores, gerentes y directores de TI. Fruto del debate del grupo de discusión se realizan los siguientes cambios:

- Con respecto a la propuesta de elementos de la CEPTIU, se recomienda un orden cronológico del listado, el cambio en la redacción de varios elementos y el incremento de dos nuevos elementos.
- Con respecto a la propuesta de ventajas de la CEPTIU, se determina un ordenamiento de las mismas considerando su importancia estratégica de mayor a menor, se cambia la redacción de varias de ellas y se añaden cuatro nuevas ventajas a la propuesta.

Como resultado de esta primera etapa se construye el listado final de 16 elementos (Tabla 1) y 21 ventajas (Tabla 2) de la CEPTIU.

Elemento	Descripción del elemento
E01	Incluye criterios que son públicos y conocidos antes de la convocatoria.
E02	Tiene en cuenta los objetivos estratégicos de la universidad.
E03	Tiene en cuenta las necesidades relacionadas con el mapa de procesos de la universidad y los factores críticos de éxito.
E04	Tiene en cuenta si los proyectos cumplen con las leyes y normativas tanto internas como externas (responsabilidad social, transparencia de la información, códigos de ética y vinculación con la sociedad) de la universidad.
E05	Tiene un procedimiento que relaciona las interdependencias entre proyectos.
E06	Dispone de indicadores para medir el impacto estratégico o valor que aportan los proyectos a la universidad una vez que se han finalizado.
E07	Mide los riesgos que suponen los proyectos para la universidad.
E08	Tiene en cuenta el coste íntegro (presupuesto y recursos propios) requerido para abordar los proyectos.
E09	Incluye criterios objetivos para ayudar al Equipo de Gobierno a seleccionar y priorizar los proyectos.
E10	Tiene definido un procedimiento o flujo de toma de decisiones donde los roles y responsabilidades de todos los que participan están claros y bien definidos.
E11	Tiene un procedimiento para involucrar a los responsables de TI y del negocio desde el principio del proyecto.
E12	Dispone de un procedimiento de comunicación con los grupos de interés implicados en los proyectos.
E13	Dispone de una oficina (grupo de expertos) de apoyo a la gestión de la cartera (tanto al CIO como a los solicitantes de los proyectos).
E14	Hace el seguimiento de los proyectos para establecer su rendimiento y su éxito final.
E15	Dispone de un procedimiento dinámico de re-planificación para ajustarse a nuevas circunstancias.
E16	Tiene en cuenta las personas implicadas en un proyecto y determina cómo deben contribuir a su éxito.

Tabla 1 – Elementos de la cartera estratégica de proyectos de TI

Ventaja	Descripción de la ventaja
V01	Contribuir a que los proyectos se alineen con los objetivos estratégicos de la universidad.
V02	Conseguir establecer el impacto estratégico o valor que aportan los proyectos a la universidad.
V03	Contribuir a que el Equipo de Gobierno se implique en la selección y priorización de los proyectos.
V04	Contribuir a gestionar y minimizar los riesgos que suponen los proyectos para la universidad.
V05	Facilitar la transición entre equipos de gobierno, en materia de proyectos TI.
V06	Promover la comunicación con los grupos de interés implicados en los proyectos.
V07	Promover el seguimiento de los proyectos para establecer su rendimiento y su éxito final.
V08	Contribuir a establecer la importancia estratégica de las TI para la universidad.
V09	Permitir equilibrar (balancear) las inversiones de la universidad.
V10	Permitir definir las responsabilidades y el flujo de decisión para poder incorporar proyectos a la cartera.
V11	Ayudar a identificar las sinergias entre proyectos TI.
V12	Contribuir a democratizar la solicitud de proyectos TI abriéndolos a la comunidad universitaria.
V13	Facilitar la planificación de los proyectos a medio y largo plazo.
V14	Contribuir a identificar todas las necesidades relacionadas con los procesos de la universidad.
V15	Contribuir a que los proyectos cumplan con las leyes y normativas tanto internas como externas (responsabilidad social, transparencia de la información, códigos de ética y vinculación con la sociedad) de la universidad.
V16	Contribuir involucrando a los responsables de TI y del negocio desde el principio del proyecto.
V17	Asegurar que la planificación estratégica se implemente, se mantenga y se pueda llevar a cabo con éxito.
V18	Favorecer que el Equipo de Gobierno decida sobre la continuidad, modificación o cancelación de un proyecto.
V19	Facilitar la gestión eficiente de los recursos humanos y financieros asignados a los proyectos (evitando sobrecargas).
V20	Permitir evitar la aparición de infraestructuras TI o sistemas de información aislados.
V21	Promover el control centralizado y facilita la elaboración del presupuesto a corto y medio plazo.

Tabla 2 – Ventajas de la cartera estratégica de proyectos de TI

Complementariamente, los investigadores con el grupo de discusión han determinado la relación (causa/efecto) entre los elementos y las ventajas de la cartera estratégica de proyectos de TI (Figura 1).

3.2. Análisis de los resultados de la encuesta

Según el informe UNIVERSITIC 2017 (Gómez et al., 2017), el 47.83% de las universidades españolas diseña anualmente una cartera de proyectos TI bien definidos

N		VENTAJAS DE LA CEPTIU																				
		V01	V02	V03	V04	V05	V06	V07	V08	V09	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20	V21
ELEMENTOS DE LA CEPTIU	E01						■						■			■						
	E02	■	■						■					■	■			■				
	E03											■		■	■					■	■	
	E04	■	■		■				■							■						
	E05											■							■	■		■
	E06	■	■			■		■	■													
	E07				■			■						■					■			
	E08									■				■						■		■
	E09	■		■						■												■
	E10						■				■			■			■			■		
	E11		■	■					■		■						■					
	E12						■		■													
	E13						■		■		■					■						
	E14				■		■	■	■									■		■		
	E15				■				■										■			■
	E16						■				■									■		

Figura 1 – Matriz de relación causa-efecto entre elementos y ventajas de la CEPTIU.

que es priorizada y aprobada por el equipo de gobierno de la universidad. En este estudio participaron 49 universidades, que conforman el 66.2% de las universidades españolas. Si nos detenemos únicamente en las universidades públicas, en el informe han participado 46 universidades, lo que representa el 86% de las universidades públicas del sistema universitario español. Nuestra investigación ha tomado como población las universidades públicas españolas que tienen implantada una cartera de proyectos de TI. Del informe UNIVERSITIC 2017 se puede estimar que unas 23 universidades públicas españolas tienen implantada la cartera de proyectos de TI. En esta investigación se ha obtenido respuesta de 19 universidades públicas españolas que tienen implantada la cartera.

El instrumento para recogida de la información cuantitativa contiene 21 preguntas, correspondientes a las 21 ventajas, que deben ser respondidas según una escala de Likert: N1 totalmente de acuerdo, N2 de acuerdo, N3 ni de acuerdo ni en desacuerdo, N4 en desacuerdo y N5 totalmente en desacuerdo). Con las respuestas de los 19 responsables de TI de las universidades españolas que tienen implantada una cartera de TI, se genera la tabla 3 con el porcentaje de población que está de acuerdo con la propuesta de cada una de las 21 ventajas de la CEPTIU. Al considerar los niveles de acuerdo N1 (Totalmente de acuerdo) y N2 (De acuerdo), se observa que los responsables de TI de las universidades públicas del SUE, en un alto porcentaje (columna %(N1+N2)) están apoyando la propuesta de las 21 ventajas puestas a consideración por los investigadores y el grupo de expertos en gobernanza de las TI. Los porcentajes de N3 (columna %N3) determinan en menor grado la existencia o no de las ventajas puestas a consideración, siendo el punto medio de elección (ni de acuerdo, ni en desacuerdo). Finalmente, el porcentaje de desacuerdo (columna %(N4+N5)) es muy bajo, por lo que se puede afirmar que la propuesta de las 21 ventajas de la CEPTIU puestas a consideración son reales y se puede percibir su existencia tras una adecuada implementación de la CEPTIU.

Ventaja	N1	%(N1)	N2	%(N2)	%(N1+N2)	N3	% (N3)	N4	N5	%(N4+N5)
V01	19	100.00%	0	0.00%	**100.00%**	0	0.00%	0	0	**0.00%**
V02	15	78.95%	3	15.79%	**94.74%**	0	0.00%	1	0	**5.26%**
V03	11	57.89%	6	31.58%	**89.47%**	1	5.26%	0	1	**5.26%**
V04	10	52.63%	7	36.84%	**89.47%**	1	5.26%	1	0	**5.26%**
V05	8	42.11%	4	21.05%	**63.16%**	6	31.58%	0	1	**5.26%**
V06	8	42.11%	6	31.58%	**73.68%**	5	26.32%	0	0	**0.00%**
V07	11	57.89%	5	26.32%	**84.21%**	2	10.53%	1	0	**5.26%**
V08	12	63.16%	5	26.32%	**89.47%**	0	0.00%	1	1	**10.53%**
V09	8	42.11%	5	26.32%	**68.42%**	5	26.32%	0	1	**5.26%**
V10	10	52.63%	5	26.32%	**78.95%**	3	15.79%	0	1	**5.26%**
V11	11	57.89%	6	31.58%	**89.47%**	1	5.26%	1	0	**5.26%**
V12	6	31.58%	4	21.05%	**52.63%**	6	31.58%	2	1	**15.79%**
V13	14	73.68%	5	26.32%	**100.00%**	0	0.00%	0	0	**0.00%**
V14	7	36.84%	9	47.37%	**84.21%**	2	10.53%	0	1	**5.26%**
V15	8	42.11%	7	36.84%	**78.95%**	3	15.79%	0	1	5.26%
V16	12	63.16%	6	31.58%	**94.74%**	0	0.00%	1	0	**5.26%**
V17	11	57.89%	6	31.58%	**89.47%**	1	5.26%	1	0	**5.26%**
V18	9	47.37%	5	26.32%	**73.68%**	3	15.79%	0	2	**10.53%**
V19	9	47.37%	8	42.11%	**89.47%**	1	5.26%	1	0	**5.26%**
V20	6	31.58%	7	36.84%	**68.42%**	5	26.32%	0	1	**5.26%**
V21	13	68.42%	5	26.32%	**94.74%**	0	0.00%	0	1	**5.26%**

Tabla 3 – Porcentaje de población por nivel de acuerdo, respecto a las ventajas de la CEPTIU

Para determinar de una forma más clara el análisis de niveles de acuerdo, por cada una de las universidades se realiza una suma de los niveles de acuerdo (1 a 5) que han sido respondidos en todas las 21 ventajas, es decir si una universidad estuviera totalmente de acuerdo (valor=5) con las 21 ventajas propuestas, la suma sería igual a 105 (21 x 5). De esta forma generamos una nueva escala de valores agrupados, con el fin de valorar de forma general el nivel de acuerdo de las 19 universidades, respecto a la existencia de las 21 ventajas de la CEPTIU propuestas. Al realizar las sumas de cada una de las universidades se determina que 15 universidades (78.9%) suman un valor entre 85 y 105, lo cual corresponde al nivel más alto de percepción (totalmente de acuerdo), le sigue con un 15.8% (3 universidades) que están de acuerdo con las 21 ventajas propuestas, finalmente solo una universidad (5.3%) consideran a la propuesta en un punto medio (ni de acuerdo, ni en desacuerdo), el nivel de desacuerdo a nivel general no existe en ninguna universidad. En conclusión, el 94.7% de universidades, es decir 18 de las 19 universidades encuestadas consideran que las 21 ventajas de la cartera estratégica de proyectos de TI propuesta por los investigadores son adecuadas y son causas de tener implementada una CEPTIU ideal (Figura 2).

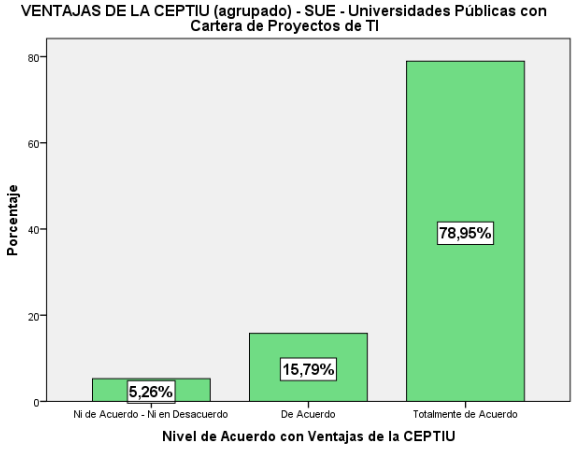


Figura 2 – Porcentaje de valoración por agrupación de las ventajas de la CEPTIU

Adicionalmente, se pregunta a las 19 universidades públicas del SUE que tienen implementada una cartera de proyectos de TI, si su cartera tiene los elementos considerados. Los resultados aparecen en la figura 3. Los elementos E02 (*tiene en cuenta los objetivos estratégicos de la universidad*) y E04 (*tiene en cuenta si los proyectos cumplen con las leyes y normativas tanto internas como externas de la universidad*) son los elementos que la mayoría de las universidades (16 de 19) posee en su cartera de proyectos de TI. Los elementos E06 (*dispone de indicadores para medir el impacto estratégico o valor que aportan los proyectos a la universidad una vez que se han finalizado*) y E13 (*dispone de una oficina de apoyo a la gestión de la cartera*) son los menos existentes.

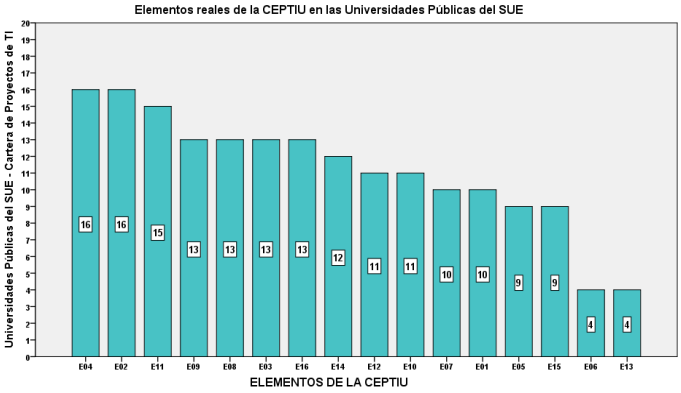


Figura 3 – Existencia de elementos de la CEPTIU en las Universidades públicas del SUE

En la figura 4 se presentan los resultados de la existencia o no de los elementos propuestos en la investigación en las 19 universidades que tienen una cartera de proyectos de TI y

que han respondido el cuestionario. Se puede observar que dos universidades (la U11 y la U19) consideran tener todos los elementos propuestos por los investigadores. Por el contrario, una universidad (la U09) a pesar de tener una cartera de proyectos de TI considera no tener ninguno de los elementos propuestos.

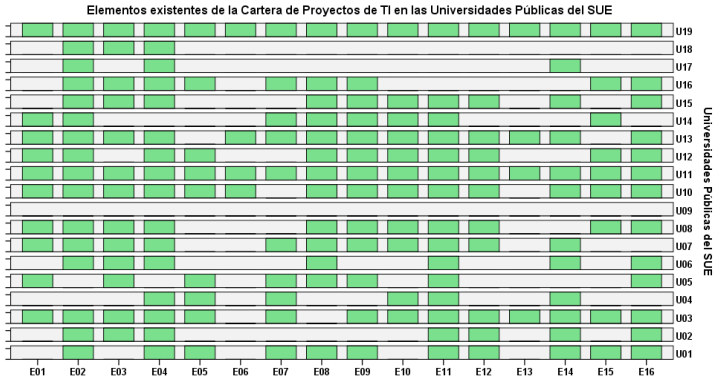


Figura 4 – Elementos existentes en las Universidades Públicas del SUE

De la misma forma, se preguntó a las universidades públicas del SUE si perciben la existencia de las 21 ventajas consideradas por los investigadores, como consecuencia de la existencia de una cartera de proyectos de TI implementada en su universidad (Figura 5). Las ventajas V01 (*contribuir a que los proyectos se alineen con los objetivos estratégicos de la universidad*) y V21 (*promover el control centralizado y facilita la elaboración del presupuesto a corto y medio plazo*) son percibidas en la mayoría de las universidades (17 de 19) y la ventaja V12 (*contribuir a democratizar la solicitud de proyectos TI abriéndolos a la comunidad universitaria*) es la menos percibida.

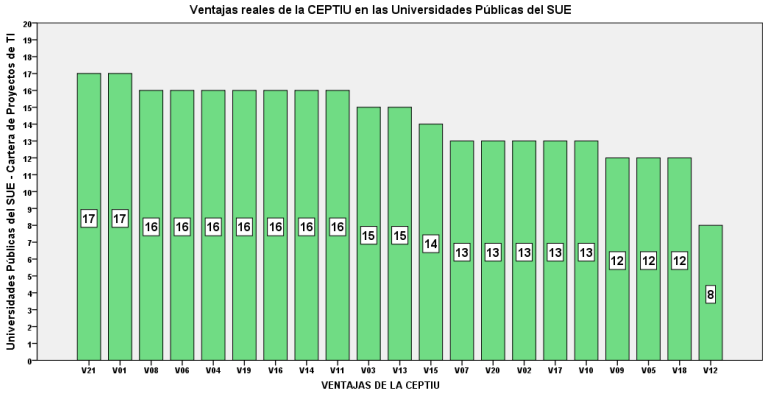


Figura 5 – Existencia de ventajas de la CEPTIU en las Universidades públicas del SUE

El último análisis realizado en esta investigación es la influencia entre los elementos reales y las ventajas reales existentes en las universidades públicas del SUE. Se toma como punto de referencia la relación causa-efecto obtenido construida y validada por los investigadores y los

expertos en gobernanza de las TI (Figura 1). Esta matriz nos permite visualizar la relación de las dos variables categóricas (elementos y ventajas). Se procede a realizar el análisis por cada elemento (parte tangible de la CEPTIU) con cada ventaja relacionada, según la construcción y validación realizada por los investigadores y el grupo de expertos (Figura 1). Por ejemplo se analiza a las universidades que si tienen el elemento E01 con las ventajas V06, V12 y V15 percibidas por las mismas universidades, con un análisis de frecuencias. El proceso de análisis se realiza con cada uno de los elementos (Tablas 4 - 19). Se puede observar que las universidades que tienen los distintos elementos, también asumen las ventajas relacionadas.

Ventaja	SI tiene	%(SI tiene)	NO tiene	%(NO tiene)
V06	10	100.00%	0	0.00%
V12	8	80.00%	2	20.00%
V16	10	100.00%	0	0.00%

Tabla 4 – Elemento E01 - 10 universidades públicas del SUE

Ventaja	SI tiene	%(SI tiene)	NO tiene	%(NO tiene)
V01	16	100.00%	0	0.00%
V02	12	75.00%	4	25.00%
V08	14	87.50%	2	12.50%
V13	14	87.50%	2	12.50%
V14	14	87.50%	2	12.50%
V17	13	81.25%	3	18.75%

Tabla 5 – Elemento E02 – 16 universidades públicas del SUE

Ventaja	SI tiene	%(SI tiene)	NO tiene	%(NO tiene)
V11	11	84.62%	2	15.38%
V13	12	92.31%	1	7.69%
V14	12	92.31%	1	7.69%
V19	13	100.00%	0	0.00%
V20	10	76.92%	3	23.08%

Tabla 6 – Elemento E03 – 13 universidades públicas del SUE

Ventaja	SI tiene	%(SI tiene)	NO tiene	%(NO tiene)
V01	16	100.00%	0	0.00%
V02	12	75.00%	4	25.00%
V04	14	87.50%	2	12.50%
V08	14	87.50%	2	12.50%
V15	13	81.25%	3	18.75%

Tabla 7 – Elemento E04 – 16 universidades públicas del SUE

Ventaja	SI tiene	%(SI tiene)	NO tiene	%(NO tiene)
V11	9	100.00%	0	0.00%
V18	7	77.78%	2	22.22%
V19	8	88.89%	1	11.11%
V20	6	66.67%	3	33.33%
V21	8	88.89%	1	11.11%

Tabla 8 – Elemento E05 – 9 universidades públicas del SUE

Ventaja	SI tiene	%(SI tiene)	NO tiene	%(NO tiene)
V01	4	100.00%	0	0.00%
V02	4	100.00%	0	0.00%
V05	4	100.00%	0	0.00%
V07	4	100.00%	0	0.00%
V08	4	100.00%	0	0.00%

Tabla 9 – Elemento E06 – 4 universidades públicas del SUE

Ventaja	SI tiene	%(SI tiene)	NO tiene	%(NO tiene)
V04	9	90.00%	1	10.00%
V07	8	80.00%	2	20.00%
V13	8	80.00%	2	20.00%
V18	8	80.00%	2	20.00%

Tabla 10 – Elemento E07 – 10 universidades públicas del SUE

Ventaja	SI tiene	%(SI tiene)	NO tiene	%(NO tiene)
V09	11	84.62%	2	15.38%
V13	12	92.31%	1	7.69%
V19	13	100.00%	0	0.00%
V21	12	92.31%	1	7.69%

Tabla 11 – Elemento E08 – 13 universidades públicas del SUE

Ventaja	SI tiene	%(SI tiene)	NO tiene	%(NO tiene)
V01	12	92.31%	1	7.69%
V03	12	92.31%	1	7.69%
V09	11	84.62%	2	15.38%
V21	12	92.31%	1	7.69%

Tabla 12 – Elemento E09 – 13 universidades públicas del SUE

Ventaja	SI tiene	%(SI tiene)	NO tiene	%(NO tiene)
V06	11	100.00%	0	0.00%
V10	10	90.91%	1	9.09%
V13	10	90.91%	1	9.09%
V16	11	100.00%	0	0.00%
V19	10	90.91%	1	9.09%

Tabla 13 – Elemento E10 – 11 universidades públicas del SUE

Ventaja	SI tiene	%(SI tiene)	NO tiene	%(NO tiene)
V02	12	80.00%	3	20.00%
V03	13	86.67%	2	13.33%
V08	14	93.33%	1	6.67%
V10	12	80.00%	3	20.00%
V16	15	100.00%	0	0.00%

Tabla 14 – Elemento E11 – 15 universidades públicas del SUE

Ventaja	SI tiene	%(SI tiene)	NO tiene	%(NO tiene)
V06	11	100.00%	0	0.00%
V08	10	90.91%	1	9.09%

Tabla 15 – Elemento E12 – 11 universidades públicas del SUE

Ventaja	SI tiene	%(SI tiene)	NO tiene	%(NO tiene)
V06	4	100.00%	0	0.00%
V08	4	100.00%	0	0.00%
V10	4	100.00%	0	0.00%
V15	4	100.00%	0	0.00%

Tabla 16 – Elemento E13 – 4 universidades públicas del SUE

Ventaja	SI tiene	%(SI tiene)	NO tiene	%(NO tiene)
V04	11	91.67%	1	8.33%
V06	12	100.00%	0	0.00%
V07	12	100.00%	0	0.00%
V08	11	91.67%	1	8.33%
V17	11	91.67%	1	8.33%
V19	10	83.33%	2	16.67%

Tabla 17 – Elemento E14 – 12 universidades públicas del SUE

Ventaja	SI tiene	%(SI tiene)	NO tiene	%(NO tiene)
V04	8	88.89%	1	11.11%
V08	8	88.89%	1	11.11%
V18	8	88.89%	1	11.11%
V21	8	88.89%	1	11.11%

Tabla 18 – Elemento E15 – 9 universidades públicas del SUE

Ventaja	SI tiene	%(SI tiene)	NO tiene	%(NO tiene)
V06	12	92.31%	1	7.69%
V10	10	76.92%	3	23.08%
V19	13	100.00%	0	0.00%

Tabla 19 – Elemento E16 – 13 universidades públicas del SUE

4. Conclusiones

Se ha establecido un listado de 21 ventajas que aporta la implantación de una Cartera Estratégica de Proyectos de TI en una universidad. Este listado es fruto de análisis de los trabajos publicados, de la revisión por parte de un grupo de expertos y de la validación de los responsables TI de las universidades públicas que tienen implantada una cartera. El 94.7% de universidades participantes en la investigación, es decir 18 de las 19 universidades encuestadas consideran que las 21 ventajas de la cartera estratégica de proyectos de TI propuesta por los investigadores son adecuadas y son consecuencia de tener implementada una CEPTIU ideal. Por lo tanto, la propuesta de estas 21 ventajas de la CEPTIU ha sido validada, tanto por el grupo de expertos como por los responsables de TI de las universidades públicas españolas que tienen implantad una cartera.

En el caso de estudio (universidades públicas del SUE), se percibe la existencia de ventajas consideradas como causa de la presencia de determinados elementos en su cartera de proyectos de TI. V01 (*contribuir a que los proyectos se alineen con los objetivos estratégicos de la universidad*) y V21 (*promover el control centralizado y facilita la elaboración del presupuesto a corto y medio plazo*) son percibidas en la mayoría de las universidades (17 de 19). Y, por el contrario, la ventaja V12 (*contribuir a democratizar la solicitud de proyectos TI abriéndolos a la comunidad universitaria*) es la menos percibida.

Para hacer realidad dichas ventajas, se ha establecido un listado de 16 elementos a incorporar en una Cartera de Proyecto de TI que pretenda ser estratégica. Con respecto al caso de estudio (universidades públicas del SUE), se observa la existencia de la mayor parte de elementos en las 19 universidades encuestadas. Los elementos E02 (*tiene en cuenta los objetivos estratégicos de la universidad*) y E04 (*tiene en cuenta si los proyectos cumplen con las leyes y normativas tanto internas como externas de la universidad*) son los elementos que la mayoría de las universidades (en 16 de las 19) posee en su cartera de proyectos de TI. Los elementos E06 (*dispone de indicadores*

para medir el impacto estratégico o valor que aportan los proyectos a la universidad una vez que se han finalizado) y E13 (*dispone de una oficina de apoyo a la gestión de la cartera*) son los menos existentes (en 4 universidades de las 19).

Finalmente, como resultado de la investigación, se ha determinado una tabla que determina la relación causa-efecto entre los elementos y las ventajas de la CEPTIU. Un alto porcentaje de universidades que afirma tener implementados en su cartera de proyectos de TI elementos pertenecientes a la propuesta, afirman de igual forma en un alto porcentaje por cada elemento, la percepción de existencia de las ventajas relacionadas a cada uno de los 16 elementos. Esta relación entre elementos y ventajas permitirá a los responsables universitarios, al diseñar su cartera de proyectos TI seleccionar aquellos elementos que les aporten aquellas ventajas que desean obtener.

Por lo tanto, la propuesta confirma su valor y funcionamiento y se valida la propuesta de 16 elementos, 21 ventajas y la relación causa efecto entre los elementos y ventajas de una cartera de proyectos de TI considerada estratégica (CEPTIU). Las universidades que carecen de la existencia de ciertos elementos pueden tomar acciones de mejoramiento que puedan permitir la introducción de dichos elementos, de esta forma la propuesta sirve de base a las universidades para establecer la estructura y funcionamiento de su cartera de proyectos de TI, además sirve de referencia para la construcción de una CEPTIU, es decir una cartera estratégica de proyectos de TI específico para universidades.

Esta investigación será la base para la construcción de un modelo que permita determinar la alineación de la Cartera de Proyectos de TI con los principios de la norma ISO/IEC 38500. Ya en esta investigación queda demostrada la fuerte relación de la Cartera de Proyectos de TI con la Gobernanza de las TI, como se puede deducir de las ventajas más consideradas y de los elementos más implantados en las carteras de proyectos TI de las universidades públicas españolas.

Referencias

Bonham, S. (2005). *IT Project Portfolio Management*. Artech House.

Cameron, B. H. (2005). IT Portfolio Management : Implications for IT Strategic Alignment. In *Proceedings of the Eleventh Americas Conference on Information Systems*.

Chiang, I. R. y Nunez, M. A. (2009). Strategic Planning and Project Selection for IT Portfolio. *Association for Information Systems*, 104–113.

Correa, J. A. y Benavides, J. T. N. (2013). ¿Cómo Impacta la implementación de una Cartera de Proyectos en el Gobierno de las TI de una Institución de Educación Superior? ICESI.

Correa, J. A. y Benavides, J. T. N. (2016). CPTI4Uv2: Modelo de Cartera de proyectos de TI para las Universidades, *TICAL 2016*, Buenos Aires, Argentina

Cubeles, A. (2007). *La Dirección del Portafolio de Proyectos TI*. Barcelona.

De Haes, S., & Van Grembergen, W. (2006). IT Governance and its Mechanisms. *Information Systems Control Journal*, 1, 27–33.

Fernández, M. O. (2010). Proyecto de arranque del Gobierno de las TI en una Universidad. Universidad de Almeria, p. 129. Almeria.

Fernández Martinez, A. y Llorens Largo, F. (2011). *Gobierno de las Tecnologías de la Información en Universidades*. CRUE Conferencia de Rectores de las Universidades Españolas. Madrid.

Fernández Martínez, A., Gumbau Mezquita, J. P., y Llorens Largo, F. (2012). *Pilot Project for Implementing Corporate Governance of IT*. Conferencia de Rectores de las Universidades Españolas (CRUE), Madrid.

Fernández Martínez, A., Hontoria Hernández, E., & Llorens Largo, F. (2014). ¿Cómo gobiernan sus Tecnologías de la Información las universidades españolas? *Cuarta Conferencia de Directores de Tecnología de Información, TICAL2014*, pp. 425–442. Cancún, México.

Gleisberg, E., Zondag, H. y Chaudron, M. R. V. (2008). An Empirical Study into the State of Practice and Challenges in IT Project Portfolio Management. Proc. of the *34th Euromicro Conference Software Engineering and Advanced Applications*, IEEE Computer Society, pp. 2.

Gómez, J., Jimenez, T., Gumbau, J. y Llorens, F. (2017). *UNIVERSITIC 2017 Análisis de las TIC en las Universidades Españolas*. Madrid.

Hernández, R., Fernández, C. y Baptista, P. (2014). Metodología de la investigación (6ª ed.). México: Editorial Mc Graw-Hill.

Jairak, K. y Praneetpolgrang, P. (2013). Applying IT governance balanced scorecard and importance-performance analysis for providing IT governance strategy in university. *Information Management & Computer Security*, 21(4), 228–249.

Karhade, P. y Shaw, M. (2007). Rejection and Selection Decisions in the IT Portfolio Composition Process: An Enterprise Risk Management Based Perspective. In *Americas Conference on Information Systems*.

Maizlish, B. y Handler, R. (2005). *Portfolio Management Step-by-Step*. (M. G. Inc, Ed.) (1st ed.). New Jersey: John Wiley & Sons, Inc., Hoboken.

Medellin, E. (2006). *Gestión de Cartera de Proyectos Tecnológicos*

PMI. (2013). *PMBOK Guía de Fundamentos para la Dirección de Proyectos*. (P. Book, Ed.) (5th ed.). Pensilvania: PMI.

Rosselet, U., & Wentland, M. (2009), Knowledge Management for IT Project Portfolio. Proceedings of the *10th European Conference on Knowledge Management* (ECKM).

Sierra Alvarez, L. (2012). *¿Cómo implantar el Gobierno de las Tecnologías de Información en Instituciones de Educación Superior?* UNIVERSIDAD ICESI, Santiago de Cali

Zhijie, X. (2012). A Method of IT Investment Portfolio Optimization in the Government Sector Integrated with IT Governance. In *Second International Conference on Business Computing and Global Informatization* (pp. 6–9).

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 11/09/2018
Aceitação/Acceptance: 09/11/2018

Técnicas de mitigación para principales vulnerabilidades de seguridad en aplicaciones web

Alex Zambrano[1], Teresa Guarda[1,2,3], Edward Vladimir Haro Valenzuela[1], Geovanni Ninahualpa Quiña[1]

aleoma1991@hotmail.com, tguarda@gmail.com, vladiharo7@gmail.com, gninahualpa@espe.edu.ec

[1] Universidad de las Fuerzas Armadas ESPE, Centro de Postgrados, 170511, Quito, Ecuador

[2] Universidad Estatal Península de Santa Elena, Santa Elena, Ecuador – UPSE, Av. La Libertad, Santa Elena, Ecuador

[2] Algoritmi Centre, Minho University, Guimarães, Portugal

Pages: 299–308

Resumen: La seguridad de la información se ha vuelto un elemento imprescindible y de gran consideración para las empresas, puesto que ahora los datos de las mismas se encuentran procesados, enviados y almacenados en aplicaciones y servicios web, por lo que es mandatorio el uso de métodos y procesos que permitan asegurar la información ante las posibles amenazas de un ataque a estos sistemas, es así que el presente artículo analiza las principales vulnerabilidades que se encuentran en las aplicaciones web, proponiendo el uso de diferentes técnicas para mitigar estas debilidades con el propósito de reducir las inseguridades que se pueden encontrar en los sistemas vía web que utilizan las distintas organizaciones o empresas a nivel mundial.

Palabras-clave: Seguridad de software, aplicaciones web, vulnerabilidades, mitigación.

Mitigation techniques for security vulnerabilities in web applications

Abstract: The security information has become an essential element of great consideration for the companies, since now the data of the companiess were processed, sent and stored in applications and web services so, it is mandatory the use of methods and processes that allow to ensure the information before the possible threats of an attack to these systems, is this way that the present article analyzes the main vulnerabilities that are in the web applications, proposing the use of different techniques to mitigate these risks with the purpose of reducing the insecurities that can be found in web systems used by different organizations or companies worldwide.

***Keywords*:** software security, web applications, vulnerabilities, mitigation.

1. Introducción

El desarrollo de las aplicaciones de software vía web (Mateu, 2004) en los últimos años, han permitido dinamizar el trabajo en las organizaciones puesto que ahora solo es necesario tener acceso a la red para ingresar a los sistemas y con esto a la información relevante y útil ya que de esta forma se gestiona ágil y rápidamente los datos de las entidades; pero por el incorrecto uso de los medios tecnológicos y los problemas que se han generado por la mala utilización, fallos de seguridad, pérdidas de información y fallos de integridad ha ocasionado que las organizaciones fijen la mira en métodos de seguridad para los sistemas de información (Baskerville, 1993) permitiendo así un crecimiento acelerado de los mismos, con el objetivo de proteger la confidencialidad, integridad y disponibilidad de los recursos de un sistema informático. Los datos de los sistemas de información son los activos más valiosos que poseen las organizaciones en la actualidad y por lo que se hace de vital importancia para las empresas salvaguardar estos recursos frente a posibles ataques, ya sean estos por las vulnerabilidades del software, compartir contraseñas, utilización de ingeniería social, denegación de servicios, entre otros, (Romero M & Hisham M., 2009) que por medio de estas inseguridades se pueden alterar o robar información de los sistemas, generando grandes perjuicios y pudiendo llevar a la empresa a perder su credibilidad y confianza de los clientes, llevando así a causar un grave daño para la entidad.

El desarrollo de las aplicaciones web implica tener en cuenta las amenazas que involucra el entorno web (Mateu, 2004), es decir que existen vulnerabilidades en la codificación del aplicativo, en la transmisión de los datos, sus comunicaciones, el almacenamiento, controles de acceso, confidencialidad e integridad (Lee, Won, Cho, Park, & Won, 2012), por este motivo es necesario la implementación de varias técnicas y mecanismos que permiten asegurar la información en la web (Bertino, Pagani, Rossi, & Samarati, 2000) que va a ser enviada, procesada y consultada por medio de los aplicativos de forma confiable y segura, así es como solo las personas autorizadas puedan visualizar o actualizar dicha información para que no exista ningún tipo de fuga o alteración de datos, puesto que de lo contrario se vería reducida la privacidad y confiablidad en el uso de estos aplicativos.

Por lo anteriormente expuesto, el escrito se centra el estudio de técnicas y mecanismos que permiten el aseguramiento de la seguridad de la información en las aplicaciones web, con el objetivo de mitigar las vulnerabilidades en las áreas de confidencialidad, control de accesos e integridad, proponiendo un análisis para su implementación en los sistemas vía web para que por medio de su adopción permitan incrementar la seguridad y protección de los datos de las organizaciones.

El presente escrito se encuentra estructurado de la siguiente manera: el análisis del problema se describe en la sección 2. La sección 3 muestra las principales vulnerabilidades de las aplicaciones web. El estudio de las diferentes técnicas y mecanismos para mitigar las vulnerabilidades y amenazas de los sistemas web en la sección 4, y finalmente las conclusiones en la sección 5.

2. Análisis de Problema

El crecimiento acelerado de la tecnología en los últimos años ha permitido que las organizaciones migren sus aplicaciones tradicionales hacia servicios enfocados a

la internet, aplicaciones web y últimamente a la computación en la nube (Feuerlicht & Tran, 2014) (Joyanes Aguilar, 2012) como se ve representado en la Figura 1; a consecuencia de este vertiginoso cambio ocurrido se ha vuelto la información un activo de importantísimo valor y un profundo cuidado para las empresas, puesto que el procesamiento de los datos y su interpretación son utilizados de forma efectiva para la toma de decisiones, reducir incertidumbre y la previsión de distintos escenarios a futuro. Según la literatura analizada existen varias vulnerabilidades y amenazas que hacen que los atacantes puedan usar cualquiera de ellas para poder realizar algún ataque y estar en la facultad de sustraerse información privilegiada de las instituciones (Zachara, 2017).

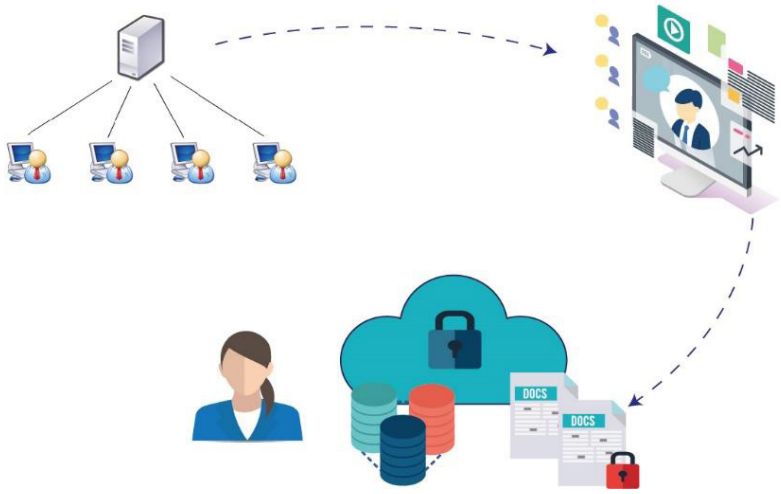


Figura 1 – Evolución de sistemas cliente servidor a sistemas web y computación en la nube.

A medida que las organizaciones buscan tener más servicios en el internet con la finalidad de incrementar su productividad y su posición en el mercado, se hace necesario que las aplicaciones que éstas utilizan se tengan que ir adaptando a las exigencias y nuevas tecnologías del entorno, sin embargo estas tecnologías se hacen cada vez más complejas existiendo una mayor dificultad para lograr que la información que se maneja esté segura, propiciando a que los esfuerzos también se vayan encaminando hacia lograr que los sistemas, datos y otros activos estén protegidos en todos sus niveles.

El auge de las organizaciones por tener su información en aplicaciones web para que sea de mayor facilidad y utilidad para sus usuarios, ha traído de igual manera a ladrones virtuales que por medio de diferentes métodos buscan descubrir y explotar las vulnerabilidades que pueden tener estos sistemas, aprovechándose de estas debilidades pueden llegar a provocar graves daños en los sistema y más aún en la información que estos poseen, por lo que la seguridad de la información se ha convertido en una necesidad imperante con el fin de ofrecer y garantizar confianza a clientes y proveedores.

3. Vulnerabilidades

Las aplicaciones web como demás sistemas informáticos presentan vulnerabilidades, las cuales son debilidades que pueden ser tanto a nivel físico como lógico, capaces de

ser identificadas después de extensas y complejas búsquedas de estos bugs o agujeros, los cuales pueden tratarse de limitaciones tecnológicas, fallas de programación a nivel de código o por falta de concientización y aplicación de las políticas de seguridad en los desarrolladores de software de las instituciones; trayendo de esta manera consigo que personas internas o externas consigan explotar los diferentes fallos con el fin de sustraer, modificar o violar la confidencialidad e integridad de los datos almacenados en las aplicaciones web como se ilustra en la Figura 2, además de otros ataques que pueden afectar la disponibilidad de la información como la denegación de servicios.

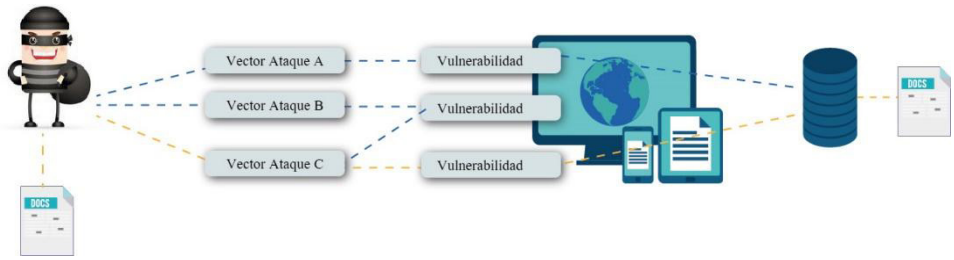


Figura 2 – Atacante usando vulnerabilidades para robar información

Las principales y más comunes vulnerabilidades de las aplicaciones web pueden ser agrupadas en los siguientes temas: inyecciones; autenticación y gestión de sesiones; exposición de datos sensibles; control de acceso; y secuencia de comandos en sitios cruzados.

3.1. Inyecciones

Este tipo de fallo de seguridad ocurre cuando datos no confiables son enviados como parte de una consulta o comandos al sistema y el intérprete puede emitir una respuesta concediéndole acceso a información no autorizada (Sarhan, Farhan, & Al-Harby, 2017) engañándolo de esta manera para realizar acciones no deseadas o contempladas por el sistema, estas inyecciones pueden ser tipo SQL en su gran mayoría. Las inyecciones pueden producir pérdidas de información, falla de integridad en los datos, por lo que debemos considerar el valor de los datos que son afectados por estas vulnerabilidades, a continuación se detalla más ampliamente estas debilidades. Las inyecciones SQL radican en enviar códigos SQL modificados por medio de formularios, con el fin de alterar las búsquedas hacia la base de datos generando así posibilidades de conseguir o alterar la información (Kaur & Kour, 2016).

3.2. Autenticación y Gestión de Sesiones

Este tipo de vulnerabilidad en las aplicaciones web radican en que las credenciales de acceso de lo usuario pueden estar almacenados en textos claros, es decir sin ningún tipo de cifrado o hash (Talti, 2015), además de que esta información puede transmitirse por canales que no son seguros o no tienen ningún tipo de seguridad. De la misma manera las sesiones pueden ser atacadas puesto que existen brechas o fallos de seguridad por

las cuales los identificadores, contraseñas de sesiones son expuestas por las URL, de la misma manera se pueden establecer sesiones sin expiración o caducidad dejando así un acceso a otros individuos capaces de usar la misma sesión para realizar cualquier tipo de actividad no autorizada.

3.3. Exposición de Datos Sensibles

La exposición de datos sensibles hace referencia a que dentro de las aplicaciones web la mayoría de los desarrolladores al programar dejan en texto claro datos relevantes para la organización como puede ser tarjetas de crédito, información de proveedores, datos personales, credenciales de autenticación, entre otros, además no sólo se hace mención al almacenamiento en las bases de datos sino también a los canales de transmisión de los datos y la utilización de técnicas de cifrado débiles u obsoletas que representan un considerable riesgo.

3.4. Control de Acceso

El control de acceso de las aplicaciones web es un mecanismo que en gran parte se ha basado en la creación y utilización de roles para brindar una significativa seguridad a los procesos del sistema (Ferraiolo, Sandhu, Gavrila, Kuhn, & Chandramouli, 2001), pero lo cual no lo hace completo, puesto que las vulnerabilidades que se presentan también están a nivel de las funcionalidades, es decir existen problemas puesto que pueden visualizarse navegaciones no autorizadas en la dirección URL del sistema, concediéndoles de esta manera a usuarios que pueden ser auténticos paso a funciones que no están autorizados o que llegan a provocar alteración de esa información en el cambio de alguna función o modificación de parámetros en la dirección web.

3.5. Secuencia de Comandos en Sitios Cruzados

Los ataques Cross-site scripting (XSS) (Grossman, Hansen, Petkov, Rager, & Fogie, 2007) facultan la inserción de scripts que se pueden ejecutar en el lado del cliente en páginas o aplicaciones web que son accesibles por otros usuarios, popularmente son usados para el hurto de sesiones. Este tipo de vulnerabilidad permiten a los atacantes incrustar código generalmente JavaScript en variables de tipo Get o Post, un ejemplo de este tipo es donde el dato a consultar aparece después como resultado de esta búsqueda (Martin, Björn, & Joachim, 2008).

4. Técnicas de Mitigación

La seguridad de la información para cualquier organización ya sea de carácter pública, privada o de cualquier otra índole es fundamental, puesto que los datos son los elementos informáticos más valiosos que ésta posee y los cuales deben ser salvaguardados ante posibles amenazas externas o internas, es por esto que se hace ineludible el uso de tecnologías, programas y procesos para controlar y reducir las vulnerabilidades con el objetivo de minimizar los riesgos en la alteración o pérdida de información como en la Figura 3 se ve reflejado existen varios métodos de encriptación y protocolos de transferencia aportando mayor seguridad en los aplicativos.

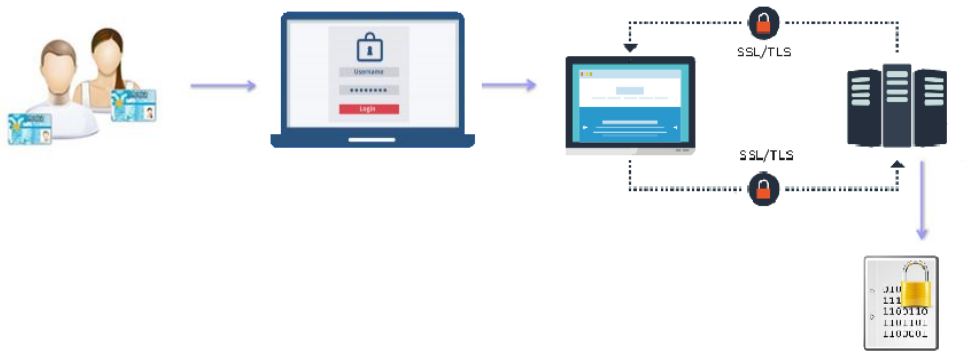


Figura 3 – Seguridad en aplicaciones web

Existen diversos mecanismos y técnicas de mitigación para las diferentes vulnerabilidades expuestas en la sección anterior, por lo que a continuación se establecen los procedimientos utilizados y que permiten mitigar estas debilidades de las aplicaciones (Guarda, Bustos, Torres, & Villao, 2018).

4.1. Inyecciones

Las inyecciones siguen siendo los ataques más habituales en las aplicaciones, por lo que se han generado varias técnicas para la mitigación de estas amenazas, entre las más utilizadas podemos mencionar: para los desarrolladores de sistemas es necesario reestructurar o modificar el código para evitar la generación de consultas dinámico para que no puedan ser afectado por este tipo de ataque, de la misma manera OWASP (The OWASP Foundation, 2001) nos provee una biblioteca para varias operaciones y métodos de seguridad que un desarrollador necesita para construir aplicaciones seguras, en este caso esta librería ESAPI (The OWASP Foundation, 2009) permite definir los parámetros que van a pasar a los controles de seguridad, es decir que nosotros creamos expresiones regulares que luego estas pasarán a ser validadas y en el caso de ser admitidas se procesarán reduciendo así códigos sospechoso, además pueden utilizarse validaciones de entrada, reduciendo con esto cualquier tipo de dato ingresado que no ajuste a las especificaciones implementadas para las peticiones ya sean estas: longitud de caracteres, sintaxis, tipo de entrada para evitar de esta forma la inclusión de caracteres especiales.

Existe además otra herramienta utilizada para mitigar las vulnerabilidades por este tipo de amanezcas como es AMNESIA (Halfond & Orso, 2005) la cual consiste en examinar puntos clave en el código de aplicación que tienen sentencias SQL, específicamente consultas, para posteriormente construir modelos de consultas que se podrían generar, con el objeto de que en tiempo de ejecución de la aplicación estas sean validadas con estos modelos para que se validen y solo se lleguen a procesar las consultas legítimas.

4.2. Autenticación y Gestión de sesiones

La autenticación es un factor importante de seguridad por tal motivo se hace necesario que en las aplicaciones web y en especial en todas sus páginas se considere preciso su

validación, es decir que en todas las páginas sea ineludible esta verificación, de esta manera se reduce la exposición de páginas a personas no autorizadas que pudieran conocer URLs de la aplicación; también es importante que las contraseñas no sean almacenadas en texto claro en las bases de datos pues esto conlleva a un riesgo muy alto al estar expuestas, es necesario que se use técnicas para asegurar esta información y una de los procedimientos más utilizados y que ha aminorado estos ataque es la utilización de cifrados con la función salt (Boonkrong & Somboonpattanakit, 2016), que son bits aleatorios que conjuntamente con las contraseña al producirse el hash seguro de esta combinación estamos asegurando aún más la información frente a cualquier tipo de ataque de diccionario o fuerza bruta.

De la misma manera como desarrolladores las técnicas más usadas para la gestión de sesiones se basan en la programación y utilización de librerías como ESAPI (The OWASP Foundation, 2009) que nos permiten asegurar estas vulnerabilidades con los controles que nos ofrecen como que no existan sesiones concurrentes al sistema, se invaliden las sesiones si el usuario ha cerrado su sesión o se finalicen en un determinado tiempo de inactividad o de un prolongado período de tiempo que se haya iniciado su actividad, entre otras.

4.3. Exposición de Datos Sensibles

Para mitigar este tipo de vulnerabilidades hay que tener presente si es realmente necesario almacenar este tipo de información privada puestos que, estos datos son de gran valor y que no pueden ser alterados o robados, por esta razón es necesario que se determinen cuáles son los elementos indispensables que se deban almacenar y cómo tienen que ser guardados, para esto existen técnicas como la encriptación, utilizar algoritmos seguros como la encriptación híbrida, la utilización de AES y RSA y la utilización de hash (Ramaraj, Karthikeyan, & Hemalatha, 2009), además es necesario que no solo los datos estén seguros sino de la misma manera los canales de comunicación por lo que la utilización de certificados SSL/TLS (Lee, Malkin, & Nahum, 2007) permite establecer comunicaciones más fiables a través de Internet, ya que los datos son cifrados por medio de métodos criptográficos mientras se navegan en aplicaciones web.

4.4. Control de Acceso

Las técnicas que se deben implementar para los controles de acceso deben ser desde el desarrollo del sistema, puesto que allí es donde se especifican quienes serán los encargados de otorgar los permisos, de verificar que las reglas estén basadas en controles específicos y que no se pueda manipular información a la cual no tienen los permisos correspondientes, es por esto que existen mecanismos de seguridad como los basados en roles (Sandhu, Coyne, Feinstein, & Youman, 1996), los cuales permiten crear un rol al cual se le asignan varios permisos de acuerdo a sus procesos y funciones en el sistema y luego éstos roles son asignados a los usuarios finales del aplicativo, sin embargo no proporcionan una seguridad completa por lo cual es necesario implementar otros mecanismos de seguridad.

4.5. Secuencia de Comandos en Sitios Cruzados

Como se ha visto en las sección anterior este tipo de ataques traen serias consecuencias por lo que es principal tener en cuenta cuales son las técnicas que nos ayudan a reducir los potenciales asaltos a la información de nuestra aplicación: una de las estrategias que

se ha venido implementando es la estrategia usada por OWASPGuard (The OWASP Foundation, s.f.) que permite verificar la integridad de las solicitudes en los sistemas web a partir de la utilización de un token único que forma parte de un parámetro HTTP que se lo compara con el valor almacenado en el inicio de sesión por parte del usuario; de igual manera existente más librería como ESAPI (The OWASP Foundation, 2009) que también permite al desarrollador utilizar otros mecanismos para protegerse ante estas amenazas, como volver a pedir un campo para la legitimidad de que la persona desea ingresar a la página solicitada o validar las entradas de los formularios

5. Conclusiones

La seguridad que se le debe brindar a la información de los sistemas es fundamental, puesto que ahora la mayoría de datos están en el internet y eso trae consigo una exposición que puede conducir a un eventual riesgo de sufrir pérdidas de confidencialidad e integridad, por lo que se vuelve obligatorio para los directores, jefes de unidades tecnologías y personal de desarrollo de software que adopten políticas de seguridad, aplicación de estándares, mecanismos de protección e instalación de tecnologías que salvaguarden los activos de los sistemas informáticos. Es por tal razón que este estudio realizado sobre las técnicas de mitigación para las más comunes vulnerabilidades que se presentan en las aplicaciones web apuntan a que no existe un solo mecanismo que provee seguridad integral para los sistemas, razón por la que se hace evidente la necesidad de aplicar varias técnicas de seguridad, las cuales deberán inicializar desde el proceso de desarrollo del aplicativo hasta la finalización del ciclo de vida.

El resultado del análisis de las diferentes técnicas que se usan para reducir los principales fallos de seguridad se lo deben plantear al inicio del proyecto puesto que así se conseguirá que en el transcurso de las especificaciones y requisitos se revisen todas las medidas necesarias que se deberán abarcar y que los desarrolladores tienen que implementar, generando así mecanismos de protección; el establecimiento de planes de seguridad refuerzan las prácticas sobre el desarrollo seguro de aplicaciones a la vez que afianzan la utilización de librerías, tecnologías y demás herramientas con la finalidad de brindar confianza y seguridad a los clientes de estos sistemas.

La aplicación de los mecanismos de mitigación de los sistemas de información deberá verse planteado de una forma integral, puesto que esto no solo recae en el personal de tecnologías de la información, sino que se debe generar los protocolos para todas las personas que laboran en la institución promoviendo la adopción de normas de seguridad, ya que de esta manera se puedan cerrar brechas para que no exista fuga de información y exposición de vulnerabilidades que se pueden encontrar en las entidades tanto a nivel físico como de sistemas.

Referencias

Baskerville, R. (1993). Information Systems Security Design Methods: Implications for Information Systems Development. *ACM Computing Surveys, XXV*, 376–413.

Bertino, E., Pagani, E., Rossi, G. P., & Samarati, P. (2000). Protecting Information on the Web. *Association for Computing Machinery, 43*, 189–199.

Boonkrong, S., & Somboonpattanakit, C. (2016). Dynamic salt generation and placement for secure password storing. *IAENG International Journal of Computer Science, 43*, 27–36.

Cova, M., Felmetsger, V., & Vigna, G. (2007). Vulnerability Analysis of Web-Based. *Test and Analysis of Web Services*, 363–394.

Ferraiolo, D. F., Sandhu, R., Gavrila, S., Kuhn, R., & Chandramouli, R. (2001). Proposed NIST standard for role-based access control. *ACM Transactions on Information and System Security, 4*, 224–274.

Feuerlicht, G., & Tran, H. T. (2014). Enterprise application management in cloud computing context. *Association for Computing Machinery, 16*, 517–523.

Grossman, J., Hansen, R., Petkov, P., Rager, A., & Fogie, S. (2007). *XSS Attacks: Cross Site Scripting Exploits and Defense*. Syngress.

Guarda, T., Bustos, S., Torres, W., & Villao, F. (2018). Botnets the Cat-Mouse Hunting. The 2018 International Conference on Digital Science (pp. 408-416). Montenegro: Springer, Cham.

Halfond, W. G., & Orso, A. (2005). AMNESIA: Analysis and Monitoring for NEutralizing SQL-Injection Attacks. *Proceedings of the 20th IEEE/ACM international Conference on Automated software engineering, 20*, 174–183.

Joyanes Aguilar, L. (2012). *Computacion en la Nube : estrategias de Cloud Computing en las empresas*. México: Alfaomega Grupo Editor.

Kaur, P., & Kour, K. P. (2016). SQL Injection: Study and Augmentation. *Proceedings of 2015 International Conference on Signal Processing*, 102–107.

Lee, H. K., Malkin, T., & Nahum, E. (2007). Cryptographic strength of ssl/tls servers: current and recent practices. *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*, 83–92.

Lee, T., Won, G., Cho, S., Park, N., & Won, D. (2012). Detection and Mitigation of Web Application Vulnerabilities Based on Security Testing. *International Conference on Network and Parallel Computing, IX*, 138–144.

Martin, J., Björn, E., & Joachim, P. (2008). XSSDS: Server-Side Detection of Cross-Site Scripting Attacks. *Computer Security Applications Conference,*, 335–344.

Mateu, C. (2004). *Desarrollo de Aplicaciones Web*. Barcelona: Eureca Media.

Ramaraj, E., Karthikeyan, S., & Hemalatha, M. (2009). A Design of Security Protocol using Hybrid Encryption Technique (AES- Rijndael and RSA). *International Journal of The Computer, the Internet and Management, 17*, 78–86.

Romero M, B. D., & Hisham M., H. (2009). Security Vulnerabilities and Mitigation Strategies for Application Development. *Sixth International Conference on Information Technology: New Generations*, 235–239.

Sandhu, R., Coyne, E., Feinstein, H., & Youman, C. (1996). Role-based access control models. *Computer, 29*, 38–47.

Sarhan, A. A., Farhan, S. A., & Al-Harby, F. M. (2017). Understanding and Discovering SQL Injection. *International Conference on Human Factors in Cybersecurity, 593*, 45–51.

Talti, E. I. (2015). Cracking more password hashes with patterns. *IEEE Transactions on Information Forensics and Security*, 1656–1665.

The OWASP Foundation. (2001). *OWASP*. (OWASP Foundation) Recuperado el 23 de Agosto de 2018, de https://www.owasp.org/index.php/Main_Page

The OWASP Foundation. (2009). *OWASP*. Recuperado el 23 de Agosto de 2018, de https://www.owasp.org/images/8/82/Esapi-design-patterns.pdf

The OWASP Foundation. (s.f.). *OWASP*. (OWASP Foundation) Recuperado el 23 de Agosto de 2018, de https://www.owasp.org/index.php/OWASP_CSRFGuard_Project/es

Yu, Q., Liu, X., Bouguettaya, A., & Medjahed, B. (2008). Deploying and managing Web services: issues, solutions and directions. *Springer - VLDB Journal, 17*, 537-572.

Zachara, M. (2017). Identification of scanning and attacks against web applications with graph-based modeling of users' behavior. *IEEE International Conference on Cybernetics, III*.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 30/10/2018
Aceitação/Acceptance: 31/12/2018

Desarrollo backend para aplicaciones web, Servicios Web Restful: Node.js vs Spring Boot

Edward Haro[1], Teresa Guarda[2,1,3], Alex Omar Zambrano Peñaherrera[1], Geovanni Ninahualpa Quiña[1]

vladiharo7@gmail.com, tguarda@gmail.com, aleoma1991@hotmail.com, gninahualpa@espe.edu.ec

[1] Universidad de las Fuerzas Armadas ESPE, Centro de Postgrados, 170511, Quito, Ecuador

[2] Universidad Estatal Península de Santa Elena, Santa Elena, Ecuador – UPSE, Av. La Libertad, Santa Elena, Ecuador

[3] Algoritmi Centre, Minho University, Guimarães, Portugal

Pages: 309–321

Resumen: Como una alternativa más simple a SOAP y a WSDL, REST (Representational State Transfer) se ha ido implementando en proveedores de Web incluso en grandes proveedores como son Google, Yahoo! y Facebook, haciendo más fácil la transmisión de datos entre el cliente y el servidor y viceversa ya que es una arquitectura fácil de hacer y aprender debido a que usa métodos explícitos de HTTP. En este trabajo se va a realizar una comparación entre los frameworks Node.js y Sprint Web Services que son dos tecnologías que nos permiten desarrollar servicios Restful, se va a realizar pruebas de los principales métodos de HTTP (GET, POST, DELETE, PUT) desarrollando servicios web de cada tecnología y comparando características como robustez, integración con otras tecnologías, seguridad, soporte, mantenimiento, estabilidad y ejecución. Obteniendo así un reporte final en el que podamos escoger el framework que mejor se adapte a nuestras necesidades.

Palabras-clave: Node.js; Spring; Restful; Web services, Backend.

Backend development for web applications, Restful Web Services: Node.js vs Spring Boot

Abstract: As a simpler alternative to SOAP and WSDL, REST (Representational State Transfer) has been implemented in Web providers even in large providers such as Google, Yahoo! and Facebook, making it easier to transmit data between the client and the server and the other way as it is an easy to do and learn architecture because it uses explicit HTTP methods. In this work we will make a comparison between the frameworks Node.js and Sprint Web Services that are two technologies that allow us to develop Restful services, we will perform tests of the main HTTP methods (GET, POST, DELETE, PUT) developing web services of each technology and comparing features such as robustness, integration with other technologies, security, support, maintenance, stability and execution. Obtaining a final report in which we can choose the framework that best suits our needs.

***Keywords**:* Node.js; Spring; Restful; Web services, Backend.

1. Introducción

El desarrollo de servicios web se ha convertido en un estándar en la programación al momento de compartir información entre diferentes aplicaciones de software, ya que para distribuir sus funcionalidades estos son independientes del hardware, el sistema operativo, el lenguaje de programación y al igual que son independientes para proveer las funcionalidades, son independientes al momento de ejecutarse y desarrollarse para su uso. (Hamad, Saad, & Abed, 2010).

REST en un tipo de arquitectura orientada a recursos para el desarrollo de servicios web reemplazando especificaciones como SOAP (Pai, Asnika, & R, 2014).

Debido al gran incremento de aplicaciones desarrolladas, para las compañías es muy complicado el hecho de tener que desarrollar la misma aplicación para diferentes plataformas. (Xanthopoulos & Xinogalos, 2013). Gracias al desarrollo de los servicios web REST, no es necesario realizar una programación exclusiva para cada plataforma, ni limitada por el lenguaje, el hardware la ejecución sino que es suficiente con exponer el servicio web para ser consumido por los clientes.

Teniendo como objetivo la creación de servicios web REST, empresas como Spring y Node.js, han desarrollado frameworks que facilitan el desarrollo para ser expuestos para su utilización.

Por un lado esta Spring, que es un framework de la plataforma de Java que provee soporte para el desarrollo de aplicaciones Java ya que se encarga del manejo de la infraestructura permitiendo que el desarrollador se enfoque en la aplicación.

Y por otro lado tenemos a Node.js, una plataforma de ejecución la cual permite ejecutar JavaScript de parte del servidor que al combinarse con Express.js que es un framework que funciona sobre Node.js permitiendo generar enrutamiento para trabajar sobre el protocolo HTTP.

Con el planteamiento anterior el problema de esta investigación es identificar cuál de los dos frameworks es el más óptimo que se debe utilizar al momento de empezar a desarrollar una aplicación web.

Para esto en este documento se va a realizar una comparación entre los dos frameworks, describiendo el desarrollo del backend de una aplicación web, usando herramientas para probar como Postman que es un cliente para servicios web, para publicar los servicios se va a usar el servidor node.js y Spring Boot.

2. Contenido

Para el desarrollo de la comparación de los frameworks Spring vs Node.js, se realiza un análisis de las principales características como son robustez, integración con otras tecnologías, seguridad, soporte, mantenimiento, estabilidad y ejecución. Tomando en cuenta la arquitectura de cada tecnología.

2.1. REST

Los servicios web basados en la arquitectura REST (Representational State Transfer), son una alternativa más simple a SOAP y a WSDL, la cual transmite los datos sobre el protocolo estandarizado HTTP. Varias empresas como Google, Facebook y Yahoo! son casos de éxito al migrar sus servicios a esta tecnología. REST fue publicado en el año 2000 por Roy Fielding en la Universidad de California en una conferencia en la cual presentaba acerca de principios arquitectónicos de software para usar a la web como una plataforma de procesamiento distribuido (Seta, 2008).

El cliente tiene acceso a los recursos usando una única URI y una representación del recurso es devuelto con un estado de la transferencia de la información. (Pai, Asnika, & R, 2014)

REST no está estrictamente relacionado con el protocolo HTTP, pero es con el que más comúnmente se asocia.

Los servicios web basados en REST cumplen los siguientes principios:

- **Utiliza métodos de HTTP de manera explícita:** Cada método tiene una función específica: GET su uso es para obtener un recurso desde el servidor, POST se usa para crear un nuevo recurso en él servidor, PUT actualiza o cambia el estado de un recurso, DELETE elimina un recurso.
- **REST no mantiene estado:** Es necesario tener un escalamiento para cubrir la demanda constante y en crecimiento mediante balanceador, clúster y servidores con alta disponibilidad con el objetivo de distribuir las peticiones entre los equipos para disminuir el tiempo de respuesta.
- **URI en formato de directorios:** La estructura de los URI se la crea de una manera similar a directorios que es fácil de entender lo cual permite poder utilizarlos de manera intuitiva.
- **Representaciones:** Los servicios web REST emiten una respuesta en formato JSON o XML los cuales pueden ser usados por cualquier aplicación desarrollada en cualquier lenguaje.

Existen diferentes herramientas que se pueden utilizar al momento de desarrollar un servicio REST, lo importante es tener en claro en que plataforma se va a desarrollar, si será un sistema distribuido, lo más importante es emplear tecnologías de similares características que tengan compatibilidad en el sistema distribuido. Este artículo se centra en dos importantes tecnologías JAVA y JAVASCRIPT, en el caso de JAVA nos vamos a centrar en Spring Boot y para JAVASCRPIT usaremos Node.js.

2.2. Spring Boot

Es un framework Java basado en el Modelo Vista Controlador, mediante el cual gracias a los componentes y librerías que brinda hace fácil el desarrollo y despliegue de los servicios REST. Ha eliminado la necesidad de configurar la aplicación con el uso de archivos XML haciendo énfasis en el desarrollo de la misma. (Toledano López, López del Castillo, & Vázquez Sanchez, 2018)

Como las principales características de Spring Boot tenemos que nos permite crear aplicaciones de Spring independientes, al poseer servidores de aplicación embebidos como Tomcat, Jetty o Undertow, no es necesario generar un WAR para su ejecución, también permite configurar automáticamente bibliotecas de Spring y de terceros, no es necesario realizar una configuración en archivos XML, por lo que es más fácil y rápido la integración con otros proyectos de Spring.

Para realizar la parte práctica de la implementación de una aplicación mediante el uso de Spring Boot vamos a usar como IDE para el desarrollo Spring Tool Suite Versión: 3.9.5.RELEASE, el lenguajede programación va a ser Java, JDK 1.8.0_171, para la gestión de las dependencias se va a usar Gradle 4.9 y como gestor de base de datos vamos a usar MariaDB 10.3.8.

A continuación se va a generar un ejemplo de la implementación con Spring Boot:

Configuramos el archivo principal: settings.gradle

```
rootProject.name = 'api-restful-spring'
```

Luego configuramos el proyecto con sus dependencias en el archivo: build.gradle

```
buildscript {
  repositories { mavenCentral() }
  dependencies {
     classpath("org.springframework.boot:spring-boot-gradle-plugin:
2.0.3.RELEASE") } }
plugins { id 'java-library' }
apply plugin: 'java'
apply plugin: 'eclipse'
apply plugin: 'idea'
apply plugin: 'org.springframework.boot'
apply plugin: 'io.spring.dependency-management'
bootJar { baseName='api-restful-spring' version='1.0.0-SNAPSHOT' }
sourceCompatibility = 1.8
targetCompatibility = 1.8
dependencies {
  api 'org.apache.commons:commons-math3:3.6.1'
  implementation 'com.google.guava:guava:23.0'
  implementation 'org.springframework.boot:spring-boot-devtools'
  implementation 'org.springframework.boot:spring-boot-starter-web'
  implementation 'org.springframework.boot:spring-boot-starter-data-
jpa'
  implementation 'org.springframework.boot:spring-boot-starter-data-
rest'
     implementation   'org.springframework.security:spring-security-
crypto'
  implementation 'org.projectlombok:lombok:1.18.2'
```

```
  implementation 'org.apache.commons:commons-lang3:3.7'
  implementation 'org.apache.commons:commons-collections4:4.2'
  implementation'mysql:mysql-connector-java'
  testImplementation 'junit:junit:4.12'
  testImplementation 'org.projectlombok:lombok:1.18.2'
   testImplementation 'org.springframework.boot:spring-boot-starter-
test'
}
repositories { jcenter() mavenCentral() }
```

Creamos la clase principal que permite ejecutar la aplicación, como podemos ver en la Figura 1.

```
1  /**
4  package com.isabellasoft.apirestfulmusifyspring;
5
6  import org.springframework.boot.SpringApplication;
8
9  /**
10  * @author Edward Haro
11  *
12  */
13
14 @SpringBootApplication
15 public class ApiRestfulMusifySpringApplication {
16     public static void main(String[] args) {
17         SpringApplication.run(ApiRestfulMusifySpringApplication.class, args);
18     }
19 }
20
```

Figura 1 – Configurar la aplicación ejecutable

En la Figura 2 vamos a mapear una entidad de la base de datos.

La Figura 3, nos muestra la implementación del controlador con el servicio REST de tipo GET.

Finalmente configuramos la conexión con la base de datos y las opciones del servidor que se va a levantar.

```
#Server configuraion
server.port=8180
#Database configuration
spring.datasource.url=jdbc:mysql://localhost:3306/musify
spring.datasource.username=eharo
spring.datasource.password=Password01
```

```java
/**
package com.isabellasoft.apirestfulmusifyspring.model;

import java.io.Serializable;

/**
 * @author Edward Haro
 *
 */
@AllArgsConstructor
@Data
@Entity
@Table(name = "user")
public class User implements Serializable {
    private static final long serialVersionUID = 6911959779305918597L;
    public static final String COLUMN_ID = "id";
    public static final String COLUMN_NAME = "name";
    public static final String COLUMN_EMAIL = "email";
    public static final String COLUMN_STATUS = "status";
    public static final String COLUMN_PASS = "password";
    public static final String COLUMN_SURNAME = "surname";
    @Id
    @GeneratedValue(strategy = GenerationType.IDENTITY)
    private Integer id;
    @Column(name = COLUMN_NAME)
    private String name;
    @Column(name = COLUMN_SURNAME)
    private String surname;
    @Column(name = COLUMN_EMAIL)
    private String email;
    @Column(name = COLUMN_PASS)
    private String password;
    @Column(name = COLUMN_STATUS)
    private String status;
}
```

Figura 2 – Mapeo de entidades

```java
import org.slf4j.Logger;

/**
 * @author Edward Haro
 *
 */
@RestController
@RequestMapping("/musify")
public class UserController {
    private static final Logger LOG = LoggerFactory.getLogger(UserController.class);
    @Autowired UserService userService;

    @GetMapping(value = UserMappingRoutes.GET_USER_BY_ID)
    public ResponseEntity<Object> getUserById(@PathVariable(name = User.COLUMN_ID) String id) {
        try {
            User optionalUser = userService.getUserById(id);
            if (optionalUser != null) {
                return new ResponseEntity<>(optionalUser, HttpStatus.OK);
            }
            return new ResponseEntity<>("No existe un usuario para el id: ".concat(id), HttpStatus.OK);
        } catch (Exception e) {
            LOG.error(e.getMessage());
            return new ResponseEntity<>(e.getMessage(), HttpStatus.INTERNAL_SERVER_ERROR);
        }
    }
}
```

Figura 3 – Controlador para crear el Web service REST de tipo GET

Como podemos ver en cinco pasos tenemos levantado un servicio web REST de tipo GET para obtener recursos desde la base de datos.

4. Node.js

Node es un entorno de ejecución orientado a eventos asíncronos de JavaScript diseñado para construir aplicaciones de red escalables (Node.js, 2018).

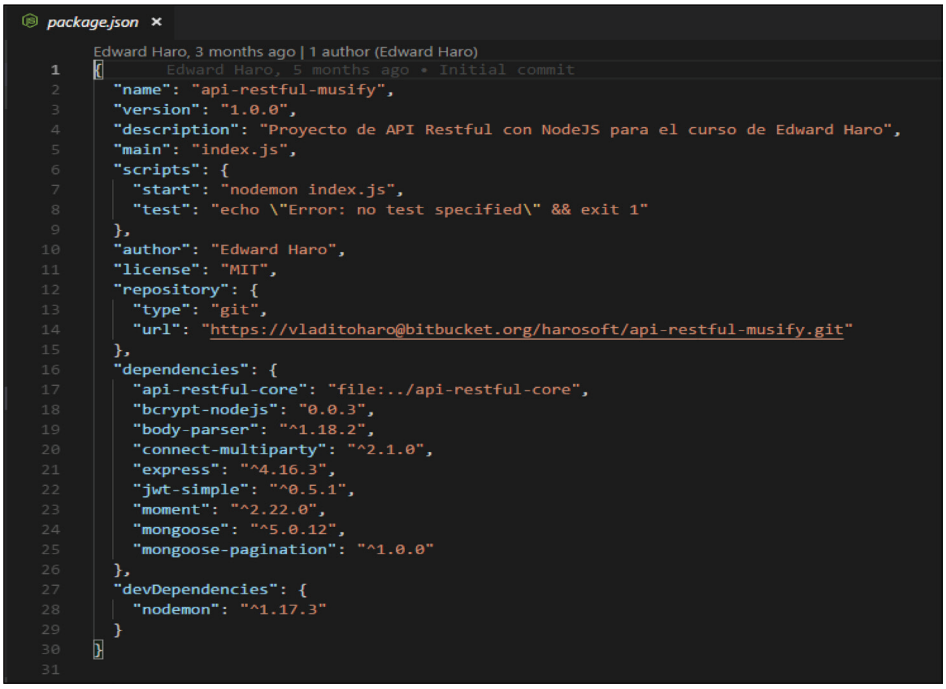


Figura 4 – Configuración de dependencias

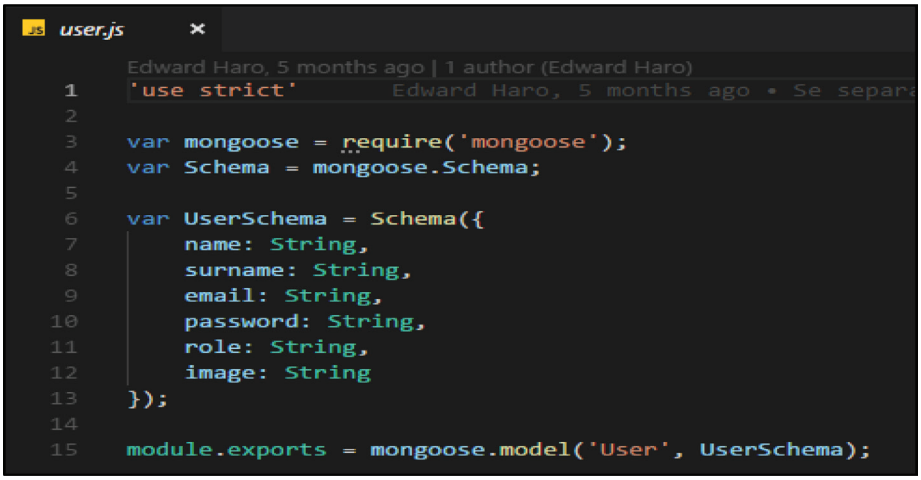


Figura 5 – Mapeo de entidades

Como las principales características de Node.js tenemos que permite desarrollar aplicaciones de red escalables, impulsada por eventos asíncronos, si no hay trabajo por hacer Node.js permanece dormido hasta que haya una conexión para que se active una llamada al servicio, todo esto es posible ya que se usa JavaScript del lado de servidor. (w3schools.com, 2018)

Para realizar la parte práctica de la implementación de una aplicación mediante el uso de Node.js vamos a utilizar como IDE de desarrollo Visual Studio Code Versión: 1.25.1, como lenguaje de programación JavaScript en la plataforma de Node.js Versión 8.11.2, para gestionar las dependencias que se va a usar para el desarrollo del proyecto utilizaremos NPM Versión 6.4.0 y finalmente como gestor de base de datos MongoDB 3.6.

A continuación se va a generar un ejemplo de la implementación con Node.js:

En la Figura 4 se muestra las dependencias con las que se va a levantar el servicio REST con la configuración principal del proyecto.

En la Figura 5 se va a realizar el mapeo del documento que se genera en la base de datos para realizar las transacciones.

A continuación se debe configurar la aplicación principal como se muestra en la Figura 6.

```
You, a few seconds ago | 2 authors (You and others)
'use strict'      Edward Haro, 5 months ago • Se separa en clases en lugar de solo funciones
const express = require('express');
const bodyParser = require('body-parser');
const UserRoutes = require('../routes/UserRoute');
You, a few seconds ago | 2 authors (You and others)
class Application {
    constructor() {
        this.application = express();
        this.init();
    }
    init() {
            this.loadRoutes();
            this.configBodyParser();
            this.configMiddlewareCabeceras();
            this.buildRoutes();
        }
        //cargarRutas
    loadRoutes() { this.userRoutes = new UserRoutes(); }

    //Configuar BodyParser
    configBodyParser() {
            this.application.use(bodyParser.urlencoded({ limit: '50mb', extended: true }));
            this.application.use(bodyParser.json({ limit: '50mb' }));
        }
        //Configuramos un middleware para las cabeceras, next para salir del metodo
    configMiddlewareCabeceras() {
            this.application.use((req, res, next) => {
                res.header('Access-Control-Allow-Origin', '*');
                res.header('Access-Control-Allow-Headers', 'Authorization, X-API-Key, Origin, X-Request-With, Content-Type, Accept,
                res.header('Access-Control-Allow-Methods', 'GET, POST, OPTIONS, PUT, DELETE');
                res.header('Allow', 'GET, POST, OPTIONS, PUT, DELETE');
                next();
            });
        } //rutas base
    buildRoutes() { this.application.use('/api', this.userRoutes.api); }
}
module.exports = Application;
```

Figura 6 – Creación de la aplicación principal

Una vez que se ha configurado la aplicación principal, se debe realizar la implementación del server que va a conectarse con la base de datos y levantar la aplicación para ser usada en un puerto especificado en el código, tal y como lo podemos ver en la Figura 7

En la Figura 8 se muestra la configuración de las rutas que se van a poder configurar para el acceso a los servicios REST.

```
You, a few seconds ago | 2 authors (Edward Haro and others)
'use stric'
const mongoose = require('mongoose');
const Application = require('../core/Application');

You, a few seconds ago | 2 authors (Edward Haro and others)
class Server {
    constructor() {
        this.port = null;
        this.mongooseConnect = mongoose;
        this.app = new Application();
        this.init();
    }
    init() {
        this.port = process.env.PORT || 3977;
        this.connectMongoose();
    }
    connectMongoose() {
        try {
            this.mongooseConnect.Promise = global.Promise;
            this.mongooseConnect.connect('mongodb://localhost:27017/app-musify', (err, res) => {
                if (err) {
                    throw err;
                } else {
                    console.log('La conexion a la base de datos esta correcta.');
                    this.app.application.listen(this.port, () => {
                        console.log('El Servidor del API Rest de musica escuchando en http://localhost:' + this.port);
                    });
                }
            });
        } catch (err) { console.log(err); }
    }
}
module.exports = Server;
```

Figura 7 – Creación de servidor y conexión con la base de datos

```
UserRoute.js
Edward Haro, 5 months ago | 1 author (Edward Haro)
'use stric'      Edward Haro, 5 months ago • Se separa en clases en lugar de solo funciones

const express = require('express');
const UserController = require('../controllers/UserController');
const Authenticated = require('../middlewares/Authenticated');

const multipart = require('connect-multiparty');
/**
 * Maneja las rutas definidas utilizando expres
 */
Edward Haro, 5 months ago | 1 author (Edward Haro)
class UserRoute {
    constructor() {
        this.api = express.Router();
        this.userController = new UserController();
        //Middleware para subir ficheros, la ruta hace referencia desde la raiz, tener en cuenta
        this.md_upload = multipart({ uploadDir: './uploads/users' });
        this.loadRoutes();
    }

    loadRoutes() {
        this.api.get('/pruebaController', Authenticated.INSTANCE.ensureAuth, this.userController.pruebas);
        this.api.get('/get-image-user/:imageFile', this.userController.getImageFile);

        this.api.post('/register', this.userController.saveUser);
        this.api.post('/login', this.userController.loginUser);
        this.api.post('/upload-image-user/:id', [Authenticated.INSTANCE.ensureAuth, this.md_upload], this.userController.uploadImage

        this.api.put('/update-user/:id', Authenticated.INSTANCE.ensureAuth, this.userController.updateUser);

    }

}

module.exports = UserRoute;
```

Figura 8 – Configuracion de las rutas para el servicio REST

Finalmente se debe configurar un controlador que va a tener la lógica necesaria que los servicios REST van a proveer. Con el código que se muestra a continuación.

```
'use strict'
const fs = require('fs');
const path = require('path');
const User = require('../models/user');

/**
 * Controlador para los relacionado con las transacciones a la base de datos para
el documento Usuario
 * @author eharo
 */
class UserController {
    constructor() {}
        /** Actualizar usuario */
    updateUser(req, res) {
        let userId = req.params.id;
        let update = req.body;
        if (userId != req.user.sub) {
            return res.status(500).send({ message: "No tiene permisos para
actualizar este usuario" });
        }
        User.findByIdAndUpdate(userId, update, (err, userUpdated) => {
            if (err) {
                res.status(500).send({ message: "Error al actualizar el usuario"
});
            } else {
                if (!userUpdated) {
                    res.status(404).send({ message: "No se ha podido actualizar al
usuario" });
                } else {
                    res.status(200).send({ user: userUpdated });
                }
            }
        });
    }
}
module.exports = UserController;
```

5. Spring vs Node.js

En esta sección se presenta un estudio comparativo entre spring framework y node.js. En este caso vamos a comparar un framework desarrollado en Java como Spring Boot y un framework hecho en JavaScript como Node.js, a continuación se muestra una analogía entre el entorno de ejecución de Java y JavaScript.

En la tabla 1 se realiza una comparación entre el entorno de ejecución de java que se ejecuta en la Máquina Virtual Java, mientras que Node.js se ejecuta en la plataforma V8. El funcionamiento es muy similar ya que ambas tecnologías funcionan con librerías que realizan funciones específicas.

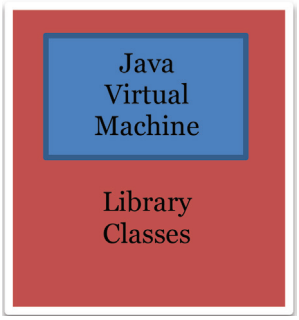
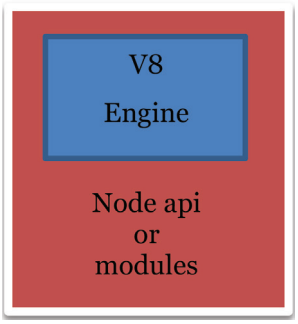
Java Runtime Environment	Node.js
	

Tabla 1 – Entorno de ejecución de Java vs Node.js

¿Por qué Spring Boot?	¿Por qué Node.js?
Es un proyecto creado a partir de Spring. Permite desarrollar y arrancar de manera rápida y sencilla, aplicaciones basadas en Spring. Se crea una clase principal que va a ser la encargada de ejecutar la aplicación. Las configuraciones se hacen a través de anotaciones. En un momento se puede tener una aplicación totalmente funcional. Open Source Nivel empresarial. Facil para realizar debug en excelentes IDE's como Eclipse, IntelliJ. Gran cantidad de librerías. Excelente seguridad. Excelente para desarrollo REST API. Empaquetamiento a través de Maven o Gradle.	Puede generar páginas con contenido dinámico. Puede crear, abrir, leer, escribir, eliminar y cerrar archivos en el servidor. Puede agregar, eliminar, modificar registros de la base de datos. Los archivos .js contienen métodos que se ejecutaran de acuerdo a ciertos eventos. Fácil y sencillo de aprender e implementar. En la actualidad es considerado lo mejor para el desarrollo de REST API. Poco código para mantener. Fácil de testear con aplicaciones con plugins como Mocha/Chai. Empaquetamiento a través de NPM. Modelo de no bloqueo de I/O. Trabaja con casos síncronos y asíncronos. Fácil integración con bases de datos No SQL.

Tabla 2 – Por qué usar cada tecnología

En la tabla 2 podemos comparar características de cada framework en las que podemos ver características de cada framework con las funcionalidades que cada tecnología ofrece.

En la tabla 3 se encuentra el detalle de conexión entre cada tecnología con las bases de datos disponibles.

Spring Boot + Bases de datos	Node.js + Bases de datos
Provee un extenso soporte para trabajar con bases de datos SQL.	Provee un extenso soporte para trabajar con bases de datos SQL.
Si es un almacenamiento relacional.	Cuando la base de datos almacena documentos.
Bases de datos como: MariaDB, MySql, Postgres, Oracle, DB2.	Bases de datos como MongDB, Cassandra, CouchDB, Redis.

Tabla 3 – Relacion con las bases de datos

En la tabla 4 se realiza una comparación de cada tecnología al momento de integrar seguridad en las aplicaciones que desarrollan con cada tecnología.

Spring Boot + Seguridad	Node.js + Seguridad
Se puede usar una gran cantidad de librerías que manejan la seguridad como Kerberos, Spring Security, Token.	Los paquetes disponibles para el manejo de la seguridad son pocos.

Tabla 4 – Con respect a la seguridad

6. Conclusiones

En base a la información recolectada, las tablas comparativas y los ejemplos realizados podemos sacar las siguientes conclusiones: las dos tecnologías son excelentes al momento de desarrollar servicios REST. Ya que las dos son fáciles de implementar y levantan un servicio REST en pocos minutos dependiendo de la funcionalidad que vaya a brindar; Spring tiene un nivel de madurez de varios años más sobre Node.js. Ya que el desarrollo del framework se ha basado en Java; Spring es más robusto al momento de hablar de seguridad, ya que posee su propio conjunto de aplicaciones para seguridad como es Spring Security; Spring Boot tiene integración con bases de datos transaccionales con las cuales se puede realizar aplicaciones empresariales debido a la posibilidad de utilizar las transacciones de las bases de datos para asegurar la integridad de la información.

Referencias

Hamad, H., Saad, M., & Abed, R. (2010). Performance Evaluation of RESTful Web Services for Mobile Devices. *Int. Arab J. e-Technol, 1*(3), 72–78.

Node.js, F. (2018, 8 26). *About Node.js®*. Retrieved from https://nodejs.org/es/about/: https://nodejs.org/es/about/

Pai, M. V., Asnika, S., & R, S. (2014). RESTful Web Services. *International Journal of Advanced Information Science and Technology (IJAIST), 24*(24), 46–50.

Rodriguez, A. (2008, 11 6). Restful web services: The basics. *IBM developerWorks*, 1-11. Retrieved from http://www.gregbulla.com/TechStuff/Docs/ws-restful-pdf.pdf

Seta, L. D. (2008, 11 13). *Introducción a los servicios web RESTful*. Retrieved from Dos Ideas: https://dosideas.com/noticias/java/314-introduccion-a-los-servicios-web-restful

Toledano López, O. G., López del Castillo, D. C., & Vázquez Sanchez, Á. A. (2018). Capa de Servicios para la Plataforma de Procesamiento de Datos Educativos Masivos de La Faultad 4 de la Universidad de las Ciencias Informáticas. La Habana.

w3schools.com. (2018, 08 26). *Node.js Tutorial*. Retrieved from https://www.w3schools.com/nodejs/nodejs_intro.asp: https://www.w3schools.com/nodejs/nodejs_intro.asp

Xanthopoulos, S., & Xinogalos, S. (2013, Septembe). A comparative analysis of cross-platform development approaches for mobile applications. *In Proceedings of the 6th Balkan Conference in Informatics*, 213–220.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 13/09/2018
Aceitação/Acceptance: 15/11/2018

Estudio de la Identidad Federada y su Aplicación en el Ecuador

Gabriel Patricio Luna Arias[1], Teresa Guarda[2,1,3], Lucía Liliana Valencia Quezada[1], Geovanni Ninahualpa Quiña[1]

gabriel.luna@itdospuntocero.net, tguarda@gmail.com, lilyanny3@gmail.com, gninahualpa@espe.edu.ec

[1] Universidad de las Fuerzas Armadas ESPE, Centro de Postgrados, 170511, Quito, Ecuador

[2] Universidad Estatal Península de Santa Elena, Santa Elena, Ecuador – UPSE, Av. La Libertad, Santa Elena, Ecuador

[3] Algoritmi Centre, Minho University, Guimarães, Portugal

Pages: 322–335

Resumen: Internet aloja la mayoría de los servicios a los que acceden personas y empresas. Minimizar los problemas de privacidad es uno de los tópicos más importantes en el entorno tecnológico actual y la percepción de que la información es accesible a terceros no autorizados afecta negativamente al usuario, por lo que aún existe mucha desconfianza en la sociedad ecuatoriana con respecto a las transacciones en la web. La gestión de identidades es fundamental en las organizaciones y es una tarea costosa que requiere tiempo y personal calificado para su implementación. Las infraestructuras de telecomunicaciones buscan ser más distribuidas y seguras, pero se vuelven más complejas. Con la llegada de nuevas tecnologías como la computación en la nube, aparecen nuevos requisitos y problemas de seguridad. Este estudio revisa las características para mejorar la confidencialidad de los datos del usuario, presenta los conceptos en los que se basa la identidad federativa, las arquitecturas existentes y propone el uso de esta tecnología en el país y las perspectivas futuras. Las conclusiones de este documento describen las implicaciones de esta tecnología en Ecuador.

Palabras-clave: Administración de Identidad; Privacidad en Internet; SAML 2.0.

Study of the Federated Identity and its Application in Ecuador

Abstract: Internet hosts most of the services that people and companies access. Minimizing privacy problems is one of the most important topics in the current technological environment and the perception that the information is accessible to unauthorized third parties negatively affects the user, so there is still much distrust in Ecuadorian society regarding the transactions on the web. Identity management is fundamental in organizations and is a costly task that requires time and qualified personnel for its implementation. Telecommunications infrastructures seek to be more distributed and secure, but they become more complex. With the arrival of new technologies such as cloud computing, new requirements and security problems

appear. This study reviews the characteristics to improve the confidentiality of user data, presents the concepts on which the federative identity is based, the existing architectures and proposes the use of this technology in the country and future perspectives. The conclusions of this document describe the implications of this technology in Ecuador.

***Keywords**:* Internet; Identity; Distributed; Ecuador

1. Introducción

La computación en la nube ha irrumpido en el mundo de la tecnología y las escenas de negocios, prometiendo enormes ventajas técnicas y económicas. Uno de las principales ventajas de la computación en la nube es que representa un modelo de informática de servicios tanto públicos como privados, capaz de ofrecer aprovisionamiento bajo demanda de recursos informáticos.

El presente estudio enfatiza en la autenticación que es un paso esencial y busca identificar con precisión al usuario que realiza una solicitud. Tradicionalmente, cada proveedor de servicios administra las identidades de los usuarios por separado. Esto complica a la mayoría de usuarios ya que tienen que manejar múltiples credenciales y la carga de trabajo del personal de help desk se eleva por peticiones de recuperación de claves (Villarreal, Villarreal, Westphall, & Werner, 2017). Como una solución a estos problemas, el concepto de gestión federada de la identidad ha surgido en los últimos años. Eso es un término general que actualmente se refiere a un conjunto de tecnologías y estándares que permiten a los usuarios interactuar con muchos servicios iniciando sesión en solo uno, donde cada servicio puede exigir prueba de autenticación y autorización de la transferencia de datos personales. Actualmente, OAuth es el protocolo principal para transferir datos de usuario entre proveedores de identidad como Google, Twitter, Facebook que alojan datos de usuario y las entidades que requieren dicha información. OAuth tiene entre sus defectos que no ofrece ninguna vinculación entre el IdP (Identity Provider) y los SP (Service Provider) en términos de solicitudes de identidad o datos de un usuario y el IdP se considera confiable para proporcionar información autorizada en nombre del usuario, adicionalmente un IdP antiético puede recopilar datos sobre el usuario al registrar las transacciones entre él y los SP. Los SP también vinculan al usuario con su IdP, impidiendo el uso seudónimo o anónimo de su servicio.

Considerando estas razones, es imperativo contar con mecanismos de seguridad más avanzados que permitan a los usuarios beneficiarse de un proveedor de identidad en la nube no necesariamente confiable, y aun así preservar su privacidad y el control sobre su información. La identidad federada permite ir más allá de un entorno SSO ya que brinda privacidad y confiabilidad a la autenticación. (Jensen & Jaatun, 2013)

2. Conceptos y Motivaciones

2.1. Conceptos

El continuo desarrollo de Internet sigue acercando a las personas y generando productos y servicios. Dentro del mundo virtual, se han desarrollado nuevas herramientas que han cambiado los conceptos relacionados con las transacciones. Para establecer relaciones, es

necesario que las partes involucradas tengan algún conocimiento el uno del otro. Conocer la otra parte es importante para saber qué nivel de confianza se puede depositar en ella y va a cumplir o no su parte en la transacción. En las relaciones tradicionales, donde hay algún tipo de contacto directo entre las partes, el intercambio de información ocurre naturalmente.

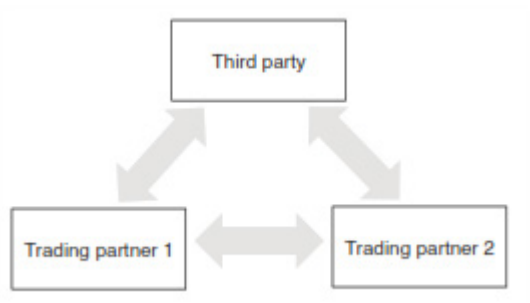


Figura 1 – Relación comercial que involucra a un intermediario

En los nuevos servicios proporcionados por la infraestructura de red, es aún más importante conocer a la otra parte involucrada en la comunicación o las transacciones; la respuesta a preguntas relacionadas con el socio comercial o proveedor de servicios involucra conceptos sobre entidad, identidad, identificador y credencial. Conocer estos conceptos permite comprender cómo funcionan las operaciones de autenticación y autorización, y cómo se pueden aplicar en la provisión de servicios futuros. En una transacción, debe haber cierto nivel de confianza entre las partes, si no puede establecerse directamente, es necesario interactuar a través de un intermediario de confianza para ambas partes, como se muestra en la figura 1. En las relaciones comerciales tradicionales, por ejemplo, esta función la realiza una institución financiera en caso de compras con cheque o tarjeta de crédito. En el mundo virtual, muchos intercambios se realizan sin el contacto directo entre las partes. En este caso, para establecer relaciones de confianza, es importante involucrar a un tercero. Este tercero es responsable de proporcionar información sobre el socio, implementando el concepto de identificación.

Las identidades de las partes involucradas son controladas y monitoreadas por un sistema IdM (Identity Management). Un IdM es un marco que gestiona una colección de identidades, su autenticación, su uso y la información relacionada con la identidad. Los sistemas IdM tradicionales se utilizan como mecanismos de autenticación y autorización, siendo el primer mecanismo responsable de establecer la confianza entre el sistema y la identidad, y el segundo responsable de proporcionar los permisos de identidad para realizar ciertas acciones en el sistema. Sin embargo, se han propuesto nuevos modelos, algunos optimizados para cumplir los objetivos del usuario, otros optimizados para abordar problemas relacionados con la infraestructura de red o los requisitos de aplicaciones y servicios.

2.1.1. Entidad, identidad, identificador y credencial

Como la definición del sistema está involucrada en los siguientes términos, se define previamente como un conjunto de software y hardware que almacena información

sobre entidades, credenciales e identidades y procesa las operaciones relacionadas con sus ciclos de vida. Un sistema generalmente protegerá datos o servicios, y por lo tanto empleará autenticación y autorización.

Una entidad puede ser una persona, un servicio de red, un dispositivo informático de red o un dispositivo de teléfono móvil. Las entidades existen en el mundo real. Usan credenciales y tienen un ciclo de vida separado de cualquier identidad asociada, identificador o ciclo de vida de credenciales. Una credencial se usa para probar una identidad de un sistema. Puede haber varios tipos de credenciales, pero todas están involucradas en garantizar un sistema que una entidad realmente tenga el derecho de usar una identidad particular. Algunas credenciales se establecen entre la entidad y el sistema y algunas son emitidas por terceros.

La identidad es un instrumento utilizado por una entidad para proporcionar información sobre sí mismo al sistema. Una identidad es un concepto virtual y no existe en el mundo real. Siempre está asociado con una entidad o generalmente formado por un identificador único. Este identificador se usa para demostrar la propiedad de la identidad, a través de credenciales, y más información sobre la cual un sistema puede basar las decisiones que toma sobre la entidad asociada. Una entidad puede tener varias identidades en el mismo sistema para ser utilizada con diferentes propósitos. Por ejemplo, un usuario puede tener una identidad de usuario común y otra como contador de la empresa. Un identificador es un índice único para una identidad. Típicamente, un identificador es utilizado por un sistema para referirse a una identidad. Debe ser único dentro de cualquier sistema dado, pero puede reutilizarse en varios. La Fig. 2 muestra la relación entre entidad, identidad e identificador.

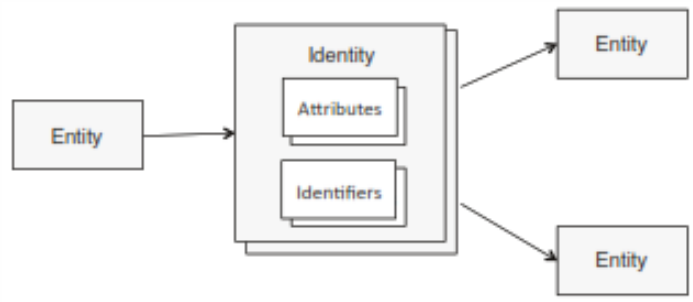


Figura 2 – Relación entre entidad, identidad e identificador

Los conceptos de identificadores y credenciales pueden ser confusos y, en algunos casos, incluso pueden usarse uno en vez del otro. La principal diferencia entre identificadores y credenciales es el hecho de que un identificador debe ser necesariamente único y la credencial no. Incluso siendo único, el identificador puede ser la unión de otros elementos no únicos.

2.1.2. Ciclo de vida

Existe un ciclo de vida definido para credenciales, y de hecho para entidades, identidades e identificadores. Este ciclo de vida abordará lo que sucede en las diferentes fases:

nacimiento, vida y muerte. El ciclo de vida de la entidad implica los mismos conceptos del ciclo de vida de las personas. La dificultad está en sincronizar sus identidades y credenciales asociadas con esto. Sin embargo, hay varios umbrales y etapas intermedios, con sus propios problemas. Por ejemplo, diferentes reglas se aplican a ciertas edades, como votar a los 16 y conducir a los 18 años. Las entidades también pueden ser máquinas o teléfonos móviles. En este caso, la credencial se crea en la fabricación o se activa al momento de la compra.

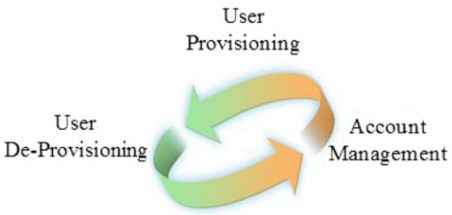


Figura 3 – Ciclo de vida de la identidad

En el ciclo de vida de la identidad, las identidades están asociadas a las entidades. La etapa de nacimiento generalmente es el registro de una entidad en un sistema. La etapa de la muerte puede no ser una cesación simple, ya que es posible que sea necesario conservar los registros después de que se finalice una cuenta. La etapa de la vida puede incluir actualizaciones y suspensiones de credenciales. Las identidades pueden terminar incluso si la entidad asociada todavía está viva. Por ejemplo, una persona puede cerrar una cuenta bancaria o conservarla incluso si la entidad se ha extinguido, a fin de mantener los registros. El ciclo de vida del identificador sigue el ciclo de vida de la identidad asociada. Sin embargo, los identificadores pueden conservarse indefinidamente para garantizar que no se reasignen, lo que podría causar confusión o permitir el fraude más adelante. Las credenciales tienen su propio ciclo de vida, bastante separado del ciclo de vida entidad / identidad asociada. Las credenciales son emitidas a las entidades por los proveedores, o a los menos acordados entre ellas. Un proveedor podría ser un sistema, emitir credenciales para permitir que sus usuarios sean autenticados, o un tercero de confianza en el que tanto el sistema como la entidad confíen, para proporcionar un enlace adecuado para la autenticación entre las dos partes. A veces las credenciales necesitan mantenimiento, como cambios regulares de una contraseña, renovación de certificados digitales, pasaportes o licencia de conducir (vida). Las credenciales pueden morir por vencimiento o revocación (muerte). También pueden cancelarse temprano debido a algún problema, ya que la entidad ya no existe, o una credencial comprometida olvidada o descubierta por un adversario. La dificultad con la revocación es que la credencial misma ya no transporta información de validez suficiente por sí misma, por lo tanto, son necesarias las fuentes externas, como las listas de revocación de certificados para verificar su validez. Si estas fuentes no están disponibles, habrá una alta probabilidad de que una credencial no se muera. En este caso, será imposible resolver las disputas sobre su uso histórico. (Habiba, Masood, Awais Shibli, & A Niazi, 2014)

2.1.3. Operaciones

Existe un ciclo de vida definido para credenciales, y de hecho para entidades, identidades e identificadores. Este ciclo de vida abordará lo que sucede en las diferentes fases:

nacimiento, vida y muerte. Las cuatro operaciones principales de un sistema IdM son: identificación, autenticación, autorización y contabilidad. La identificación es la acción de una entidad que proporciona una identidad al sistema a través de un identificador. La identificación le dice al sistema quién podría estar tratando de conectarse, pero no ofrece ninguna prueba. Por lo tanto, el próximo paso para el sistema es obtener una prueba de esta identidad y verificarla. Para obtener una prueba de esta identidad, el sistema realiza el proceso de Autenticación donde la entidad de conexión proporciona credenciales para la identidad reclamada, conocida solo. Las credenciales no son necesariamente únicas. Por ejemplo, varios usuarios pueden casualmente tener la misma contraseña, pero solo un usuario en particular debe saber cuál es su contraseña. Una vez completado el paso de autenticación, el sistema puede estar seguro de que la entidad es el propietario legítimo de la identidad elegida. Cada entidad tendrá un conjunto de acciones permitidas en el sistema. Estos permisos se conocen como privilegios y se determinan en la etapa de Autorización. El sistema debe buscar los privilegios relevantes para la identidad autenticada. Por lo tanto, debe haber un enlace de identidad a los privilegios. Esta información de enlace debe estar disponible para el sistema, y puede obtenerse de una base de datos o contener una credencial o documento provisto por la entidad. También hay un concepto adicional de Contabilidad, que es una grabación de lo que sucedió una vez que se otorgó la autorización. Consiste en un conjunto de registros, cada uno vinculando una identidad establecida con una acción.

2.1.4. Usuario, proveedor de identidad y proveedor de servicios

En un sistema IdM general, podemos identificar tres entidades involucradas: usuario, proveedor de identidad y proveedor de servicio. Un usuario (U) es una entidad que utiliza un servicio proporcionado por un proveedor de servicios. Los usuarios usan sistemas IdM para acceder a los servicios que requieren la certificación de sus atributos por un tercero. Esto es común en el acceso a Internet porque los usuarios no confían en la seguridad de la transmisión de datos. Por lo tanto, IdM busca aumentar la confianza en los procesos que involucran Internet. El proveedor de identidad (IdP) es una entidad que controla las credenciales del usuario y proporciona servicios de autenticación. Las entidades que proporcionan cierto grado de identificación, como las Autoridades de Certificación (CA), generalmente brindan algunos servicios que cumplen con la definición de los sistemas IdM. Dichos servicios generalmente integran una cadena de seguridad requerida por algún tipo de negocio en línea, como un servidor de autenticación o cifrado SSL.

El Proveedor de servicios (SP) es una entidad que ofrece uno o más servicios a los que acceden usuarios potenciales y utiliza los servicios IdP para autenticar a un usuario. Las razones por las que los SP implementan un sistema IdM coinciden, en parte, con las razones por las cuales los usuarios deben usarlas. Las organizaciones también valoran las oportunidades para aumentar la seguridad. Además, es importante que ambas partes estén seguras sobre la autenticidad del socio de comunicación. Particularmente, cuando un fraude puede prevenirse utilizando este medio. Las reglamentaciones, como la protección de la privacidad, deben ser respetadas por los SP. Un sistema IdM puede ayudar a una organización a garantizar y simplificar el cumplimiento de la legislación. A fin de influir en la cantidad y la calidad de los datos transmitidos, la cantidad de información personal proporcionada por el usuario debe reducirse mediante el uso de seudónimos y credenciales. Los seudónimos son identificadores de la entidad.

2.2. Escenarios de Aplicación

El continuo desarrollo de Internet sigue acercando a las personas y generando productos y servicios. Consideremos dos escenarios del mundo real que requieren autenticación y que en un ambiente de identidad federada serían sumamente sencillos, un banco en línea y un cine en línea. El llegar a este nivel de integración tecnológica debe ser nuestra meta.

El primer escenario, es decir, un banco en línea, requiere una divulgación completa de información personalmente identificable; mientras que el cine en línea debe verificar solo la edad del usuario y la ubicación geográfica.

Banca en línea: Considere a un usuario que quiere abrir una cuenta bancaria con un banco en línea. En lugar de ir en persona para identificarse, el usuario desea usar su inicio de sesión federado emitido por una autoridad, es decir, el proveedor de identidad.

Para que una solicitud de apertura de cuenta sea legal, el banco debe obtener los datos completos del usuario, incluidos su nombre completo y dirección actual. En tal escenario, el usuario debe divulgar casi toda su información privada requerida por el banco para poder usar el inicio de sesión federado para acceder a su cuenta bancaria.

Cine en línea: En nuestro segundo escenario, consideramos a un usuario que desea acceder a servicios en línea que no requieren todos sus datos, sino una prueba de que el usuario tiene más de cierta edad y no es un robot. Para acceder al servicio, en lugar de divulgar su información personal, para preservar su privacidad, solo desea proporcionar pruebas de que tiene más de la edad requerida para la película seleccionada y de que está registrada con cierta autoridad.

Como se aprecia en ambos ejemplos dependiendo de la aplicación o servicio al cual el usuario quiera acceder se comparte cierta cantidad de información con las plataformas de las entidades interesadas (Alkhalifah & Al Amro, 2015).

3. Arquitecturas de Identidad Federada

El enfoque principal de los sistemas analizados es proporcionar autenticación federada al tiempo que se protege la privacidad de los usuarios, de modo que el IdP no pueda correlacionar la solicitud de autenticación con el SP de destino. Más específicamente, el IdP no puede evaluar el comportamiento en línea del usuario al rastrear las solicitudes de autenticación.

Desvinculación del perfil. El IdP no debería poder correlacionar las solicitudes de autenticación con los usuarios correspondientes. En otras palabras, el IdP no debería ser capaz de evaluar el comportamiento del usuario.

Divulgación selectiva.- Un proveedor de servicios no debe aprender nada más allá de la información que se requiere para proporcionar el servicio. En algunos casos, esta información puede ser requerida para la verificación de elegibilidad, por ejemplo, requisito de edad.

No suplantación.- La propiedad de no suplantación requiere que nadie, excepto el usuario que está en posesión de un secreto, pueda ser autorizado. El IdP tampoco debería ser capaz de hacerse pasar por sus usuarios.

Despliegue.- El sistema debe ser desplegable desde una perspectiva técnica. Dadas las diversas soluciones de identidad federada implementadas (OpenID y SAML), los sistemas se deben implementar incrementalmente como un complemento de la infraestructura existente.

Teniendo estos puntos fundamentales en cuenta, se analizará algunas soluciones desarrolladas (Vapen, Carlsson, Mahanti, & Shahmehri, 2016).

3.1. UnlimitID

UnlimitID ofrece la creación de pseudoidentidades desvinculables que están vinculadas a un RP con el uso de credenciales de múltiples presentaciones.

Este sistema admite la divulgación selectiva de atributos (por ejemplo "mostrar edad") sin la necesidad de revelar el seudónimo, de modo que los atributos se puedan verificar sin una pseudoidentidad persistente, evitando efectivamente la recolección de atributos a lo largo del tiempo por el SP o el IdP. Además, un IdP puede permitir a los usuarios tener múltiples pseudo-identidades con un SP, cada uno de ellos con diferentes atributos. Al igual que OAuth, no define los métodos de autenticación ni cómo almacenar y sincronizar de manera segura los dispositivos. No controla los valores de los atributos transferidos, por lo que un usuario puede transferir valores que reducen su conjunto de anonimato o que se identifican de manera única. Se asume la confidencialidad de los tokens en almacenamiento y en tránsito, así como el anonimato del tráfico de red (como el que proporcionan las redes proxy anónimas) y un entorno de usuario seguro con primitivas criptográficas confiables.

En pocas palabras, UnlimitID logra la desvinculación en dos fases:

Primero, un usuario se autentica en el IdP y se le emite una credencial que codifica una clave secreta a largo plazo y un par clave / valor que denota un atributo. La credencial también incluye una fecha de vencimiento después de la cual tendrá que volver a emitirse. Se supone que todos los problemas de credenciales contienen cuatro campos: una clave secreta Sub desconocida para el IdP, un par de atributos Key y Value y un tiempo de expiración Exp.

En una segunda fase, el usuario ciega la credencial y la regresa de nuevo en el IdP para generar una pseudo identidad específica para un servicio de SP. En este punto, el IdP actúa en su rol normal en el protocolo OAuth, utilizando la pseudoidentidad y los atributos no rastreables para respaldar la autenticación, y el usuario puede autenticarse en el SP utilizando el flujo estándar de OAuth (Isaakidis, Halpin, & Danezis, 2016).

3.2. BlindIdM

Es definido como un sistema IDaaS que garantiza la privacidad y el control del usuario incluso cuando el almacenamiento y el procesamiento de datos se realizan mediante nubes que no son de confianza necesariamente.

Preserva la privacidad, donde el proveedor de identidad en la nube puede ofrecer un servicio de información de identidad sin conocer la información personal real de los usuarios. El sistema usa SAML 2.0 como protocolo de gestión de identidad subyacente y el reencriptado de proxy como un medio para lograr un manejo ciego de la información

de identidad; De esta forma, el proveedor de la nube transforma los atributos encriptados por la organización anfitriona en textos cifrados para el servicio.

BlindIdM también cumple con la visión de IDaaS, donde las competencias se distribuyen en el ecosistema de la nube debido al enfoque híbrido que estamos utilizando, donde la autenticación se lleva a cabo localmente por la organización anfitriona y los mecanismos de protección de datos (basados en el reencriptado del proxy) se distribuyen en todas las entidades involucradas. Además, todas las entidades siguen el protocolo de gestión de identidades y la organización anfitriona actúa como la fuente autorizada de información. Puede soportar otras variaciones, como la externalización del almacenamiento de atributos cifrados a un proveedor especializado de almacenamiento en la nube; el proveedor original de identidad en la nube podría entonces actuar como un mero intermediario en la interacción de la identidad (Nuñez & Agudo, 2014).

3.3. PRIMA: Privacy-Preserving Identity and Access Management at Internet-Scale

PRIMA, se define como un inicio de sesión federado que preserva la privacidad, es desplegable y eficiente a escala de Internet. Se basa en una versión modificada de la construcción criptográfica presentada en Oblivion, asegurando una exposición mínima de datos privados y escalabilidad.

Este sistema que propone, un proveedor de identidad que emite credenciales a los usuarios después del registro. Los usuarios almacenan localmente y usan estas credenciales cuando desean consumir servicios, sin requerir ninguna interacción con los proveedores de identidad. Este enfoque no solo garantiza la privacidad de los usuarios sino también la interacción innecesaria entre los proveedores de servicios y los proveedores de identidades. PRIMA permite a los usuarios tener control total sobre la información compartida con los proveedores de servicios permitiendo a los usuarios otorgar o revocar el acceso en cualquier momento. Aún más, para preservar la privacidad y soportar varios escenarios del mundo real, PRIMA introduce un motor de inferencia para generar pruebas específicas de los atributos. Más específicamente, el motor de inferencia genera pruebas que prueban la posesión de un atributo particular sin revelar información privada.

Además, un proveedor de servicios no puede obtener más información de la requerida o permitida por el usuario. La técnica criptográfica que se utiliza es tal que un proveedor de servicios puede verificar un número arbitrario de atributos con una simple verificación.

Evaluaciones integrales del sistema para muestran la viabilidad de este enfoque, mostró que un proveedor de identidad puede procesar de 1,426 a 3,332 solicitudes por segundo cuando el tamaño de la clave varía de 1024 a 2048 bits, respectivamente (Rizwan Asghar, Backes, & Simeonovski, 2018).

3.3. Privacy Token

Este proyecto invierte el escenario actual en el que los usuarios se ven obligados a aceptar las políticas definidas por los SP al permitirles expresar sus preferencias de privacidad. Estas preferencias están unidas a sus datos y son utilizadas por el SP para alinear sus acciones o solicitar permisos específicos.

Presenta un mecanismo para los sistemas de gestión de identidad que instruye a los usuarios sobre los posibles usos de sus datos personales por parte de los proveedores de servicios y les permite establecer sus preferencias de privacidad. Estas preferencias se convierten en una estructura estandarizada y legible por máquina, llamada token de privacidad, que luego se envía al SP junto con otros datos de autenticación. Solo se espera que los SP adapten sus sistemas de recopilación de datos para interpretar y cumplir las preferencias expresadas en el token de privacidad, que ya pueden leer y comprender una vez que tiene el mismo formato que los otros tokens utilizados por OpenId Connect.

Se espera que el modelo se implemente en sistemas IdM y se use en entornos de nube federados para permitir la privacidad del usuario permitiéndole controlar su información. Por lo tanto, también se espera que aumente su confianza en los SP de la nube y, en consecuencia, promuevan una mayor adopción del paradigma (Villarreal, Villarreal, & Westphall, 2017).

4. Plataforma escogida para la implementación en Ecuador

Se Tras la rápida revisión de las tecnologías y del enfoque que dan cada una de las soluciones anteriormente descritas se considera como la alternativa más viable y completa a BlindIdM ya que funciona plenamente como IDaaS que es el servicio que se busca posicionar y alentar en el presente estudio. La adopción del enfoque de Gestión de identidad como servicio constituye una gran oportunidad tanto para la industria de la computación en nube como para la gestión de identidades ya que ofrece la oportunidad de definir un modelo más flexible para la gestión de la identidad que aproveche todas sus características.

Se debe tener en cuenta los diversos desafíos de seguridad y privacidad que se producen en la intersección de la gestión de identidad y la computación en la nube. Una simple reubicación de un proveedor de identidad en la nube no es suficiente; la existencia de una única entidad que actúa como proveedor de identidad no aprovechará todas las posibilidades que ofrece la nube. Además, esta entidad monopoliza los roles, servicios y procesos involucrados en la gestión de la información de identidad, aumentando los riesgos asociados; la flexibilidad inherente de la nube, por lo tanto, se pierde. Es necesario descomponer estos roles, servicios y procesos y distribuirlos y organizarlos en todo el ecosistema de la nube para lograr un IDaaS apropiado.

Esto nos permite lograr una combinación a la medida y una personalización de las necesidades, los riesgos y las responsabilidades de todas las partes involucradas en las interacciones de identidad que permite:

- Optimizar el uso de los recursos, incluidos los proveedores de la nube y las infraestructuras internas.
- Mejorar la privacidad y la confidencialidad de la información de identidad y reduciendo los riesgos de seguridad.
- Incrementar el control sobre los procesos de gestión de identidad.
- Abrir nuevas oportunidades de negocios para los proveedores de la nube.

A continuación se describirá con mayor detalle su funcionamiento y las tecnologías en las cuales está basada.

El presente sistema está basado en dos tecnologías principales, SAML 2.0 como protocolo de gestión de identidad subyacente y reencriptación de proxy como protección criptográfica. Se escogió SAML por su amplia adopción, su extensibilidad y sus mecanismos arraigados para establecer confianza entre las entidades, y reencriptación de proxy porque permite garantizar la confidencialidad de los datos de los usuarios con respecto al proveedor de la nube mientras se preserva la capacidad de suministrar el servicio de identidad.

SAML 2.0 (Lenguaje de marcado de aserción de seguridad) es un marco estándar basado en XML que permite la descripción e intercambio de información de identidad entre diferentes dominios de seguridad. Con SAML, la información de identidad se expresa en forma de aserciones, que son un conjunto de afirmaciones sobre un tema; estas declaraciones cubren diferentes aspectos, tales como autenticación, autorización y atributos de identidad. El marco SAML también especifica los protocolos para emitir e intercambiar aserciones, como el protocolo de Solicitud de Autenticación, un protocolo de solicitud / respuesta que permite a las entidades solicitar una declaración de autenticación y, opcionalmente, atributos de identidad. En este protocolo, el solicitante envía un mensaje de solicitud de autenticación de SAML a un proveedor de identidad, que a su vez responde con una respuesta SAML que contiene aserciones sobre la solicitud.

En BlindIdM se usa el perfil SSO del navegador web y el enlace HTTP POST como base para las interacciones de identidad. El protocolo de Solicitud de Autenticación también permite el enfoque híbrido para la autenticación que se ha escogido, ya que esta posibilidad se considera en la especificación SAML

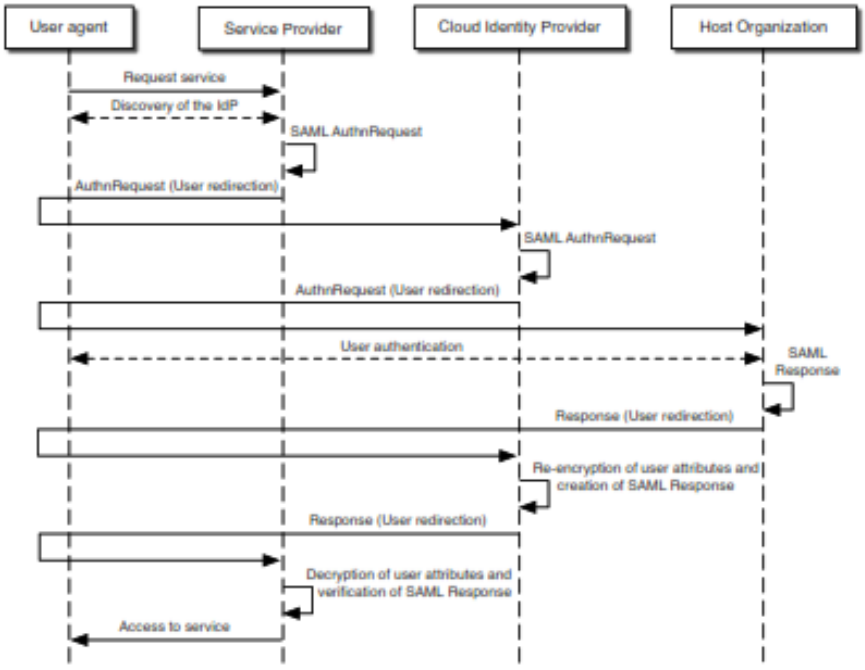


Figura 4 – Diagrama de secuencia de autenticación en el Sistema BlindIdM

En este caso, el proveedor de identidad en la nube actúa como un proveedor de identidad proxy, y la organización anfitriona es el proveedor de autenticación. Los atributos SAML se utilizan para expresar información de identidad sobre el tema de la afirmación. SAML permite la expresión de metadatos para proveedores de servicios y proveedores de identidad utilizando la especificación de metadatos SAML.

En cuanto al re-cifrado de proxy es un esquema de cifrado asimétrico que permite a un proxy transformar textos cifrados bajo la clave pública de Alicia en textos cifrados bajo la clave pública de Bob. Para ello, el proxy recibe una clave de reencriptación que hace que este proceso sea posible. La noción de reencriptación de proxy es bidireccional pero no es resistente a las colusiones. Gracias a la interacción de estas tecnologías se obtiene el siguiente diagrama de secuencia de una petición de autenticación:

A continuación se abordan las consideraciones específicas de implementación en Ecuador.

Se sugiere montar la plataforma en una entidad pública que cuente con la suficiente infraestructura y recursos para alojar una gran cantidad de usuarios. En Ecuador se ha considerado dos posibles entidades; el SRI (Servicio de Rentas Internas) y el Registro Civil ya que manejan una gran cantidad de usuarios y una constante modificación de sus atributos en sus bases de datos respectivas. Considerando que tiene una base de datos mucho más completa se elige al Registro Civil como la entidad llamada a alojar la plataforma BlindIdM. Analizando un poco más a detalle la infraestructura para brindar alta disponibilidad sugerimos levantar dos instancias; una en Quito y otra replica en Guayaquil, la una será respaldo de la otra, dado que en estas ciudades se cuenta con recursos energéticos y de red suficientes para el normal funcionamiento del IDaaS tal cual se grafica en la figura 5.

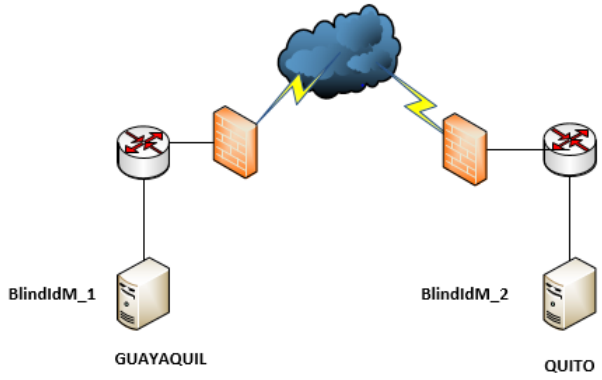


Figura 5– Modelo de Implementación del Sistema BlindIdM

La entrega del servicio consideramos que no sea directa por parte del estado sino que brinde las facilidades para que las instituciones privadas que deseen puedan conectar su nube a la plataforma y brindar el servicio al usuario final que solo necesitará manejar un usuario y un método de autenticación únicos para acceder a servicios como pago de

servicios básicos, afiliaciones a locales comerciales, pago de impuestos, matriculación vehicular, transacciones bancarias, etc.

5. Conclusiones

Este estudio muestra una descripción general de los proyectos y arquitecturas existentes, las especificaciones y herramientas consolidadas que se usan hoy en día en IdM. Se describen el estado del arte en relación con los esfuerzos realizados hacia la estandarización de los sistemas IdM.

Con base a lo investigado, se observa que hay pocas propuestas existentes y algunas son bastante incipientes y solo han quedado como teoría. Las arquitecturas existentes son bastante conceptuales, y requieren una especificación más detallada y formal sobre los conceptos y el uso de estos conceptos en la práctica. Todavía deben definir con mayor claridad cómo se implementarán realmente los componentes y módulos conceptuales y cómo será su aplicación en redes heterogéneas. Se requieren especificaciones y revisiones formales para mostrar los costos relacionados con el uso de cada una de estas arquitecturas de red. Los sistemas IdM todavía tienen desafíos hoy. Más allá de estos problemas, los sistemas IdM deben ser interoperables, ya que se espera que las tecnologías heterogéneas funcionen juntas. Como aspectos primordiales, se debe establecer la confianza entre las partes involucradas en las transacciones de identidad, evitar el abuso de identidades y hacer que las identidades estén disponibles solo para el usuario correcto en el momento y lugar correctos, todo esto de forma escalable, utilizable y de costo -manera efectiva.

Sobre las herramientas de IdM existentes, se ha considerado a BlindIdM como la más completa para la implementación de una identidad federada y administración de claves en un entorno de nube intercomunitaria.

Habiendo definido las áreas de trabajo de BlindIdM, el siguiente paso comprenderá un prototipo para fomentar las discusiones con los usuarios finales, el estado y la industria para comenzar con los esfuerzos de estandarización.

Referencias

Alkhalifah, A., & Al Amro, S. (2015). Understanding the Effect of Privacy Concerns on User Adoption of Identity Management Systems. *Journal of Computers*, 172–182.

Habiba, U., Masood, R., Awais Shibli, M., & A Niazi, M. (2014). Cloud identity management security issues & solutions: a taxonomy. *Complex Adaptive Systems Modeling a SpringerOpen Journal*, 1–37.

Isaakidis, M., Halpin, H., & Danezis, G. (2016). UnlimitID: Privacy-Preserving Federated Identity Management using Algebraic MACs. *WPES'16*, 139–142.

Jensen, J., & Jaatun, M. G. (2013). Federated Identity Management—We Built It; Why Won't They Come? *IEEE Computer Society*, 34–41.

Nuñez, D., & Agudo, I. (2014). BlindIdM: A Privacy-Preserving Approach for Identity Management as a Service. *International Journal of Information Security*, 199–215.

Rizwan Asghar, M., Backes, M., & Simeonovski, M. (2018). PRIMA: Privacy-Preserving Identity and Access Management at Internet-Scale. *2018 IEEE International Conference on Communications (ICC)*, 1–6.

Vapen, A., Carlsson, N., Mahanti, A., & Shahmehri, N. (2016). A Look at the Third-Party Identity Management Landscape. *IEEE Internet Computing*, 18–25.

Villarreal, M. E., Villarreal, S. R., & Westphall, C. M. (2017). Privacy Token: A Mechanism for User's Privacy Specification in Identity Management Systems for the Cloud. *ICN 2017 : The Sixteenth International Conference on Networks* , 53–58.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 09/10/2018
Aceitação/Acceptance: 29/12/2018

Un modelo basado en datos abiertos para la creación de un Espacio Latinoamericano y del Caribe de educación superior

Nelson Piedra[1], Lucía Puertas-Bravo[1], Santiago Mengual-Andrés[2], Roser Grau[2]

nopiedra@utpl.edu.ec, lpuertas@utpl.edu.ec, santiago.mengual@uv.es, roser.grau@uv.es

[1] Universidad Técnica Particular de Loja, Knowledge Based Systems Research Group, San Cayetano Alto, 11-01-608, Loja, Ecuador.

[2] Universidad de Valencia, Departamento de Educación Comparada e Historia de la Educación, Facultad de Filosofía y Ciencias de la Educación, Valencia, España.

Pages: 336–348

Resumen: La implementación del Espacio Europeo de Educación Superior y el proceso de Bolonia generó importantes cambios en los países europeos, entre ellos la implementación de competencias y estructura de tres ciclos, reconocimiento de títulos y garantía de la calidad de la educación superior. Por su parte en Latinoamérica y el Caribe, pese a las iniciativas de creación de espacios comunes de educación superior, no han podido concretarse. La presente investigación tiene por objetivo contribuir a la creación del espacio común en Latinoamérica y el Caribe impulsando la creación de un ecosistema de datos abiertos enlazados usando el enfoque la Web Semántica y las tecnologías de Linked Open Data. Esta propuesta tiene como finalidad de diseñar el observatorio con la intención a lo futuro de que sea puesto a disponibilidad de las organizaciones interesadas y el público en general.

Palabras-clave: big data; espacio común de educación superior; datos abiertos universitarios.

A model based on open data for the creation of a Latin American and Caribbean space for higher education

Abstract: The implementation of the European Higher Education Area and the Bologna process generated important changes in European countries, including the implementation of competencies and the three-cycle structure, recognition of qualifications and quality assurance of higher education. On the other hand, in Latin America and the Caribbean, despite the initiatives to create common spaces for higher education, they have not been able to materialize. The objective of this research is to contribute to the creation of a common space in Latin America and the Caribbean by promoting the creation of an ecosystem of open data linked using the Semantic Web approach and Linked Open Data technologies. The purpose

of this proposal is to design the observatory with the intention of making it available to interested organizations and the general public.

***Keywords**:* big data; common space of higher education; open university data.

1. Introducción

La creación de un espacio Latinoamericano y del Caribe de educación superior nace, de la experiencia iniciada por Europa en 1987. Primero con el programa Erasmus y, más tarde, en 1998 con la Declaración de Sorbona y con la Declaración de Bolonia en 1999; donde se reflexionó sobre la necesidad de la convergencia de un sistema educativo universitario, la estructuración de las titulaciones y la creación del sistema de créditos ECTS como un sistema útil para favorecer la movilidad académica por los países europeos. Sobre el Espacio Europeo de Educación Superior -EEES-, su importancia e implementación existen números estudios de los que podemos citar a Gonzalez & Liebana, 2008; Fernández, 2009; García, 2009; Mansilla, 2009, Stockwell & Bengoetxea & Tauch, 2011; Alonso-Saéz & Araita-Loroño, 2017.

Han sido varias las organizaciones que en Latinoamérica y el Caribe e Iberoamérica, han planteado propuestas para una integración regional, cada una de éstas han formulado varios acuerdos, resoluciones y decisiones sobre educación superior, ciencia y tecnología. En este mismo esfuerzo podemos encontrar iniciativas con relación a convenios para reconocimiento de títulos y creación de espacios comunes de educación superior, además de proyectos como Tuning América Latina 2004-2007. Las propuestas de espacios comunes incluyen una integración birregional considerando a los países Iberoamericanos como una forma de reducir la pobreza y combate a la exclusión social (Colomer, 2010).

Para Schwartzman (2007) la diversidad de la educación superior en Latinoamérica y el Caribe es muy grande, tenemos instituciones públicas y privadas, otras llamadas como complejas con enseñanza de grado, pos – grado e investigación científica y simples con solamente cursos de graduación. Sin embargo, nos preguntamos si, estas diferencias que marcan la realidad académica y científica en nuestros países y de nuestras instituciones, están expresadas en las leyes que regulan la educación superior o políticas de ciencia y tecnología especialmente en los últimos años.

Considerando las iniciativas propuestas por los organismos anteriormente mencionados, en los que se propone una integración de educación superior y tomando en cuenta que, actualmente no se cuenta con información sistematizada, completa y confiable del sistema académico que rige en las Universidades de Iberoamérica, consideramos necesario, entregar a las organizaciones que han mostrado iniciativas en la creación de un espacio común de educación superior, un observatorio, utilizando tecnologías de la Web Semántica (datos enlazados) en el dominio de la Legislación Universitaria: políticas de ciencia y tecnología, proyectos de investigación, publicaciones, grupos de investigación, eventos, becas, movilidad entre otros.

Creemos que ello facilitará el acceso e intercambio de información universitaria a través de una plataforma semántica de datos abiertos y en formatos legibles automáticamente por agentes máquina (“machine-readable” format). La creación del observatorio con

datos abiertos favorecerá a la movilidad de estudiantes y docentes en Iberoamérica, generación de fondos, construcción de programas académicos de doble titulación, trabajo en red, entre otros.

2. Esfuerzos de Espacios comunes de educación superior en Europa y Latinoamérica y el Caribe.

2.1. La Unión Europea

Contrariamente a lo sucede en el contexto europeo, en Latinoamérica y el Caribe no existe una organización de similares características y vinculación legal como es la Unión Europea.

La Unión Europea goza de personalidad jurídica, tiene su fundamento en valores de respeto de la dignidad humana, libertad, democracia, igualdad, estado de derecho y respeto a los derechos humanos, así como la libre circulación de personas. Se señala una sola nacionalidad para los ciudadanos pertenecientes a los Estados de la Unión, además están representados a través del Parlamento Europeo entre otras.

2.2. El Proceso de Bolonia y la enseñanza superior en Europa

Diez años tardó Europa en incorporar el Espacio Común Europeo de Educación Superior –EEES- en sus países miembros. Su origen se configuró a través de una serie de conferencias de los ministros encargados de la educación superior. Las reuniones y Declaraciones de la Sorbona en 1998 y Bolonia en 1999 fueron las más representativas que dieron inicio al objetivo de crear este espacio.

Actualmente 48 países participan en el EEES, 19 de ellos no forman parte de la Unión Europea. La Declaración de Bolonia se articuló en torno a seis acciones relativas a:

- Un sistema de grados académicos fácilmente reconocibles y comparables. Incluye la creación de un suplemento común al título superior para mejorar la transparencia;
- Un sistema basado fundamentalmente en dos ciclos: un primer ciclo orientado al mercado laboral con una duración mínima de tres años y un segundo ciclo (máster), al que se accede sólo si se completa el primer ciclo;
- Un sistema de acumulación y transferencia de créditos similar al sistema ECTS (European Credit Transfer System) utilizado para los intercambios Erasmus;
- La movilidad de los estudiantes, docentes e investigadores: la supresión de todos los obstáculos a la libertad de circulación;
- La cooperación en lo que respecta a la garantía de la calidad;
- La dimensión europea en la enseñanza superior: aumento del número de módulos, cursos y planes de estudios cuyo contenido, orientación u organización tengan una dimensión europea.

La adopción de esta declaración fue un proceso voluntario de las naciones adherentes y se fundamenta en el respeto a la diversidad de las culturas, las lenguas, los sistemas nacionales de enseñanza y la autonomía universitaria. Los Estados miembros conservan todas sus competencias sobre el contenido de la enseñanza y la organización de

sus sistemas educativos, así como su diversidad cultural y lingüística. (European Commission, 2018; Declaraciones de Bolonia, Praga, Berlín, Bergen, Londres, Lovaina, Budapest, Bucharest, Yerevan y Paris)

Un hecho notable de esta integración regional constituye sin duda que se trató de un proceso intergubernamental y voluntario de los países miembros y las instituciones de educación superior, las que fueron sumándose logrando la armonización de sus sistemas. El conseguir que todos los países participantes, crearán un único sistema, flexible, comparable y compatible de titulaciones, que facilite la movilidad docente, fue un proceso que involucró muchos años y esfuerzo para su implementación.

3. El Espacio Latinoamericano y del Caribe de educación superior

Como antecedente es menester mencionar que la historia de las universidades en Latinoamérica y el Caribe se remonta a inicios de la conquista, fue creada antes que el resto de la educación, impartiendo por muchísimo tiempo educación postsecundaria. Las primeras fueron fundadas por órdenes religiosas, con influencia del modelo de la Universidad de Salamanca.

A finales del siglo XIX se produjo una división del sistema universitario entre los conservadores criollos y el grupo laico y liberal, en donde se diferenciaban las universidades nacionales, que eran las antiguas estatales y las católicas, vinculadas a órdenes religiosas. Surgen desde esta época dos "modelos" de universidades que fueron diferenciándose a través de la historia, y que se forjaron en dos instituciones como lo son la de Salamanca y la de Alcalá de Henares, dividiendo a la educación superior en pública y privada, esta última fundamentalmente católica. (Tunnermann, 1996; Arocena, 2000)

3.1. Iniciativas de integración regional en Latinoamérica y el Caribe

Son varias las organizaciones que encontramos en la región, en orden de su creación se anotan las más influyentes:

- Sistema de Integración Centroamericana - SICA (Carta de San Salvador 1951)
- Comunidad Andina – CAN (Acuerdo de Cartagena, 1969)
- Asociación Latinoamericana de Integración – ALADI (Tratado de Montevideo, 1980)
- Mercado Común del SUR – MERCOSUR (Tratado de Asunción, 1991)
- Asociación de Estados del Caribe – AEC (Convenio de AEC, 1994)
- Unión de Naciones Suramericanas – UNASUR (Declaración del Cusco, 2004)
- Comunidad de Estados Latinoamericanos y Caribeños – CELAC (Cumbre de América Latina y el Caribe sobre integración y desarrollo, 2010)

Estas organizaciones, tienen entre sus objetivos la integración de los países que las conforman, en el ámbito económico, social, cultural, político, científico, tecnológico entre los principales. Sin embargo, a excepción de la Comunidad Andina en la que los países de Ecuador, Bolivia, Colombia, Chile y Perú en los que existe el pasaporte andino y se promueve la libre circulación de sus ciudadanos, no se ha podido eliminar las fronteras para libre circulación y tampoco se ha podido armonizar la normativa, tal como lo hizo la Unión Europea.

3.2. Iniciativas de espacios comunes en Latinoamérica y el Caribe e Iberoamérica.

En nuestra región, el proceso de integración ha involucrado la suscripción de numerosos convenios bilaterales, multilaterales y regionales sobre el reconocimiento, convalidación y homologación de estudios. Los convenios que acogen a un mayor número de países que los han suscrito se destacan:

- Acuerdo sobre títulos académicos: convenio suscrito por Bolivia, Colombia, Ecuador, Perú y Venezuela el 17 de julio de 1911.
- Convenio regional de convalidación de estudios, títulos y diplomas de educación superior en América Latina y el Caribe: suscrito el 19 de julio de 1974.
- Convenio Andrés Bello: suscrito por Bolivia, Colombia, Chile, Ecuador, España, Panamá, Perú y Venezuela

Por otro lado, podemos anotar iniciativas de espacios comunes:

- Espacio de Encuentro Latinoamericano y Caribeño de Educación Superior (ENLACES): Surge de la Conferencia Regional de Educación Superior del 2008 y está coordinado por UNESCO – IESALC y la UDUAL.
- Espacio Euro-latinoamericano para la Educación Superior Ciencia, Tecnología y Educación / ALC-UE.
- Espacio Iberoamericano de conocimiento.

4. Propuesta: Construir un ecosistema universitario de datos abiertos enlazados

4.1. La Web: un espacio para colaborar a gran escala

Internet y la Web han hecho posible que dispongamos de una infraestructura de comunicación que permite que nos relacionemos, compartamos y colaboremos, en los más diversos campos de la actividad humana: educación, negocios, ciencia, relaciones sociales, diversión, y todo lo que uno pueda imaginar. Tenemos acceso a millones de recursos, sin tener que preocuparnos por restricciones de hardware, software, infraestructura de red, idioma, cultura, localización geográfica, o habilidades de los usuarios. Este entorno es el resultado del esfuerzo continuo de la comunidad internacional comprometida con el desarrollo de estándares abiertos que aseguran el crecimiento de la Web a largo plazo. Este crecimiento no ha alterado el modelo descentralizado, heterogéneo, sencillo, diverso, compatible, y abierto al crecimiento, la contribución, y los usos no previstos. En este empeño ha sido importante los esfuerzos que se canalizan a través del consorcio W3C (Ward Wide Web Consortium).

El concepto de “Datos Abiertos” facilita el acceso a datos públicos, y tiene el potencial de garantizar una gestión más eficiente. Además, los datos abiertos proporcionan una interfaz para la colaboración e integración de sistemas distribuidos y heterogéneos, entre diferentes actores: ciudadanos, gobiernos, empresa privada, academia, entre otros. En este sentido, el uso de datos abiertos tiene como objetivo avanzar en el conocimiento y la implementación de diferentes servicios. En orden a alcanzar este objetivo, las plataformas tecnológicas y los repositorios deben disponer de un conjunto de características que

permitan que la integración y reuso de información pública sea factible y económico. Cuando los datos son abiertos significa que se han reducido las barreras de acceso a datos con el fin de garantizar a cualquiera la reutilización más amplia posible.

4.2. Presencia de silos de información en los datos universitarios

La información es clave para crear una cultura de transparencia, colaboración y participación. En este contexto, la Web puede ser imaginada como una base de datos global, gigante, distribuida y amplia, que contiene información de todos los campos de la actividad humana. Sin embargo, si se desea construir una aplicación basada en la integración de información procedente desde diferentes fuentes de información, con frecuencia se enfrentará a actividades que son muy costosas y que demandan demasiado tiempo de desarrollo. Identificamos tres causas de esto:

SILOS: Las fuentes de información son diseñadas como "silos" que operan en forma asilada al resto de bases de datos. Sus responsables no desean que otros toquen la base de datos, y asumen que ellos son los únicos con capacidad de entender el esquema y el significado de los datos. Esto es obsoleto.

HETEROGENEIDAD: Diversidad de formatos de representación, almacenamiento e intercambios de información. Cuando se desea reducir costos asociados a la transferencia de información entre sistemas, los estándares desempeñan un rol crucial. Ya sea el transmisor o el receptor, tienen que convertir o mapear sus datos a una estructura y formato que pueda ser entendida por el sistema que recibirá el flujo de información. Este mapeo o conversión para entregar información significativa se debe hacer en por los menos tres niveles diferentes: sintaxis usada para describir los datos, esquemas/ estructuras de datos y vocabularios utilizados. Un escenario ideal sería un entorno totalmente armonizados, en el que las capas están basadas en un estándar único, pero lo cierto es que en la actualidad nos enfrentamos a muchos estándares de facto.

LIMITACIONES AL REUSO: Muchos datos universitarios aún están encerrados en aplicaciones. El problema técnico más común de hoy en día en las arquitecturas de información es que los metadatos y los esquemas de información no están bien separados de la lógica de la aplicación. Esto provoca que los datos no se puedan reusar tan fácilmente como se debería. Se debe dejar de diseñar bases de datos que se enfocan en aplicaciones que usan los datos y más bien enfocarse en la descripción significativa de los datos en sí.

4.3. Integrar datos universitarios usando Linked data

La Web desde su origen fue pensada como una herramienta para compartir y acceder a información globalmente (Abrams, 1998; Connolly, 2000). Su creador, Tim Berners-Lee planteó un reto social: construir confianza a gran escala (Berners-Lee, 1989). El W3C reconoce que la confianza es un fenómeno social, que puede ser apoyado con el diseño de tecnologías que fomenten la confianza y la responsabilidad de forma abierta. Desde la visión de esta propuesta, la Web Semántica (Berners-Lee., et al., 2001) tiene el potenciar de resolver dos importantes problemas presentes en la información universitaria que se comparte actualmente a través de la Web: falta de estructuración y significado en el contenido Web y heterogeneidad de fuentes de información (Piedra, 2016). Mediante las tecnologías semánticas se podría mejorar la capacidad de procesar contenido

heterogéneo y distribuido, razonar con este, combinarlo y realizar deducciones lógicas para resolver problemas cotidianos automáticamente (Bizer, et. al., 2009; Gómez-Pérez A., et al., 2003)

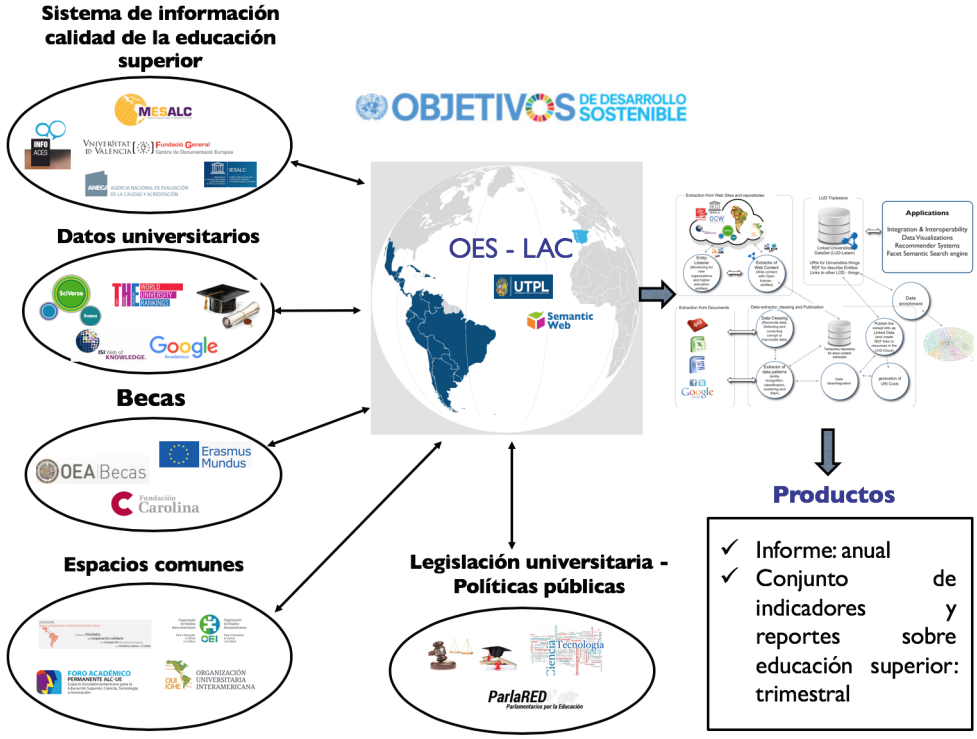


Figura 1 – Visión general sobre integración de datos en el ecosistema de educación superior de América Latina y el Caribe

La implementación de la propuesta de este trabajo se basa en los 4 principios de diseño planteados por Tim Berners-Lee para la publicación de datos en la Web (Heath, et al.,2001). El primer principio hace referencia al uso de un estándar para nombrar los recursos. Este puede estar expresado por una dirección URL, por un nombre uniforme o por ambos. Los recursos pueden ser documentos digitales, archivos, imágenes, objetos del mundo real como personas, entidades, organizaciones, o conceptos abstractos como cargos, conocimientos, etc. El segundo enunciado, manifiesta que Linked Data usa el protocolo HTTP como un mecanismo de acceso universal a las URI's. El tercer enunciado se refiere al uso de RDF para ofrecer información de los recursos, es un modelo común que permite hacer afirmaciones sobre los recursos y hace posible que estos recursos pueden ser nombrados por las URI's antes mencionadas. Finalmente, el último principio de diseño de Linked Data permite conectar datos relacionados y evita la existencia de datos aislados. Esencialmente, abrir datos significa que los datos están abiertos a cualquier tipo de aplicación y uso, y esto se puede lograr si se utilizan un enfoque de arquitectura de datos abiertos, además de estándares abiertos para describir metadatos como RDF.

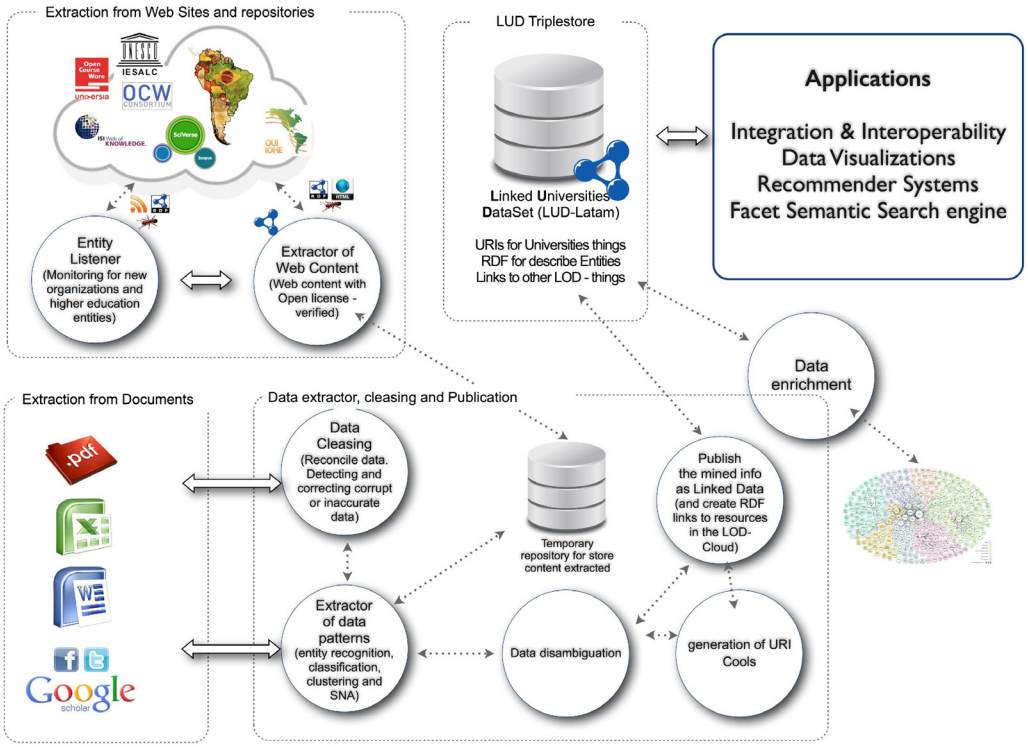


Figura 2 – Plataforma semántica y colaborativa para acceso en línea a datos y documentos universitarios abiertos y enlazados. Alcance: instituciones que formen parte del Espacio de Educación Superior de América Latina y el Caribe. El uso de Linked Data agrega una capa semántica a los silos de información (Piedra N., 2016).

4.4.Ciclo de vida propuesto para la publicación de datos enlazados

Siguiendo la visión de la Web Semántica y las tecnologías de Linked Data (Berners-Lee, et al., 2001), los datos deben estar publicados en un espacio global que permita que estos sean utilizados por varios usuarios y aplicaciones (Heath, et al., 2011). En el contexto de este trabajo, la publicación de datos enlazados sigue un ciclo de vida presentado en (Piedra, et al., 2014) y descrito con mayor detalle en (Piedra, 2016). El propósito es guiar a los interesados en los pasos que deben seguir para obtener un proyecto de publicación de datos enlazados que reúna las características técnicas de esta propuesta y satisfaga las expectativas de los usuarios (ver Figura 2). El proceso es como sigue:

1. Planificación del proyecto. El proceso arranca con el visionamiento general del proyecto haciendo énfasis en la identificación de fuentes de datos que contienen la información a publicar, la normativa legal, negociando recursos para el proyecto, y estableciendo el alcance del proyecto. En esta fase se trata de adquirir conocimiento sobre el entorno de trabajo, una vía es responder preguntas como: ¿dónde están los datos?, ¿en qué formato?, ¿qué tipo de repositorio? ¿cuál es el formato en el que se entregan los datos?

Es fundamental definir el alcance del proyecto de publicación de datos enlazados, ya que proporciona un entendimiento común entre los interesados del proyecto. Se debe elaborar a partir de los entregables principales, recursos disponibles y posibles restricciones al proyecto que se han documentado en la fase de identificación de fuentes de datos. Durante esta fase, se analizan los riesgos, los supuestos y las restricciones existentes, actualizando esta información según sea necesario. Esto ayuda a gestionar las expectativas de los interesados, y proporciona la línea base del alcance, definiendo: qué datos se publicarán como Linked data y qué es lo que está excluido del proyecto, qué está fuera del alcance. Además, se recomienda describir las restricciones (legales, técnicas, funcionales, económicas) específicas asociadas al alcance, como pueden ser por ejemplo el presupuesto, fechas tope o hitos, así como los supuestos asociados con el alcance y el impacto potencial de tales supuestos en el caso de que no se cumplan. Tareas: visionamiento del proyecto; determinación de alcance; identificación, análisis y selección de fuentes de datos

2. Selección de fuentes de datos, extracción y reducción de ambigüedades.
   Una vez que se ha planeado y extraído la información que forma parte del alcance del proyecto, es importante establecer una estrategia de análisis y limpieza de datos. El propósito es identificar ambigüedades, errores e inconsistencias en la información a convertir. Entre los errores comunes están: errores de accesibilidad, ambigüedad de términos, tipos de datos malformados o incompatibles, o términos de vocabularios mal definidos. Por ejemplo, fechas en formatos incompatibles, datos de edad fuera de rango, datos vacíos, o datos con semántica diferente.
3. Modelamiento de ontología. En la fase de creación de ontologías y vocabularios semánticos controlados, se deben mapear los conceptos según el ámbito y alcance definido. Aunque estos sean reutilizados, existirán aquellas definiciones distintas que describan un mismo término, a lo cual se le puede sacar ventaja siempre que sea utilizado correctamente (Villazón, et al., 2011). Es clave recomendar la reutilización de vocabularios existente por razones de relación coste/beneficios; ejemplos de vocabularios son FOAF para describir personas, SIOC para describir semánticamente comunidades on-line, DublinCore para describir recursos de información, DCAT para describir catálogos de datos, entre otros. El re-uso permitirá que los datos a publicar se puedan conectar con otros conjuntos de datos existentes. El re-uso de vocabularios aporta facilidades para la búsqueda, recuperación, integración, visualización, análisis y representación comparada en cada campo y dominio específico. El re-uso contribuirá a crear un escenario global de datos enlazados. Tareas: mapeo de conceptos; búsqueda de recursos ontológicos y no ontológicos a reusar; diseño de estrategia para identificar a los recursos a través de URIs persistentes. En Linked Data es muy importante, mantener las URIs de manera que sigan existiendo y sirviendo a lo largo del tiempo. Se recomienda pensar en el espacio de identificación de recursos descritos en Linked Data como un espacio de URIs, abstracto y perfectamente organizado. Es fundamental diseñar una estrategia de creación de URIs y un plan que asegure la persistencia de los enlaces cuando las cosas cambien.
4. Generación y publicación de datos en RDF. La siguiente fase consiste en generar datos enlazados usando RDF. El trabajo se centra en convertir las fuentes de

datos identificadas previamente a tripletas RDF utilizando los vocabularios establecidos, de esta forma se cumple con el tercer principio de Linked Data. Publicar los datos implica seguir un proceso que permita de una forma organizada aportar con información valiosa y estructurada a la Web Semántica, para tomar en consideración la publicación de datos es necesario plantearse preguntas que permitan visualizar el alcance del proyecto: ¿Con qué cantidad de datos se cuenta?, ¿Cuál es el estado de los datos a vincular? ¿Cómo se almacenarán los Datos?, ¿Quiénes son los potenciales usuarios de esos datos? ¿Cuáles pueden ser los usos que den a esa información?. Tareas: extracción de información; ejecución de tareas de limpieza de datos que ayuden a detectar posibles errores en la información; generación de datos en RDF; publicación de datos enlazados en un RDF triplestore.

5. Enlazado de datos a fuentes de datos RDF externas. Para que los datos abiertos permitan alcanzar altos beneficios, es crucial poner los datos y la información en un contexto que crea nuevo conocimiento y que permite crear o ampliar el alcance de los servicios y las aplicaciones. Una vez que los datos han sido generados, el proceso continúa con la publicación de datos y el almacenamiento en una base de datos RDF (triplestores), además se ejecutan tareas que permitan aplicar el cuarto principio de Linked Data "incluir enlaces con otras URIS" para lograr interconectar los datos que son parte del alcance del proyecto con conjunto de datos externos.

La aplicación de este proceso promueve la estructuración de la información almacenada en los sitios web y la codificación explícita del significado de los datos publicados. Además, las computadoras podrían operar automáticamente sobre la información publicada, por ejemplo, agentes software podrían hacer inferencias sobre estos datos y apoyar a los humanos en tareas de generación de conocimiento.

5. Conclusiones

América Latina y el Caribe requieren, un espacio común de educación superior que favorezca la colaboración y la cooperación internacional para responder a las circunstancias sociales locales y regionales. Un espacio innovador y flexible en los tipos de ofertas educativas, que favorezca el crecimiento de la calidad de las instituciones y que facilite la movilidad regional con base al reconocimiento de créditos académicos obtenidos en diferentes contextos. Para lograrlo se requiere avanzar de manera decidida en el diseño de estrategias de internacionalización con mecanismos de movilidad académica, creación y operación de redes, colaboración internacional para la oferta de programas educativos, en sus modalidades presencial y virtual, así como un ambicioso programa para la promoción de la internacionalización. Todo ello mediante la movilización de profesores e investigadores preocupados por vincular investigación, docencia e innovación.

Pese a que existen en nuestra región convenios que establecen el compromiso y la voluntad de los Estados de reconocer los estudios y títulos a fin de continuación de ellos o incluso el ejercicio profesional, en la práctica no hay un reconocimiento automático, ya que no se cuenta con normativa específica y podría no haber reciprocidad. Los titulares de estudios y títulos deben realizar trámites especiales a fin de que se les reconozca sus estudios. La calidad y la diversidad de planes de estudio existentes en las universidades,

es un factor que podría incidir en esta problemática, por tanto, el proceso de integración requiere además analizar las diferentes universidades existentes en Latinoamérica y el Caribe o por lo menos su clasificación, en grandes y pequeñas, privadas y públicas, de mayor o menor calidad, que como bien sabemos difícilmente podríamos unificar un solo criterio de igualdad entre ellas.

Europa consideró a la integración de la educación superior como relevante para construir una Europa más completa e influyente y la consolidación de sociedades estables; además de mejorar la calidad y atractivo entre las instituciones de educación superior. El EEES se consolida sobre un rico y variado patrimonio cultural, autonomía universitaria y libertad académica de oportunidades y principios democráticos. Para lograrlo, la recolección de datos e indicadores fue considerada como prioritaria para medir los avances que se realizaron en su proceso de implementación. El reto que tenemos los países Latinoamericanos y del Caribe, es fortalecer las instituciones de educación superior y movilidad. El contar con un observatorio de educación superior con datos abiertos podría aportar para que los actores políticos, de gobierno, así como gestores de las instituciones u organismos interesados puedan tomar decisiones en base a datos e indicadores. Con ello podríamos adelantar un proceso que Europa lo hizo previo a la implementación del EEES.

La idea básica que fundamenta la presente propuesta implica establecer vías de colaboración entre políticos, administración pública, industria, universidades y ciudadanos, con el propósito de fortalecer la integración, la participación, la colaboración, y, una mayor transparencia. En esta vía, la gestión eficiente de datos e información son aspectos críticos en el entorno de la economía global. En particular, en sectores cuyas actividades son sensibles a compartir datos y conocimiento, ya que estas posibilidades pueden asegurar una mejor toma de decisión, asegurar desarrollo de proyectos potentes y promover mecanismos efectivos de colaboración en las funciones universitarias de investigación, docencia, vinculación y transferencia de conocimiento.

La interoperabilidad y los estándares abiertos son claves cuando se piensa cómo obtener beneficios plenamente de la integración de información compartida en un ecosistema distribuido y heterogéneo. Es en este contexto de interoperabilidad en donde la visión de la Web Semántica y los principios de Linked Data cumplen un papel fundamental. Este enfoque tiene por objeto ayudar a las organizaciones en la adopción de las medidas necesarias para que abran sus conjuntos de datos y maximizar la interconexión entre intermediarios y proveedores de conocimiento. Los datos abiertos y enlazados plantean un mecanismo importante para la gestión, la interoperabilidad y la integración de información que facilita la innovación y la creación de conocimiento a partir de datos entrelazados. Al dotar de estructura y significado a la información universitaria con semántica explícita, procesable por máquinas, es posible automatizar tareas de integración de datos y de filtrado automático preciso de la información. Para ello, es necesario hacer que la información universitaria que reside en la Web use los estándares de la Web Semántica. Anotar semánticamente un recurso digital implica incluir descripciones explícitas del significado de los datos que están contenidos en las páginas Web. Con esto, se permite que las propias máquinas tengan un nivel de comprensión de la Web suficiente como para hacerse cargo del trabajo que actualmente realizan los usuarios que navegan en la Web, que son rutinarios, demandantes de tiempo o incluso físicamente imposibles para los humanos.

Referencias

Abrams M. (1998). *World Wide Web - Beyond the Basics*, Prentice Hall. Editorial: Pais

Alonso-Sáez, I. & Arandia-Loroño, M., (2017). 15 años desde la Declaración de Bolonia. Desarrollo, situación actual y retos del Espacio Europeo de Educación Superior. *Revista Iberoamericana de Educación Superior*, *8* (23), 199–213.

Arocena, R., & Sutz, J. (2000). La universidad Latinoamericana del futuro. México Unión de Universidades de América Latina.

Berners-Lee, T. (1989) Information Management: A Proposal. Internal Project Proposal, CERN, 1989. Disponible en https://www.w3.org/History/1989/proposal.html

Berners-Lee, T., Hendler, J., & Lassila, O. (2001) The Semantic Web. Scientific American, May 2001. Doi: https://doi.org/10.1038scientificamerican0501-34

Bizer C., Heath T., Berners-Lee T. (2009) Linked Data - The Story So Far *International Journal on Semantic Web and Information Systems*. Vol, pp doi: https://doi.org/10.4018/jswis.2009081901

Colomer, A. (2010) Algunas reflexiones sobre la integración latinoamericana, a la luz de la integración europea. *Revista Aportes para la Integración Latinoamericana*, *23*, 80–89.

Connolly, D., (2000) A Little History of the World Wide Web. W3C Consortium, 2000. Disponible en http://www.w3.org/History.html.

European Commission/EACEA/Eurydice. (2018). The European Higher Education Bologna Process Implementation Report. Luxembourg: Publications Office of the European Union.

Fernández, C., & Serrano, C. (2009). El plan Bolonia. Madrid: Libros de la Catarata.

García, M. (2009) *Hacia el Espacio Europeo de Educación Superior*, netbiblo.

Gómez-Pérez A., Fernández-López M., & Corcho, O. (2003). *Ontological Engineering with examples from the areas of Knowledge*. Madrid: Springer.

González, A. & Liébana J. (2008) *Posibilidades, experiencias y retos en el Espacio Europeo de Educación Superior*. Granada: Universidad de Granada.

Gruber, T., (2007). Collective Knowledge Systems: Where the Social Web meets the Semantic Web. *Journal of Web Semantics*.

Heath, T., Bizer, C. (2011). *Linked Data: Evolving the Web into a Global Data Space*. *Synthesis Lectures on the Semantic Web: Theory and Technology*. Morgan & Claypool.

Mansilla, M. I., (2009) El European Credit Transfer System (ECTS) como medida del haber académico" (64-85), en García Manjón, J. (Coord.), Hacia el Espacio Europeo de Educación Superior, netbiblo.

Piedra, N., Tovar, E., Colomo-Palacios, R., Lopez-Vargas, J., Alexandra Chicaiza, J., (2014) "Consuming and producing linked open data: the case of OpenCourseWare", Program, Vol. 48 Issue: 1, pp.16-40, https://doi.org/10.1108/PROG-07-2012-0045

Piedra Pullaguari, N. O. (2016). Marco de Trabajo de Interoperabilidad Semántica para la Educación Abierta: un enfoque para la Integración de Sistemas de Información de Recursos Educativos Abiertos autónomos, distribuidos y heterogéneos (Doctoral dissertation, ETSI_Informatica).

Schwrtzman, S. (2007). La Universidad latinoamericana entre el pasado y el futuro. In Pantoja Morán D. (Comp.), Antología del pensamiento Latinoamericano sobre la educación, cultura y las universidades. Ciudad de México.

Stockwell, N., Bengoetxea, E., & Tauch, C., (2011). El Espacio Europeo de Educación Superior y la promoción de la cooperación académica y de la movilidad con México. *Perfiles Educativos*, *133*, 198–205.

Tunnermann, C. (1996). Breve historia del desarrollo de la universidad en América Latina, La educación superior en el umbral del siglo XXI. Caracas: Ed. CRESALC.

Villazón-Terrazas B., Vilches-Blázquez L.M., Corcho O., Gómez-Pérez A. (2011) Methodological Guidelines for Publishing Government Linked Data. In: Wood D. (eds) Linking Government Data. Springer, New York, NY

Este trabajo es parte del proyecto financiado por la Fundación Carolina, en el programa de becas de estancias cortas de investigación postdoctoral otorgado a la segunda autora.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 12/09/2018
Aceitação/Acceptance: 13/11/2018

A tecnologia ao serviço da música – Um estudo no sentido de se evitar lesões e maximizando-se a *performance*

Eliseu Silva[1], Manuel Au-Yong-Oliveira[2], Pedro Fonseca[3], Rui Garganta[4], Christopher Bochmann[1]

eliseu7silv@gmail.com, mao@ua.pt, pedro.labiomep@fade.up.pt, ruigarg@fade.up.pt, bochmann@uevora.pt

[1] Music Department, University of Évora, Largo dos Colegiais 2, 7000 Évora, Portugal.

[2] GOVCOPP, Department of Economics, Management, Industrial Engineering and Tourism, University of Aveiro, Campus Universitário de Santiago, 3810-193 Aveiro, Portugal.

[3] LABIOMEP - Porto Biomechanics Laboratory, Rua Dr. Plácido Costa, 91, 4200-450 Porto, Portugal.

[4] Faculty of Sports of the University of Porto, Rua Dr. Plácido Costa, 91, 4200-450 Porto, Portugal.

Pages: 349–362

Resumo: Todos os violinistas e violetistas enfrentam as dificuldades de tocar um instrumento não simétrico e pouco ergonómico, o que torna bastante fácil criar técnicas erradas. Foi elaborado uma metodologia científica, composta por alguns padrões de notas simples e repetíveis, e analisada a *performance* em 3D, por um sistema MoCap (Qualisys AB, Suécia) que opera a uma frequência de amostragem de 200 Hz, captando o posicionamento de marcadores retrorrefletores nas articulações da mão, punho, antebraço esquerdo, violino e arco. Os dados resultantes foram analisados no software Visual 3D (C-Motion, EUA) para extrair parâmetros e métricas relevantes. Com base neste estudo foi elaborado um modelo anatómico-funcional do posicionamento da mão esquerda, a divulgar brevemente. Com base na tecnologia, aqui discutida, apurou-se que a técnica certa leva a menos tensão e lesões, a uma melhor evolução, levando também a uma evolução mais rápida, e com menos esforço, sendo por isso mais produtivo.

Palavras-chave: Biomecânica do membro esquerdo; análise movimento 3D; MoCap Qualisys AB; angulação e nível da articulação metacarpicofalângica; programa virtuoso.

Technology at the service of music: A study on how to avoid injury while maximizing performance.

Abstract: All violinists and violists face the difficulties of playing on a non-symmetrical instrument, which makes it quite easy to create wrong techniques. A scientific methodology was developed, composed of some simple and repeatable note patterns, and this performance was analyzed in 3D by a MoCap system (Qualisys AB,

Sweden) that operates at a sampling frequency of 200 Hz, capturing the positioning of retroreflective markers on the joints of the hand, wrist, left forearm, violin and bow. The resulting data were analysed in Visual 3D software (C-Motion, USA) to extract relevant parameters and metrics. Based on this study, supported by technology, and on a soon-to-be divulged anatomical-functional model of the positioning of the left hand, we are able to state that the right technique leads to less tension and injury, to a better and faster evolution, with less effort, thus being more productive.

Keywords: Biomechanics of the left limb; 3D motion analysis; MoCap Qualisys AB; angulation and level of the metacarpophalangeal joint; virtuoso program.

1. Introdução e enquadramento

Não é comum em Portugal profissionais de música clássica procurarem na tecnologia um meio para prevenir lesões (Silva et al., 2018). As lesões são um problema real e tendem a aparecer em um, ou até em mais que um, momento da carreira de músicos profissionais. Cada vez mais se começa a tocar mais jovem, e dada a extrema competitividade, cada vez mais se treina durante mais horas. Ora, isto faz com que as lesões apareçam, mais cedo ou mais tarde. Este estudo é sobre como se poderão evitar essas mesmas lesões, no caso do violino, e baseia-se no uso e apoio em tecnologia avançada para se determinar incorreções de postura (e de colocação da mão esquerda, em particular) a corrigir o quanto antes pelo músico.

Este estudo é assim inédito pois envolve uma equipa multidisciplinar, em Portugal, na resolução de um problema recorrente na música – o das lesões músculo-esqueléticas. Este tipo de lesões, quando mais graves, poderão significar a interrupção involuntária de uma carreira de topo, obrigando assim um músico a ter que procurar outra via de subsistência que não a da *performance* em espetáculos ou para gravação. Fizeram parte da equipa especialistas em anatomia e em fisiologia, assim como um músico de topo (o primeiro autor) a par das atuais exigências de profissionais de orquestra assim como os que atuam mais a solo e/ou em grupos mais pequenos (por exemplo, em duetos ou em quartetos). Também foi necessário ter um especialista em tecnologia na equipa, para as medições e a recolha de informação.

No caso do violino, uma parte do corpo que tende a ser mais afetada é a mão esquerda, dado o ângulo acentuado e movimentos antinaturais que se têm que fazer, e durante um tempo prolongado – tanto no treino, como na *performance* em si.

A investigação sobre a mão esquerda de um violinista/violetista é também muito parca e muitos autores/violinistas nos seus métodos discutem diferentes maneiras de trabalhar a afinação ou oferecem opiniões pessoais sobre possíveis resoluções. No entanto, essas abordagens convencionais não são cientificamente e anatomicamente baseadas e não sugerem uma compreensão ergonômica e biomecânica do funcionamento da mão esquerda nos seus vários desafios. Este estudo pretende assim preencher uma lacuna que existe atualmente na literatura científica e ligado ao instrumento violino em particular.

Para Frank e Mühlen (2007) no seu trabalho de investigação intitulado *Queixas Musculoesqueléticas em Músicos: Prevalência e Fatores de Risco*, defendem que, independentemente dos instrumentos praticados, a zona corporal que mais era afetada por problemas músculo-esqueléticos, é o pulso esquerdo.

Estes resultados talvez se relacionem com o facto de que, dos vários grupos de músicos estudados, a maior parte dos estudos apontam os instrumentistas de cordas como aqueles que apresentam uma maior prevalência de problemas músculo-esqueléticos.

Demonstram ainda que comparando os dois braços de violinistas e violetistas, estes apresentam o dobro da predominância de desordens musculares no membro superior esquerdo em comparação com o membro superior direito. Isto porque os flexores da mão direita não têm grande mobilidade, apenas a tarefa de segurar e controlar os movimentos do arco.

Outros autores, Moraes e Papini (2012), defendem que a grande maioria dos problemas músculo-esqueléticos em violinistas e violetistas advêm de incorreções técnicas, ou de tensões musculares que os violinistas desenvolvem na *performance* do seu instrumento numa passagem em específico ou numa abordagem mais generalizada.

Ambos advogam que a quantidade de força, coordenação, velocidade e motricidade fina exigida na *performance* diária do instrumento feita durante horas a fio, adicionada a uma técnica pouco consciente ou até mesmo incorreta, pode efetivamente originar problemas músculo-esqueléticos bastante graves que podem resultar, em última instância, na cessação da atividade profissional, conforme já foi referido.

Este estudo pretende mostrar a relevância da tecnologia para os problemas expostos em cima, e o quão importante esta relação com a tecnologia pode ser para o bem-estar das pessoas e dos músicos em particular.

2. Revisão da literatura

Este capítulo tem como objetivo central fazer uma revisão bibliográfica sobre as várias abordagens tecnológicas que se fizeram à performance violinística conjugando-a com a tecnologia científica. Isto porque este estudo pretende efetivamente analisar a biomecânica do braço esquerdo do violinista. Desta forma foi nossa intenção perceber o que se tinha feito até então.

Efetivamente algum trabalho científico já foi realizado ao nível da performance violinística, recorrendo a diversos dispositivos e tecnologias científicas.

A pesquisa seguinte foi realizada relativamente há pouco tempo, na Austrália, em Sidney. Foi utilizada essencialmente a eletromiografia na análise de violinistas profissionais, de forma a perceberem as vantagens do aquecimento na performance. Este trabalho, desenvolvido por McCrary, Halaki e Sorkin (2016), que tem como título *Acute Warm-up Effects in Submaximal Athletes: An EMG Study of Skilled Violinists* analisou através da EMG[1] o trabalho muscular, o trabalho cardiovascular e o aquecimento, medindo os seus resultados na performance violinística e na opinião de violinistas dotados. Desta forma foi pedido a 55 violinistas desde licenciados a profissionais, que tocassem 45 segundos de excertos musicais antes e depois de 15 minutos de aquecimentos gerais, muscular, cardiovascular e técnico. Os elétrodos para medição eletromiográfica foram colocados em 16 músculos diferentes. Os resultados sonoros também foram obtidos e gravados. Os

[1] EMG – equivalente a Eletromiografia.

resultados mostraram que nenhum dos três protocolos de aquecimento tiveram resultados significativos na atividade muscular e na qualidade da performance. Os resultados demonstraram que o aquecimento traz benefícios somente no limiar da força máxima.

Um outro autor relevante no estudo biomecânico na performance violinística é Young (2003).

O seu trabalho apresentado em Young (2002a, 2002b) baseou-se essencialmente num interface denominado de *Hyperbow*. O objetivo era criar um arco de violino capaz de medir os aspetos mais intrínsecos e minuciosos dos gestos musicais que têm impacto direto no som do instrumento. De forma a poder-se observar e analisar as nuances mais ínfimas da articulação do arco, foi desenvolvido um sistema de sensores, que usam o campo eletromagnético e acelerómetros comerciais MEMS e medidores de tensão.

Já o trabalho realizado em 2007 com o título de *A Methodology for Investigation of Bowed String Performance Through Measurement of Violin Bowing Technique* (Young, 2007) é um desenvolvimento do anterior. Este estudo foca-se essencialmente sobre o braço direito do violinista, tentando descodificar a complexidade dos gestos que produzem os resultados sonoros e as diferentes nuances e cores no violino. Assim através de tecnologia avançada, são colocados um sistema de sensores num arco de fibra de carbono, de forma a serem pequenos, leves e portáteis. Estes sensores são essencialmente inerciais e de posição de forma a poder medir o movimento e aceleração. Todos os sistemas foram uniformizados em SI[2]. Este sistema é denominado de *Hyperstring*, onde para além de recolherem os vários resultados inerciais destes sensores aplicados ao instrumento e arco, também é gravado o resultado sonoro em áudio, e em cenários realistas de performance. Este sistema de medida e análise de movimento foi usado para o estudo de técnicas básicas de arco como a realização de *detaché,* de *martelé* e até *pizziccato*. Este trabalho de recolha de dados foi sistematizado por forma a criar uma abordagem mais científica no ensino e performance de alguns golpes de arco.

Outro estudo bastante semelhante é o de Großhauser, Großekathofer, e Hermann (2010) com a sua pesquisa científica sobre a mão direita e o arco do violinista: *New Sensors and Pattern Recognition Techniques for String Instruments*. Este estudo pesquisou forças de pressão na *performance* do violinista. Para isto, usou um complexo sistema de sensores acoplados ao arco e ao violino, juntamente com giroscópios acelerómetros, e outros sensores transmitidos por rádio frequência e integrados numa pequena caixa colocada no talão do arco do violino. Foi usado também um sistema de captação e análise do movimento em 3D e gravado o resultado musical. Desta forma conseguiram detetar e interpretar as diferentes forças, muitas das vezes impercetíveis e bastante complexas, que atuam durante a performance instrumental.

Desta forma analisaram o trabalho da mão direita em consonância com a mão esquerda e a coordenação de ambas.

Conseguiram analisar, por meio de sensores de pressão, a relação da força que no queixo faz durante as mudanças de posição de forma a suportar o instrumento, quando a mão esquerda exerce menos pressão para se movimentar na escala do violino.

[2] SI – Sistema Internacional.

Este estudo foi uma combinação de informações por vídeo, áudio e software de recolha de informação de pressão e movimento que poderá ser futuramente usado para o ensino e para a melhoria do estudo diário (figura 1).

Figura 1 – Imagem do arco montado com sensores de pressão

O seguinte estudo tem muitas similaridades com o anterior e é uma pesquisa científica sobre o braço direito de instrumentistas de cordas.

Foi desenvolvido no Reino Unido, pela unidade de reabilitação de Northwick Park e St Mark's Hospital Trust e dirigido pelos inverstigadores Turner-Stokes e Reid (1999) com o título de *Three-dimensional motion analysis of upper limb movement in the bowing arm of string-playing musicians*. Este estudo teve o objetivo de pesquisar o papel de cada articulação do braço direito a nível tridimensional, na performance de um instrumentistas de cordas, de forma a poder ser usado este conhecimento clinicamente, ao nível da reabilitação de problemas músculo-esqueléticos. Desta forma foi usado um dispositivo de análise biomecânica tridimensional MacReflex 3-D analysis system, onde incluía a colocação de marcadores calibrados no espaço, ao longo do braço direito, que analisavam sequencias de arcadas pré-estabelecidas. Este estudo abrangeu 39 instrumentistas de violino, viola e violoncelo. Os resultados demonstraram as diferenças claras entre os vários instrumentos e os vários instrumentistas.

A amplitude do movimento do ombro diminuía desde o violoncelo ao violino, e a máxima elevação do ombro tinha registos máximos no violoncelo ($P < 0.001$) enquanto os maiores registos de flexão no cotovelo eram obtidos nos violinistas. Os resultados demonstraram que existiam diferenças consideráveis mesmo entre instrumentistas do mesmo instrumento. Os violoncelistas, pela predisposição técnica própria do instrumento, eram mais predispostos a sintomas de problemas músculo-esqueléticos.

Um outro autor relevante no estudo da performance do violino aliado às técnicas científicas e laboratoriais é Peter Visentin (Visentin, 2010; Visentin e Shan, 2011).

O projeto desenvolvido por este autor em 2008 (Visentin, 2008) em colaboração com Shan Gongbing e Edwin B. Wasiak, tem como título original *Informing music teaching and learning using movement analysis technology*, e foi desenvolvido pela Universidade de Lethbridge, no Canadá. Este estudo explora a utilidade da tecnologia na análise do movimento como contribuição para a pedagogia do instrumento. Duas perguntas cimeiras foram investigadas:

1. Poderão as capacidades biomecânicas na performance do violino, ser caracterizadas de uma forma objetiva e precisa de forma a serem generalizadas?

2. Poderão estes dados obtidos serem úteis à pedagogia de forma a maximizar a eficiência performativa e minimizar as lesões?

Foi neste estudo usada tecnologia capaz de estudar e analisar os fenómenos biomecânicos relacionados com o arco, mais especificamente, o legato com EMG e tecnologia de análise do movimento em 3D.

Marcadores foram colocados no ombro e no cotovelo direito. Os resultados obtidos em violinistas profissionais foram então adaptados com objetivos pedagógicos. Os resultados salientam também a importância de uma colaboração mais sistematizada entre a análise da performance e o ensino, e o perigo do trabalho excessivo, que aumenta exponencialmente o risco de lesão.

Um dos poucos estudos que se focam sobre o funcionamento da mão esquerda foi desenvolvido por Visentin, Li, Tardif, e Shan, (2015). Este projeto foi denominado de *Unravelling mysteries of personal performance style: biomechanics of left-hand position changes (shifting) in violin performance* e analisa as mudanças de posição na mão esquerda de violinistas.

Este estudo pretendeu trazer uma abordagem científica à pedagogia do violino em especial à mão esquerda. Pretende examinar e medir cientificamente através da captura de movimento a 3D aspetos relacionados mais especificamente com as mudanças de posição, como, "quando" e "o quê", um movimento que é essencial na performance do violino.

Foram analisados 540 mudanças de posição entre 6 violinistas profissionais, sobre o estudo numero 11 de Kreutzer com 12 câmaras VICON MX40 de captura de movimento a 3D (VICON Motion Systems, Oxford Metrics Ltd., Oxford, England) e que seguiam os marcadores a uma velocidade de 200 *frames*/s. Demonstrou que estes movimentos continham aspetos antropométricos e ergonómicos passíveis de serem explicados e generalizados, mas também tinham uma componente individualizada e pessoal, do domínio da interpretação. Demonstraram, portanto, que cada violinista tinha o seu próprio estilo de mudança, que afetava o início do movimento, o tempo do mesmo, e casualmente o resultado estético durante a performance.

3. Metodologia

Depois de uma ponderação mais profunda sobre as estratégias e dispositivos científicos disponíveis e passíveis de serem aplicados neste projeto científico, chegamos à conclusão que aquele que melhor resultados prometia conseguir, seria a captura de movimento em 3D Qualysis. Também foi usado, numa fase experimental, a termografia (FLIR SC7000 - captação e medição do calor emitido pelo braço e pela mão – figura 5 – não avançou porque era exigido sobreuso da atividade muscular, ao longo de um tempo mais prolongado, que saía fora do âmbito do projeto; adicionalmente, os resultados não eram tão precisos), a eletromiografia (Delsys Trigno Wireless - captação da atividade muscular através de sensores colocados no braço e na mão – mas que não tiveram os resultados pretendidos porque descolavam e não mediam os músculos internos mas sim somente os músculos superficiais – teria sido necessário colocar agulhas dentro dos músculos e optou-se por evitar este processo fisicamente desconfortável) e sensores de pressão (WalkinSense - pequenos sensores, no entanto

afastados demasiado, para a afinação milimétrica necessária com o estudo do violino; igualmente, ao serem colocados em linhas paralelas, era diferente do alinhamento das cordas do violino, que se afastam no espaço e não estão paralelas). Contudo, os melhores resultados foram efetivamente conseguidos com cinemetria e com a análise tridimensional do movimento. Isto porque consegue medir com grande detalhe a variação mais sensível da posição, em termos de angulação e altura da mão em relação à escala do violino, e consegue analisar de uma forma subtil todo o funcionamento biomecânico da mão esquerda.

Existe também a possibilidade de relacionar dados e resultados de diferentes parâmetros e nos três eixos X Y Z que correspondem aos três eixos do movimento das variadas articulações que seriam analisadas, como se pode ver na figura 2.

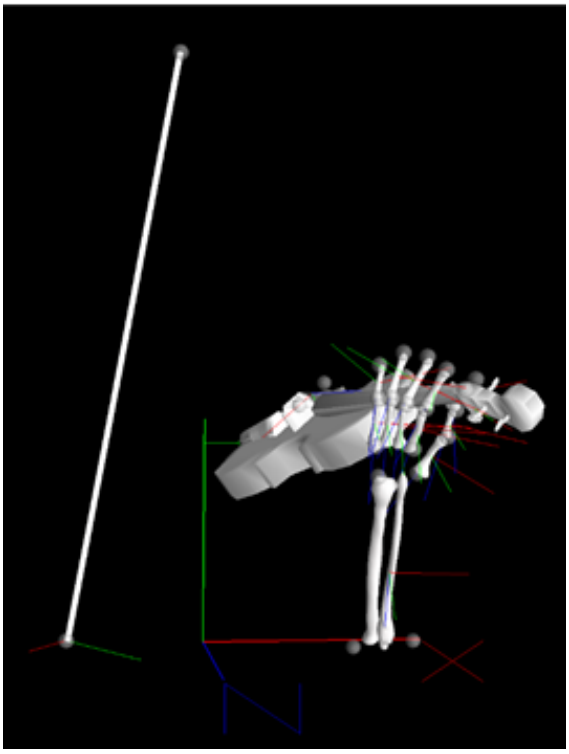


Figura 2 – Marcadores e sistema vetorial no eixo dos X, Y, Z

Deste modo delineamos um conjunto de exercícios bastante práticos e intuitivos, em forma de padrão, capazes de serem facilmente repetidos em forma de ciclo, sem haver grande necessidade de estudo.

Assim, conseguiríamos eliminar a variável da repetição (a necessidade de estudar em casa, que não existia, dado os exercícios serem acessíveis) como forma de aumentar o tipo de preparação na execução do exercício (cada exercício A, B e C foi repetido, e analisado, três vezes em laboratório).

Assim, os objetivos principais dos exercícios propostos seriam:

1. Analisar o comportamento mecânico da mão com diferentes alturas na mesma corda em relação ao ponto de contacto com o braço do violino, numa mesma posição e com os dedos numa só corda.
2. Perceber qual o comportamento digital, da mão e do punho, com diferentes angulações, numa mesma posição.
3. Analisar e comparar o posicionamento da mão nas diferentes aberturas intervalares possíveis entre o primeiro e o quarto dedo, tendo em atenção a altura e a angulação da mesma em relação ao braço do violino. Os intervalos podem oscilar entre intervalos de 4ª perfeita, intervalos de 5ª perfeita, intervalos de 6ª menor.

Em cada um dos seguintes grupos de exercícios em A, B e C (com amplitudes de dedos diferentes – "A" com amplitude de quarta perfeita, "B" de quinta perfeita e "C" com sexta menor – sendo as amplitudes medidas as que se verificavam entre o dedo indicador e o dedo mínimo) foram analisados exatamente os mesmos parâmetros por forma a tentarmos encontrar respostas para as perguntas que colocamos nos objetivos iniciais. Sendo assim as secções que privilegiamos foram, inclusive, as seguintes:

- Foi dado ênfase ao funcionamento do dedo mínimo (figuras 3 e 4), que corresponde ao dedo 4 na dedilhação violinística, isto porque é o dedo mais frágil e com menos força muscular. Para além do mais, nos exercícios de quarta perfeita, o A1 mi, A2 mi, A1 sol e A2 sol, este dedo, nos exercícios propostos e apenas num ciclo, realizam apenas um batimento. Portanto os parâmetros são mensuráveis mais facilmente por cada ciclo. É também o dedo mais solicitado nas aberturas superiores a quartas perfeitas, pois está no extremo do primeiro dedo e, portanto, está sujeito a maiores tensões.

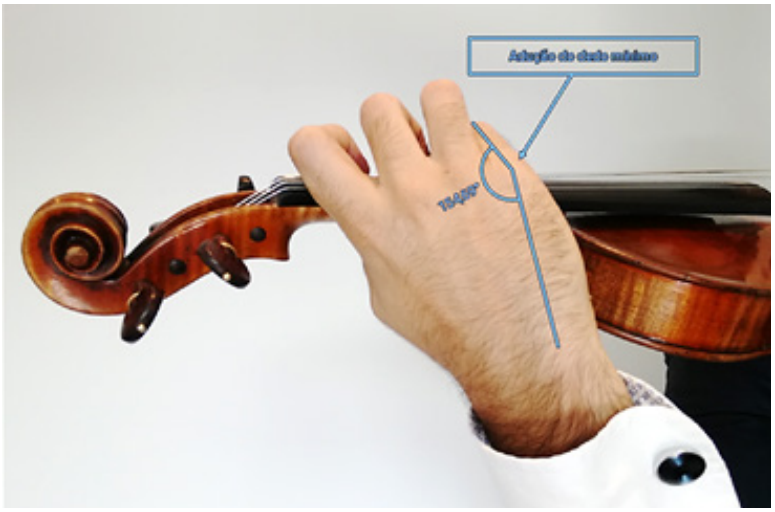


Figura 3 – Adução do dedo mínimo na performance violinística

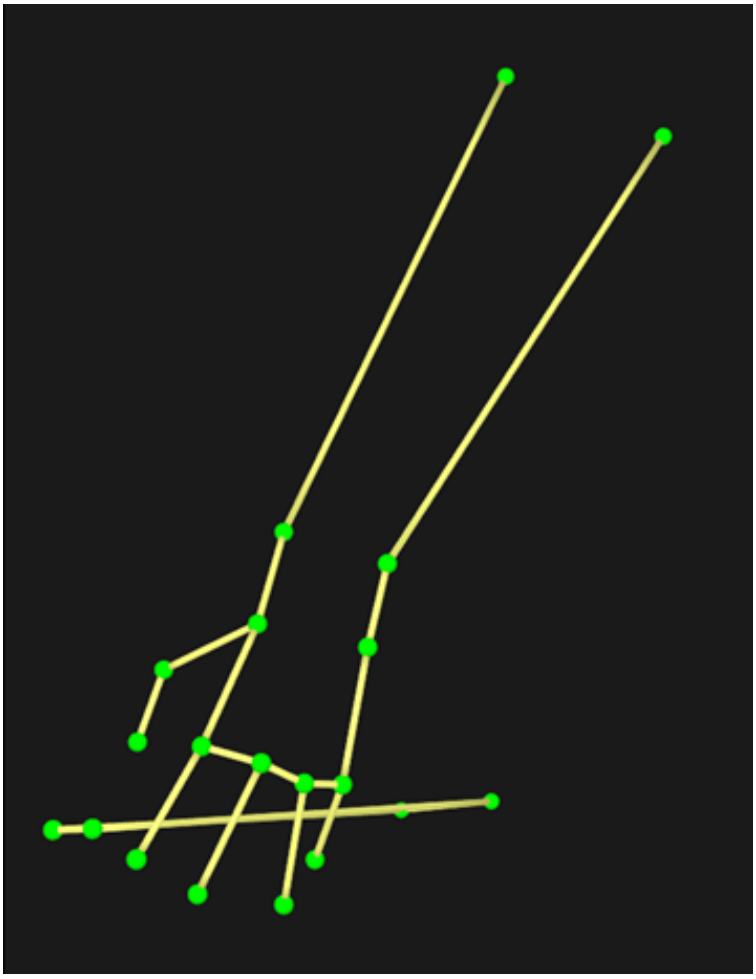


Figura 4 – Adução do dedo mínimo na *performance* violinística em imagem computorizada

Analisou-se também a abdução do dedo mínimo, a adução do dedo indicador, a abdução do indicador, a flexão do punho, a extensão do punho, a adução do punho (ou desvio ulnar do carpo), a abdução do punho (ou desvio radial do carpo), a altura da mão em relação ao braço do violino, e a angulação da mão com o braço do violino.

O estudo com o dispositivo de análise a três dimensões Qualisys teve o objetivo de analisar duas variáveis principais no funcionamento mecânico da mão em relação ao violino procurando responder às quatro questões de investigação, que passamos a enunciar:

1. Como controlar a posição da mão nas diferentes cordas, de modo a obter uma relação intervalar idêntica, com o mínimo de tensão muscular e máxima eficiência?

2. Existirá algum modo mecânico-funcional da mão capaz de controlar a abertura dos dedos, nos intervalos de grande extensão, que diminua a tensão muscular palmar e aumente a eficácia performativa?
3. Existirá algum modo estrutural e posicional para controlar a diminuição progressiva da relação intervalar, que acontece ao longo da escala do violino?
4. Será possível usar uma tecnologia específica de forma benéfica e para resolver as questões acima descritas?

Inicialmente foram analisadas doze pessoas de vários níveis, idades, géneros e dimensões antropométricas diferentes. Contudo, as variáveis que se pediam para alterar não eram executadas de uma forma regular, o que originava um desvio padrão considerável e por isso pouca precisão nos resultados. Optou-se então por fazer um estudo piloto com um só violinista (o primeiro autor deste estudo).

4. Discussão e conclusões

A termografia analisa a atividade biomecânica através da distribuição e variação do calor no organismo analisado (foi usada numa fase experimental, tais como a eletromiografia e os sensores de pressão o foram).

Já que os seres humanos na realização de atividades mais ou menos intensas libertam uma quantidade de energia maior, esta energia libertada, é traduzida na variação da temperatura corporal, que pode ser analisada de uma forma mais pormenorizada, tanto na dimensão como na variação dessa mesma energia com dispositivos ultra-desenvolvidos na detecção do calor.

Um dos dispositivos passíveis de serem utilizados para eventualmente analisar a tensão, cansaço e esforço dos músculos dos dedos da mão e do braço seria o FLIR SC7000.

Esta é uma câmara termográfica de infravermelho, um dispositivo sem contacto direto, e que deteta energia infravermelha (calor) convertendo-a num sinal eletrónico, que é então processado para produzir uma imagem térmica num monitor, que realiza cálculos de temperatura.

O calor detectado por esta câmara de infravermelhos pode ser medido com bastante precisão, o que lhe permite monitorizar não só o desempenho térmico, mas também identificar e avaliar a gravidade relativa dos problemas relacionados com o calor. Em relação às mais recentes inovações referentes à tecnologia de detecção, a incorporação de um sistema de imagiologia, funcionalidades automáticas e o desenvolvimento de software infravermelho, oferece mais soluções de análise térmica e de custo-benefício do que nunca.

Após um estudo ponderado sobre as vantagens da utilização deste dispositivo chegamos à conclusão que a termografia seria uma forma de avaliar o cansaço muscular.

Em relação a este método de medição termodinâmico, tentamos uma abordagem analisando a variação do calor no dorso da mão na altura da performance.

Executamos um exercício bastante simples, mas repetido consequentemente para percebermos a fadiga muscular e as variações de cor na imagem.

Isto é somente um exemplo de como a tecnologia poderá ajudar a resolver problemas do dia-a-dia de músicos de topo e outros (figura 5).

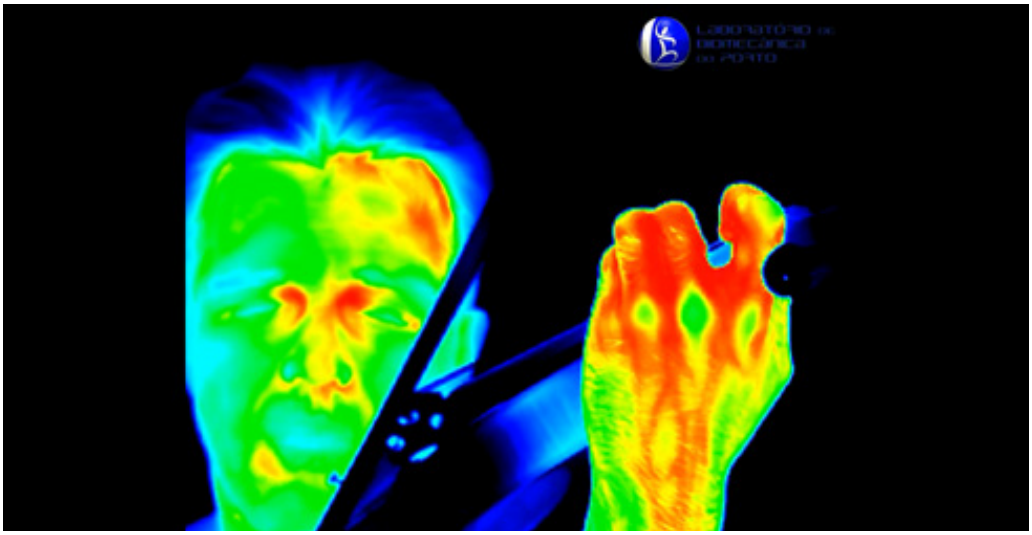


Figura 5 - Imagem resultante da termografia da performance da mão esquerda

Decisivo para este projeto científico foi usar o dispositivo Qualisys Oqus Camera series que analisa o movimento tridimensionalmente.

Este sistema também é utilizado amplamente por outras áreas cientificas e industriais como no desenvolvimento de automóveis. Por exemplo, o design de interiores pode ser melhorado com o uso de tecnologia de captura de movimento tanto na ergonomia como para avaliar o conforto e segurança para o motorista. Igualmente em áreas como a Animação e a Realidade virtual usam captura de movimento para aumentar o realismo e um ambiente mais interativo.

Este dispositivo laboratorial contém os seguintes elementos constituintes e suas características de funcionamento:

- O componente central do sistema de captura de movimento Qualisys tem várias câmaras com infravermelhos ópticos, Oqus, ao redor do sujeito a ser analisado em diferentes alturas e que emitem um feixe de luz infravermelha.
- tem vários marcadores pequenos e leves, retro-reflexivos, que são colocados num objeto/pessoa, nas zonas articulares.
- As câmaras emitem uma luz infravermelha para os marcadores que refletem a luz de volta para o sensor da câmara. Esta informação é então usada para calcular a posição de alvos com uma elevada resolução espacial.
- Para fazer uma reconstrução 3D de dados 2D o sistema precisa de ser calibrado. Uma varinha é movida em torno do volume de captura, enquanto um objeto de referência fixo define o sistema de coordenadas do laboratório.
- Com um sistema Qualisys, o processo de calibração é de apenas um exercício 10-20 segundos.

- Uma vez que o sistema é calibrado e que a pessoa está a ser analisada usa um conjunto de marcadores reflexivos, o movimento pode ser gravado, processado e analisado.

Os procedimentos metodológicos incluem, num primeiro momento, a filmagem de um objeto de calibração e do movimento em estudo, por câmaras colocadas em diversos planos (estudo tridimensional).

Numa segunda fase, é utilizado um sistema vídeo-analógico de medição do movimento, ou seja, um programa informático, através do qual se captará os dados por meio de um procedimento manual ou automático de digitalização dos pontos de referência anatómica do indivíduo, em cada fotograma. Este procedimento tem como objetivo a criação de imagens animadas de modelos espaciais, isto é, de um modelo que represente o sujeito através de segmentos rígidos e articulados, correspondentes aos diversos segmentos anatómicos a realizar na tarefa em estudo.

Após a digitalização das imagens, os dados serão tratados. Isto é, através de determinadas técnicas de filtragem, as informações obtidas serão corrigidas, aumentando a fiabilidade dos resultados.

Finalmente, serão recolhidos os dados de interesse para o estudo sob a forma numérica, gráfica ou pictórica.

Depois de uma reflexão sobre a abordagem cinemática e da capacidade dos dispositivos existentes no LABIOMEP consideramos que este sistema de medição e a plataforma de análise de movimento Qualisys seria a forma mais indicada para investigar os objetivos propostos para este projeto.

Pensamos assim ter discutido sobre como a tecnologia poderá servir a sociedade e a música – e, em especial, os profissionais de música (figura 6).

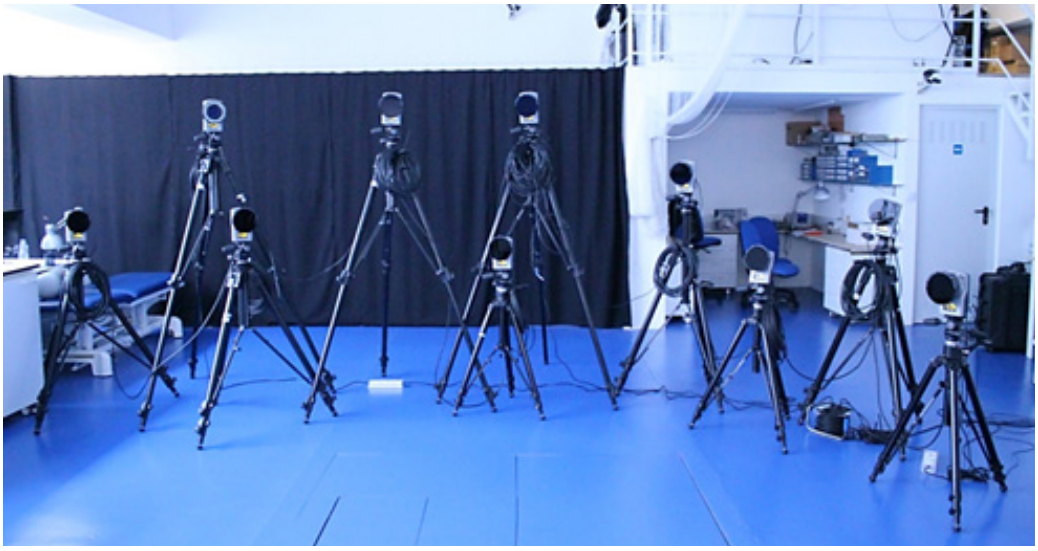


Figura 6 – Disposição das câmaras Qualisys no Labiomep

O objetivo de encontrar uma técnica mais eficiente, aplicável desde um nível mais de iniciação, até a um nível mais avançado, foi alcançado, com o uso da tecnologia de captura e análise de movimento 3D. Diminuiu-se o tempo de estudo, o esforço, aumentou-se o conforto na *performance* e a eficiência dos movimentos, podendo ser aplicado no ensino desde o início da aprendizagem. Mais importante é o potencial de diminuir as lesões, com a nova técnica, uma técnica mais natural, que consiste em ter, acima de tudo, um alinhamento dos segmentos ossiculares mais uniforme. Certas posições, em determinados trechos musicais, têm tendência a ter uma posição da mão esquerda que poderá ser menos ergonómica, o que agora se evita com os resultados apurados.

Referências

Frank, A., Mühlen, C.A. (2007). Queixas músculo-esqueléticas em músicos: prevalência e fatores de risco. *Revista Brasileira de Reumatologia*, *47*(3), 188–196.

Großhauser, T., Großekathofer, U., Hermann, T. (2010). New sensors and pattern recognition techniques for string instruments. Proceedings of the 2010 Conference on New Interfaces for Musical Expression++ (NIME 2010), Sydney, Australia, 271–271.

McCrary, J.M., Halaki, M., Sorkin, E. (2016). Acute Warm-up Effects in Submaximal Athletes: An EMG Study of Skilled Violinists. *Official Journal of the American College of Sports Medicine*.

Moraes, G. F., Papini, A. A. (2012). *Desordens musculoesqueléticas em violinistas e violistas profissionais: revisão sistemática*. Itabira, MG, Brasil: Universidade Federal de Itajubá Campus.

Silva, E., Au-Yong-Oliveira, M., Fonseca, P., Garganta, R., Bochmann, C. (2018). The musician as entrepreneur – multidisciplinary innovation with performing violinists – achieving a sustainable competitive advantage via the biomechanical enhancement of the left hand. Proceedings of the ECIE 2018, University of Aveiro, Aveiro, Portugal, 20-21 September. Costa, C., Au-Yong-Oliveira, M., e Amorim, M. (Coordenação), pp.984-994.

Turner-Stokes, L., Reid, K. (1999). Three-dimensional motion analysis of upper limb movement in the bowing arm of string-playing musicians. *Clinical Biomechanics*, July, *14*(6), 426–433.

Visentin P. (2008). Informing music teaching and learning using movement analysis technology. *International Journal of Music Education*, *26*, 73–87.

Visentin, P. (2010). A pilot study on the efficacy of line-of-sight gestural compensation while conducting music. *Perceptual and Motor Skills*, *110*, 647–653.

Visentin P., Li, S., Tardif, G., Shan, G. (2015). Unraveling mysteries of personal performance style; biomechanics of left-hand position changes (shifting) in violin performance. *PeerJ*, pre-print.

Visentin, P., & Shan, G. (2011). Applications of EMG Pertaining to Music Performance - A Review. *Arts BioMechanics*, *1*(1), 15–32.

Young, D. (2003). Wireless sensor system for measurement of violin bowing parameters. Proceedings of the Stockholm Music Acoustics Conference, August 6-9 (SMAC 03), Stockholm, Sweden, 111–114.

Young, D. (Jan 2002a). The Hyperbow: A Precision Violin Interface. International Computer Music Association.

Young, D. (2002b). The Hyperbow Controller: Real-Time Dynamics Measurement of Violin Performance. Proceedings of the 2002 Conference on New Instruments for Musical Expression (NIME-02), Dublin, Ireland, May 24–26.

Young, D. (2007). *A Methodology for Investigation of Bowed String Performance Through Measurement of Violin Bowing Technique*. PhD Program in Media Arts and Sciences, School of Architecture and Planning, Massachusetts Institute of Technology.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 13/10/2018
Aceitação/Acceptance: 11/12/2018

Análisis de la estructura intelectual y la evolución de la investigación en la interacción humano -computador: un análisis bibliométrico

Gabriel Mauricio Martínez-Toro[1], Dewar Rico-Bautista[2], Efrén Romero-Riaño[3], Claudia Jazmin Galeano-Barrera[4], Cesar D. Guerrero[5], Jorge A. Parra-Valencia[6]

gmartinez714@unab.edu.co, dwricob@ufpso.edu.co, efren.romero@saber.uis.edu.co, claudia.galeano@udes.edu.co, cguerrer@unab.edu.co, japarra@unab.edu.co

[1]Facultad de Ciencias Económicas, Administrativas y Contables. Universidad Autónoma de Bucaramanga, Cra 34#41-51, 680003, Bucaramanga, Colombia.

[2]Departamento Sistemas e Informática. Universidad Francisco de Paula Santander Ocaña, Sede Algodonal Vía Acolsure, 546551, Ocaña, Colombia.

[3]Escuela de Estudios Industriales y Empresariales. Universidad Industrial de Santander, 680002, Bucaramanga, Colombia.

[4]Centro de Estudios Tecnológicos. Universidad de Santander, 68OO3, Bucaramanga, Colombia.

[5]Centro de Excelencia y Apropiación en IoT - Nodo Oriente. Universidad Autónoma de Bucaramanga, Avenida 42 No. 48 – 11, 680003, Bucaramanga, Colombia.

[6]Grupo de Investigación Pensamiento Sistémico. Universidad Autónoma de Bucaramanga, Avenida 42 No. 48 – 11, 680003, Bucaramanga, Colombia.

Pages: 363–378

Resumen: La bibliometría, es la aplicación de modelos y métodos cuantitativos para la comprensión de las disciplinas de investigación. Los principales conceptos que se aplican dentro del desarrollo de este tipo de análisis se encuentran asociados al análisis de redes e incluyen los enlaces, número de enlaces y fuerza de enlace dentro de las redes de co-ocurrencia, co-citación, coautoría y vínculo bibliográfico. Bajo estos conceptos y en aras de visualizar la evolución de la interacción humano – computador (HCI) se desarrolla un análisis bibliométrico de la literatura publicada para conocer características de este cuerpo de conocimiento y su desarrollo en el paso del tiempo.La comprensión de la evolución y las tendencias dentro de las disciplinas de investigación requiere la combinación de métodos y herramientas que garanticen calidad y replicabilidad dentro de estos ejercicios. En ese contexto, este artículo implementa un análisis basado en una estrategia de búsqueda sistemática y exhaustiva, de la evolución de la literatura de HCI, utilizando una combinación de métodos cuantitativos, que incluye un análisis de calidad para la delimitación de la muestra de análisis de las publicaciones identificadas.

Palabras-clave: Interacción humanocomputador; Visualización científica;análisis de redes; Bibliometría.

Analysis of the intellectual structure and evolution of research in human-computer interaction: a bibliometric analysis

Abstract: Bibliometrics is the application of quantitative models and methods for the understanding of research disciplines. The main concepts that are applied within the development of this type of analysis are associated to the analysis of networks and include the links, number of links and force of link within the networks of co-occurrence, co-citation, co-authorship and bibliographic links. Under these concepts and in order to visualize the evolution of human-computer interaction (HCI), a bibliometric analysis of the published literature is developed in order to know the characteristics of this body of knowledge and its development over time, understanding the evolution and trends within the research disciplines requires the combination of methods and tools that guarantee equality and replicability within these exercises. In that context, this article implements analysis is based on a systematic and exhaustive search strategy of the evolution of the HCI literature, using a combination of quantitative methods, which includes a quality analysis for the delimitation of the simple analysis of the identified publications.

Keywords*:* computer human interaction; Scientificvisualization; Network analysis; Bibliometrics

1. Introducción

La tecnología hace parte de la vida cotidiana, ésta es usada para fines tan variados como la comunicación, el trabajo, entretenimiento, enseñanza, aprendizaje, entre otros. La comunicación entre el hombre y la máquina, para hacer posible su finalidad, ha sido motivo de estudio en los años recientes, en los cuales la mediación entre el hombre y la máquina aumenta a causa de las aplicaciones computacionales que requieren cada vez más interacción para lograr un aprovechamiento eficiente de su potencial (Rautaray & Agrawal, 2012).La interacción entre humano y la máquina ha tenido progresos y es definida como la manipulación de objetos gráficos en una interfaz usando algún artefacto que lo haga posible(Seufert et al., 2015). El teclado y mouse han sido comúnmente usados para mediar la interacción hombre-máquina(Turk, 2014), y estas tecnologías, a pesar de ser populares, se han convertido en un cuello de botella debido a las limitaciones que estos presentan con la interactividad con computadores, para lo cual, se han desarrollado técnicas que permiten la interpretación de gestos por parte de los computadores bajo dos vertientes, el uso de dispositivos vestibles (*wearables*) y los lectores de gestos basados en visión (Yeo, Lee, & Lim, 2015). Gamificación(Seaborn & Fels, 2015), atención a la superación de limitaciones físicas (Mengüç et al., 2014) y cognitivas (González-Ortega, Díaz-Pernas, Martínez-Zarzuela, & Antón-Rodríguez, 2014), comunicación no verbal (Rautaray & Agrawal, 2012), reconocimiento de objetos, dinámica humana (Herath, Harandi, & Porikli, 2017), experiencia del Usuario (UX) (Lallemand, Gronier, & Koenig, 2015), entre otras, han centrado la atención de investigadores para buscar el desarrollo de tendencias encaminadas a buscar mejorar la interacción con los computadores de una manera natural siendo la comunicación intuitiva la premisa a seguir.

El reconocimiento tridimensional es una de las tendencias de mayor interés en la investigación de la visión por computador(Pisharady & Saerbeck, 2015), en lo que respecta a patrones e interacción persona-computador, enmarcado en los enfoques de reconocimiento en 3D de Cheng, Yang, & Liu(2015): 1) reconocimiento del lenguaje de

signos; 2) manipulación virtual; 3) asistencia diaria; 4) verificación de la palma de la mano; 5) juegos de azar; y 6) la emergente interacción hombre-robot (HRI), expresado por Tsarouchi, Makris, & Chryssolouris (2016). Debido a su importancia, carácter emergente y el crecimiento de la investigación dentro de este cuerpo de conocimiento y con el objeto de visualizar la evolución en investigación asociado a la interacción humano– computador (HCI), se implementa un análisis bibliométrico de la literatura publicada dentro de la base de datos *Web of Science* (WOS) de *Clarivate Analytics*, para conocer sus características y tendencias en el paso del tiempo.La comprensión de la evolución dentro de las disciplinas de investigación de HCI, requiere la combinación de métodos y herramientas que garanticen calidad y replicabilidad. En ese contexto, este artículo implementa un análisis basado en una estrategia de búsqueda sistemática de la evolución de la literatura, utilizando una combinación de métodos cuantitativos, que incluye un análisis de calidad soportado en herramientas estadísticas, para la delimitación de la muestra de las publicaciones identificadas, con base a un análisis de Pareto (Takey & Carvalho, 2016), y un análisis de la evolución del cuerpo de conocimiento, basado en indicadores bibliométricos (Garfield, 1972; PRITCHARD, 1969),para el análisis, comprensión y visualización de las redes y relaciones entre la literatura identificada.

La bibliometría atiende cinco grandes grupos de fenómenos y su análisis: primero, en el análisis de citación para individuos, instituciones, países y disciplinas (White & Mccain, 1998); segundo el análisis de co-citación(Small, 1973) para el mapeo de países, individuos, escuelas de pensamiento; tercero la productividad científica (Sink, Tuttle, & Devries, 1984) por medio de análisis de impacto (Narin, 1987) de individuos, países e instituciones; cuarto, el análisis de productos específicos como artículos o patentes, basados en características del autor, patrones de citación y características longitudinales (Terry, 1994) y quinto, co-ocurrencia de palabras, referencias (Naghizadeh, Elahi, Manteghi, Ghazinoory, & Ranga, 2015). Este artículo se centra en el análisis de redes de co-citación y de co-ocurrencia de palabras.Debido al volumen de publicaciones en el área, desarrollar una revisión sistemática de la evolución de las publicaciones sobre la investigación en HCI, requiere de métodos y herramientas tecnológicas de soporte. En este contexto se desarrolla un estudio longitudinal y exhaustivo de la evolución de la literatura sobre HCI, utilizando un método cuantitativo, basado en un estudio bibliométrico, que complementa empíricamente las revisiones bibliográficas cualitativas existentes, por medio delprograma VOS VIEWER.

Este estudio contribuye a avanzar más allá de los análisis de citaciones tradicionales y emplea a VOS VIEWER para identificar y visualizar la estructura intelectual, las relaciones, patrones y la dinámica de la investigación en HCI en la literatura científica recuperada de la base de datos WOS.

2. Metodología

Con el fin de establecer la evolución de la investigación en HCI, se apropia un enfoque de revisión sistemática de literatura, apoyados en métodos y herramientas bibliométrica. La integración de estos métodos es complementaria(Takey & Carvalho, 2016), considerando que estos facilitan la comprensión de las características de las disciplinas(Chiu & Ho, 2007). La metodología consta de tres fases, (i)selección artículos, (ii) procesamiento de datos y (iii)análisis bibliométrico, (Acs, Anselin, & Varga, 2002). Como métodos se

implementan el análisis de co-citación y co-ocurrencia de palabras, los cuales sirven como medida de relación entre documentos (Small, 1973), visualizar la estructura de las disciplina y campos de investigación (White & Mccain, 1998), y detectar frentes de investigación (Fujita, Kajikawa, Mori, & Sakata, 2014) entre otros.

La co-citación como medida de relación establecida por los autores que citan, se define comola frecuencia con la que dos trabajos de literatura anterior, se citan juntos en la literatura posterior(Small, 1973). Para ser fuertemente co-citado, un gran número de autores debe citar los dos trabajos anteriores.La fuerza de la co-cita mide el grado de relación entre los documentos(Koehler, 2001). La frecuencia de co-citación, corresponde al número de veces con la que se citan dos documentos de manera simultánea (Liu, Yin, Liu, & Dunford, 2015). La unidad de análisis de los estudios de co-citación comprende autores, fuentes y referencias o artículos citados (van Eck & Waltman, 2010). En la visualización de redes de co-citación, por ejemplo, entre fuentes, la distancia entre dos revistas, indica su relación en términos de enlaces de citas compartidas. En ese sentido, cuanto más cerca se encuentran dos revistas, más fuerte es su relación. El vínculo de citación conjunta, es un enlace entre dos elementos,donde ambos son citados por el mismo documento(Eck & Waltman, 2013).

La co-ocurrencia de palabras clave indica la relación de proximidad entre dos o mas palabras clave dentro de un documento, se representa mediante las gráficas de redes de enlace entre términos. Hay dos atributos de ponderación estándar, a los que se hace referencia, enlaces y fortaleza. Para un término determinado, los atributos enlace y fortaleza indicanel número de enlaces de este con otros elementos y la fortaleza de los enlaces de este elemento con otros elementos(Eck & Waltman, 2013). El análisis de palabras clave es de amplia difusión dentro del análisis de contenido, tomando dentro de sussupuestos que un conjunto de palabras, señala orefleja el contenido central de la literatura de investigación(Naghizadeh et al., 2015).

2.1. Delimitación de la muestra de artículos y bibliometría

Para extraer el conjunto de artículos, se obtiene una muestra dentro de revistas indexadas en WOS, tomando como referente el factor de impacto, el cual es uno de los indicadores bibliométricos de mayor relevancia (Pudovkin & Garfield, 2002), calculado mediante el índice JCR (*Journal Citation Report*). Se implementa esta búsqueda dentro de WOS, teniendo en cuenta que comprende revistas de otras bases de datos, que implementan procesos revisión y selección de artículos tipo –*blind peer review*-, así como evaluación y seguimiento a la calidad y visibilidad de las publicaciones, por medio de indicadores de calidad e impacto. En adición, WOS permite la exportación de metadatos para el análisis bibliométrico, como: palabras clave de autor, palabras clave *plus*, filiación institucional de los autores, número de citaciones y las referencias citadas. Todos los artículos de la base de datos desde el año 2013 hasta el 2017, fueron tenidos en cuenta dentro de la búsqueda inicial.Con el fin de extraer un censo de artículos y revisiones relacionados, se genera la expresión de búsqueda TS=(“human computerinteraction”), generando como resultado 1970 artículos. De este primer conjuntode documentos, mediante exploración de los títulos, 54 fueron excluidos, por estar relacionados de forma exclusiva a temas no asociados a la interacción con máquinas.

2.2. Herramientas

La visualización de la evolución y las características de campos de investigación, por medio de software, es una práctica de creciente uso entre investigadores(Garcia Martinez & Walton, 2014), basados en los principios y conceptos de análisis de redes sociales(Borgatti & Halgin, 2011). El inicio del análisis y visualización de redes de bibliometría, se basa en los antecedentes desarrollados por el software Pajek(Martínez-Torres, 2013) para el análisis de grandes redes y que han sido probados para el estudio de: la estructura y evolución de tecnologías(Saviotti, 2005) como el software Vantage Point, la evolución de sistemas de innovación mediante CiteSpace(Liu et al., 2015), entre otros. El presente artículo, implementa el enfoque de la herramienta para la visualización y el mapeo de redes bibliométricas VOS VIEWER (Van Eck&Waltman, 2010).Al igual que con otras herramientas de software para visualizar redes, VOS VIEWER permite producir y analizar la red de co-ocurrencias de palabras clave y categorías de temas (análisis de co-palabras), redes de co-citación y coautoría entre autores, países e instituciones. Los métodos de análisis de co-citación que ofrece VOS VIEWER son el análisis de co-citación de documentos, el análisis de co-citación de autores y el análisis de co-citación de revistas, los cuales se complementan entre sí (Jose M.Merigo, Fabio Blanco-Mesa, Anna M. Gil-Lafuente, 2016), facilitando a los usuarios la detección y visualización de los patrones y tendencias estructurales y temporales, de un dominio de conocimiento en particular.

La visualización y análisis de las redes de co-citación de documentos, permiten extraer datos sobre las conexiones y su naturaleza, entre documentos co-citados. De otra parte, el estudio de la red de co-citación de autores, se centra en las interrelaciones entre los mismos, con base a lo cual es posible identificar grupos de investigadores intelectualmente interrelacionados y sus publicaciones. El grupo de documentos citados, se entiende como la base de conocimiento colectivo de una disciplina o campo de investigación en particular.

3. Resultados empíricos

Dentro de la descripción de la estructura y evolución del campo de investigación en HCI, se identifican varias características. Con base al número de publicaciones anuales, se identifica una tendencia creciente que va desde 305 (2014) hasta 509 en 2017. De acuerdo con la frecuencia de aparición de la filiación institucional de los autores de los artículos de la muestra, se identifican dos liderazgos dentro de la disciplina, uno ubicado geográficamente en Europa representado por la Universidad de Londres (41 documentos) y otro localizado en Estados Unidos de América, con la Universidad de California (29 documentos). Esta geografía es complementada por la Academia de Ciencias con 28 documentos (Asia), CNRS con 25 documentos (Europa) y la Universidad Estatal de Florida 24 documentos. Los autores con mayor número de artículos son Kim J, del InstBioMed IT Energy IT & Smart IT TechnolBest, DeptElect&ElectEngn y Lee M de KyungpookNatlUniv, SchElectEngn, ambos de Corea del Sur. A nivel de número de citaciones, Rautaray (217), Seaborn (170) y Seufert (157), son los autores con mayor representatividad.

3.1. Áreas de conocimiento

La composición de las áreas de conocimiento dentro de un campo de investigación muestrala confluencia e interacción de las disciplinas involucradas. En ese orden de ideas se presenta la figura 1, que describe la frecuencia de ocurrencia de cada una de las áreas de conocimiento de WOS, dentro de la muestra de artículos de HCI.

El análisis de co-ocurrencia de categorías temáticas permite detectar las disciplinas involucradas en el desarrollo intelectual de un determinado dominio de conocimiento y puede ser visualizado usando la herramienta de análisis de WOS soportada por *Clarivate Analitics*. El área de cada cuadro es proporcional a la frecuencia de co-ocurrencia de la respectiva categoría de sujeto, y el pequeño número localizado en la esquina superior izquierda de cada cuadro, es la frecuencia de co-ocurrencia del área.

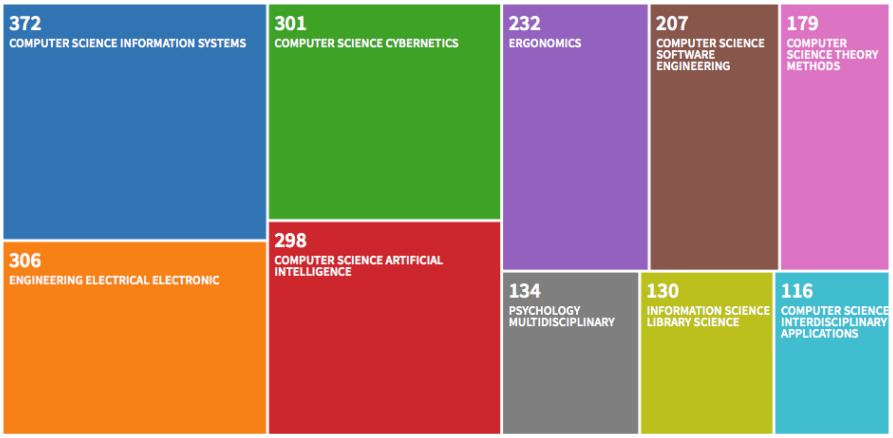


Figura 1 – Distribución de áreas de conocimiento más frecuentes dentro de la investigación en HCI- Fuente: WOS

3.2. Análisis de palabras clave

El análisis de co-ocurrencia de términos, se basa en el estudio comparativo de las palabras clave de autor y de todas las palabras clave. Esta última categoría incluye los términos reportados por el autor, más las palabras indexadas por las revistas dentro de su proceso editorial. Dentro de los resultados obtenidos por el programa VOS VIEWER se generan 3 visualizaciones de las palabras clave agrupadas inicialmente por el análisis de co-ocurrencia de las palabras clave seleccionadas por los autores (figura 2), en segundo momento el análisis de co-ocurrencia de las palabras clave de indexación proporcionadas por WOS (Figura 3) y finalmente el análisis de co-ocurrencia de la palabras clave condensadas (figura 4). Las palabras clave dan sentido al contenido central de los artículos, de tal manera que dan cuenta de los temas de investigación y las fronteras de conocimiento de este.

Dentro del análisis de co-ocurrencia, por parte de las palabras clave de autor, se resalta que dentro de la muestra de artículos “human-computerinteraction” es la palabra clave

que tiene el mayor número de ocurrencias con un total de 608 veces, seguida por “human computerinteraction” con 245 ocurrencias, “usability” con 57, “gesture recognition” con 52 y “human-computer interaction (hci) con 52 ocurrencias. Se observa que, dentro de los 10 primeros resultados, 5 de ellos refieren a HCI, pero debido a variaciones de escritura, el uso combinado del texto y la abreviatura hci, o únicamente la abreviatura, se presentan

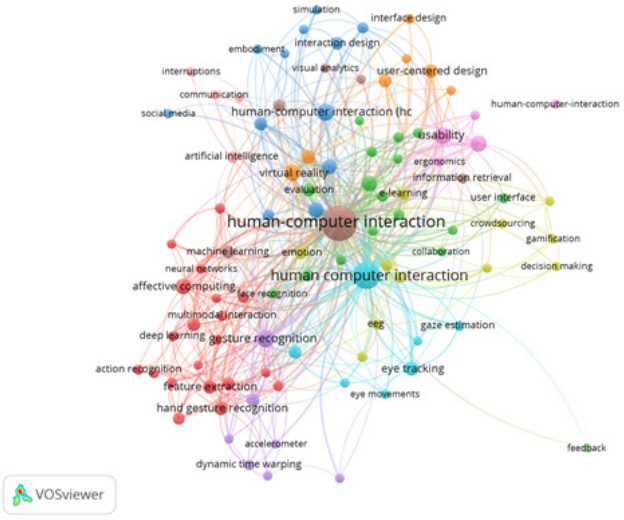


Figura 2 – Gráfico de red de co-ocurrencia de palabras clave de autor dentro de la muestra de HCI 2014-2018Fuente: VOSVIEWER

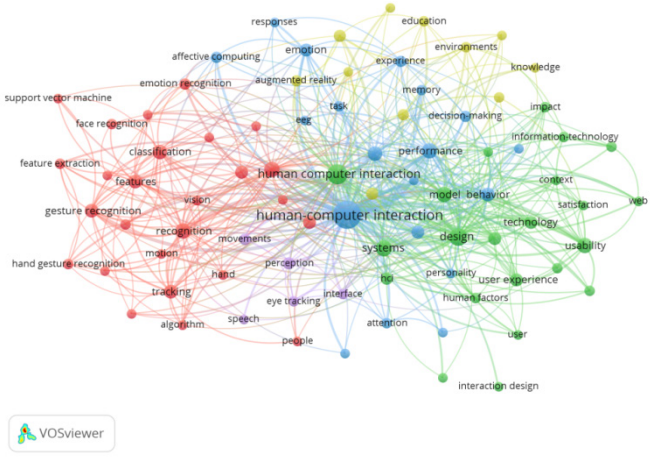


Figura 3 – Gráfico de red de co-ocurrencia de palabras clave de autor dentro de la muestra de HCI 2013-2017Fuente: VOSVIEWER

ocurrencias separadas, las cuales al combinar los resultados indican que se tendrían 938 ocurrencias, orientando que "usability" y "gesture recognition" son las temáticas de mayor exploración dentro de hci. Analizando la fortaleza de los enlaces, los cuales describen los clústeres de información, "Human-computer interaction" tiene 499 relaciones con otras palabras clave, seguida por "human computerinteraction" con 204 enlaces, "usability" con 93 enlaces, "Gesture recognition" con 84 y "user experience" con 63 enlaces.

En el análisis de palabras claves de indexación, proporcionadas por WOS, en la figura 3, se visualiza la red generada por medio de VOS VIEWER. De este resultado, a través del análisis de redes proporcionado por la herramienta, se obtiene que para el análisis de co-ocurrencia para la indexación, las palabras clave "Human-computerinteraction" y "Human computerinteraction" son las que tiene un mayor número de ocurrencias con 743 y 245 respectivamente. En esta red surgen palabras que no tuvieron aparición en el análisis previo de autor, como lo son "design", "systems", "performance", "recognition" y "model", lo cual indica la orientación de la indexación y la manera en que dirige estas temáticas a términos relacionados con resultados de la interacción humano-computador. Adicionalmente la fortaleza de los enlaces da cuenta de la importancia y la coocurrencia de palabras como "design", "model" y "system".

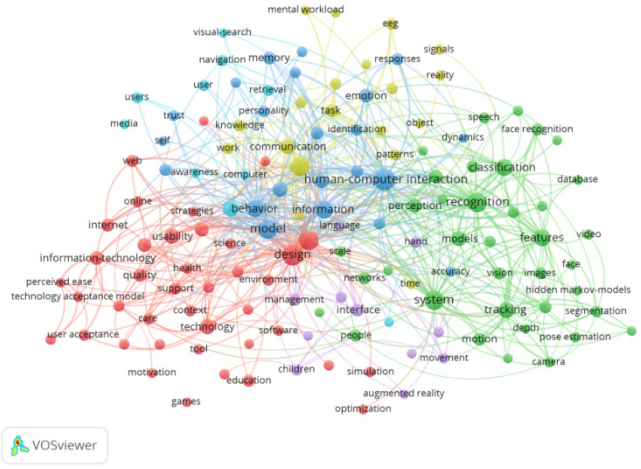


Figura 4 – Gráfico de red de co-ocurrencia de palabras clave condensadas (keyword plus) dentro de la muestra de HCI 2013-2017Fuente: VOSVIEWER

La figura 4 muestra la red de palabras condensadas, de la misma manera que los análisis de palabras clave de autor y palabras claves por indexación y en coherencia a lo anterior, la palabra "human-computerinteraction" con 148 ocurrencias, "design" con 135 ocurrencias, "systems" con 106 ocurrencias, "system" con 102 ocurrencias y "model" con 96 ocurrencias hacen parte del top 5 de esta lista. En lo que corresponde con la fortaleza de los enlaces "human-computerinteraction" con 354 conexiones, "design" con 330 conexiones, "performance" con 260, "model" con 238 y "systems" con 234 encabezan la lista.

De los análisis se infiere que HCI está fuertemente relacionado con la usabilidad, el reconocimiento de gestos, el diseño para que esto se haga posible, sin dejar de lado el rendimiento de las aplicaciones, por medio del desarrollo de modelos y sistemas que faciliten la experiencia de los usuarios finales.

3.3. Red de co-citación de revistas

Mediante la visualización de co-citación se puede determinar la relación entre las revistas a través de la distancia que se puede observar en el gráfico generado por VOS VIEWER, indicando que entre más cerca se encuentren, su relación es mayor. Respecto al vínculo entre ellas, se establecen los enlaces entre dos elementos, los cuales fueron citados en el mismo documento y de esta manera se logra identificar la frecuencia con el número de veces en que fueron citados entre revistas. A continuación, se puede observar la co-citación entre las revistas en una ventana de 5 años (2013-2017).La red de co- citación de revistas fusionadas en la colección de artículos de investigación sobre HCI, contiene 147 revistas.Después de llevar los metadatos al software VOSVIEWER, se puede observar la red simplificada de co-citación. La figura 5 muestra los 4 clúster representados en colores, el primer clúster es el más fuerte, representado por el color rojo y compuesto por 46 revistas, donde la revista "Lect notes computsc", resulta ser la revista que tiene mayor número de relaciones con su mismo clúster y con los otros, la segunda revista más fuerte dentro del clúster es la Ieee t pattern anal y en tercer lugar Pro cvprieee.

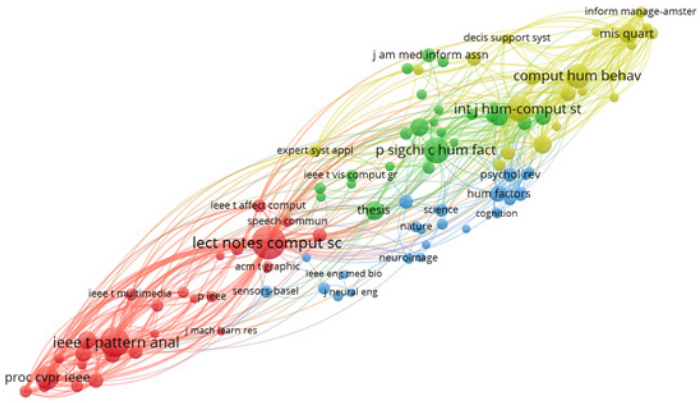


Figura 5 – Gráfico de red co-citación de fuentes 2013-2017Fuente: VOSVIEWER

El segundo clúster se encuentra representado por el color verde y está compuesto por 37 revistas y entre las revistas más representativas se encuentran Science, Psycholsci, Plosone, ergonomics. El tercer clúster, representado por el color azul y con 32 revistas muestra que la revista Computhumbehav tiene mayor número de relaciones, seguidamente la revista Int j hum- computst y en tercer lugar la revista J Perssocpsyschool; finalmente el cuarto clúster en color amarillo y con 32 revistas muestra como sus revistas más fuertes a P Sigchi c humfact ,thesis y communacm.

Origen	Citaciones	Fuerza total del enlace
Lect notes computsc	1508	29310
Ieee t pattern anal	1097	25343
P sigchi c human fact	868	13557
Int j hum-computst	810	18001
Computhumbehaw	758	17677
Proccvprieee	694	23800
Thesis	486	6834
Communacm	456	7699
Patternregogn	454	11697
J perssocpsychol	450	9979
Interactcomput	445	9632
Mis quart	441	13264
Comput vis imageund	366	10670
Humfactors	344	6193
Int j hum-computint	320	6401
Imagevisioncomput	310	8883
Computeduc	304	5106
Int j computvision	286	7708

Tabla 1 – Ranking de las 18 revistas más co-citadasFuente: VOSVIEWER

Ahora, al observar la tabla 1, nótese que las cinco primeras revistas con mayor cantidad de citaciones son:Lect notes computsc con 1508, Ieee t pattern anal con 1097, P sigchi c human fact con 868, Int j hum-computst con 810 y Computhumbehaw758. De igual manera las cinco primeras revistas con mayor fuerza del enlace son Lect notes computsc con 29310, Ieee t pattern anal con 25343, Int j hum-computst con 18001, Computhumbehaw con 17677 y Proccvprieee con 23800. Estos resultados indican que Lect notes computsc, no sólo es la revista citada con más frecuencia, sino también la revista académica con mayor fuerza en enlace.

3.4. Red de cocitación de documentos.

El análisis de la co-cita estudia las estructuras de la investigación científica, basadas en citas y citas compartidas. El análisis de co-citas permite identificar grupos de científicos y sus publicaciones, y extraer conclusiones sobre la estructura interna de disciplinas de investigación, escuelas o paradigmas. El enfoque de estudio de co-cita de documentos, se basa en la premisa de que el indicador más válido y confiable de una escuela de investigación y su evaluación científica, son sus documentos en órganos o publicaciones con procedimientos de revisión por pares (Small, 1973).

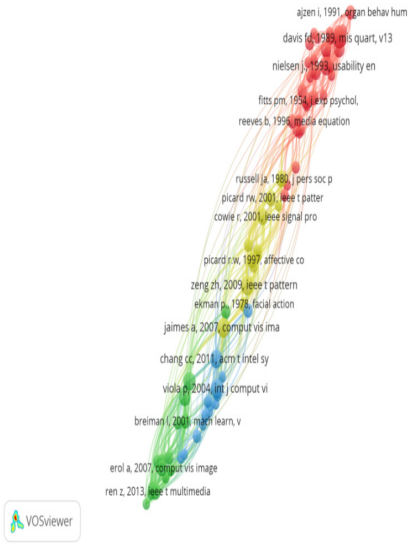


Figura 6 – Gráfico de Red de co-citación de artículos en la muestra de HCI Fuente: VOSVIEWER

Referencias citadas	Citaciones	Fuerza total del enlace
David fd, 1989,	49	189
Pavlovic vi, 1997,	41	138
Venkatesh v. 2003	31	130
Zengzh, 2009	43	128
Mitra s, 2007	34	126
Russell ja, 1980	33	121
Ren z, 2013	30	116
Jaimes a, 2007	42	115
Fornell c, 1981	28	110
Panticrw, 2003	21	109
Venkatesh v. 2000	20	106
Picardrw, 2001	28	103
Davis fd, 1989	18	102
Agarwal r, 2000	18	98
Venkatesh v. 2000	15	98
Cowie r, 2001	31	97
Erol a, 2007	29	91
Van derHeijden h, 2004	13	88

Tabla 2 – Ranking de las 18 referencias más co-citadas Fuente: VOSVIEWER

Una red de co-citación de documentos es una red de referencias citadas, y es útil en estudios de la estructura, dinámica y desarrollo de paradigmas de un campo de investigación dado. La Figura 6 muestra la estructura de la red saliente de las referencias citadas derivadas del comportamiento de cita de los autores que escribieron.Representado en colores, el primer clúster es el más fuerte, representado por el color rojo, se destaca el documento escrito por David, en 1989. El segundo clúster se encuentra representado por el color amarillo, donde se destaca el documento cuyo autor es Cowie r en el año 2001. El tercer clúster, representado por el color azul se destaca el documento cuyo autor es Chang en el año 2011; finalmente el cuarto clúster en color verde se destaca el documento cuyo autor es Erol en el año 2007. Ahora, al observar la tabla 2, nótese que las cinco primeras referencias con mayor cantidad de citaciones son:Davidfd, 1989 con 49; Pavlovic vi, 1997 con 41; Zengzh, 2009 con 43; Mitra s, 2007 con 34 y Jaimes a, 2007 con42. De igual manera las cinco primeras referencias con mayor fuerza del enlace son: David fd, 1989, con 189; Pavlovic vi, 1997, con 138; Venkatesh v. 2003 con 130; Zengzh, 2009 con 128 y finalmente Mitra s, 2007 con 126.

3.5. Red de co-citación de autores

El mapeo y análisis de la red de co-citación de autores, permite descifrar aspectos como la afiliación institucional de los investigadores que contribuyen dentro del campo de investigación en HCI, autores seminales dentro del área, la influencia de los autores dentro del campo de investigación y el estado de integración de los autores dentro del mismo. El análisis de co-cita del autor es la subcategoría de estudio que mapea las obras y por implicación, las personas que los producen. Los datos brutos son recuentos de los tiempos en que los pares de autores seleccionados se citan juntos en artículos, independientemente de cuáles de sus obras se citan.

El análisis de la co-citación de autores se centra en las interrelaciones entre autores individuales en un campo de investigación. Al medir las frecuencias de co-ocurrencia

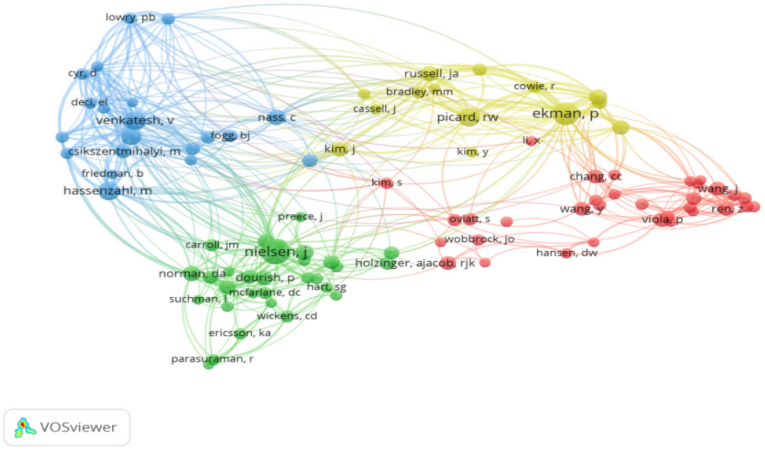


Figura 7 – Gráfico de Red de co-citación de autores Fuente: VOSVIEWER

de trabajos individuales de diferentes autores en bibliografías, se pueden identificar las interconexiones entre los autores(Liu et al., 2015). Cuanto más se citan dos autores, más se relacionan intelectualmente. El enfoque de la co-cita del autor, está guiado por la premisa de que las citas entre documentos de dos científicos representan un indicador igualmente adecuado para la proximidad en términos de contenido dentro de una escuela de investigación. VOS VIEWER se emplea aquí para generar la visualización de las redes de co-citación de autores y para representar ‘’redes institucionales invisibles” entre los investigadores de HCI.

La Figura 7 traza la red de co-citación de los autores más citados. Consta de los 105 colaboradores más citados y 191 enlaces de co-citación. El tamaño de cada nodo corresponde al número total de citas de los autores relevantes. Los nodos de mayor tamaño, indica centralidad y significa que este autor tiende a ser un erudito fundamental, cuyo trabajo vincula diferentes áreas de desarrollo.Dentro de este conjunto de autores se identifican cuatro clústeres claramente definidos, que se complementan como subconjuntos o escuelas, para conformar la estructura de red de autores de HCI. El clúster en color verde. Nielsen J se presenta como el autor de mayor

4. Conclusiones

Este trabajo es basado en datos bibliométricos relacionados con 1.970 artículos de revistas listados la base de datos *Web of Science* (WOS) de *Clarivate Analytics*y los métodos de análisis de co-palabra y co-citación del software bibliométrico VOS VIEWER para examinar el desarrollo de la investigación eninteracción humano– computador (HCI). El análisis de las palabras clave permitió explorar el contenido evolutivo de la investigación, el análisis de las co-ocurrencias permitió la identificación de las disciplinas e instituciones fuente de las principales contribuciones, la co-citación de revistas y el análisis de la co-citación de autores representan las redes que conectan a las revistas y a los autores.

Las tres principales revistas en términos de frecuencia de co-citación fueron larevista Lect notes computcsc, resulta ser la revista que tiene mayor número de relaciones con su mismo clúster y con los otros, la segunda revista más fuerte dentro del clúster es la Ieee t pattern anal y en tercer lugar Pro cvprieee.

El análisis de la citación conjunta de documentos escogió tres publicaciones de r David, en 1989; Cowie r en el año 2001; Chang en el año 2011; finalmente Erol en el año 2007. Se identificó la importancia de la publicación conjunta. El análisis de las publicaciones también permitió identificar los puntos de inflexión en los que surgieron nuevas direcciones y las frecuencias de estallido que indicaban la velocidad con la que se adoptaron las nuevas ideas, y observamos que algunas de las publicaciones de punto de inflexión con alta centralidad tienen puntuaciones de citas relativamente bajas, lo que indica la importancia de utilizar múltiples indicadores de impacto.

Finalmente, de los análisis se infiere que HCI está fuertemente relacionado con la usabilidad, el reconocimiento de gestos, el diseño para que esto se haga posible, sin dejar de lado el rendimiento de las aplicaciones, por medio del desarrollo de modelos y sistemas que faciliten la experiencia de los usuarios finales.

Agradecimientos

Los autores desean agradecer la colaboración de todos los socios dentro del proyecto Centro de Excelencia y Apropiación en Internet de las Cosas (CEA-IoT). Los autores también desean agradecer a todas las instituciones que apoyaron este trabajo: el Ministerio de Tecnología de la Información y Comunicaciones – MinTIC de Colombia y al Departamento Administrativo de Ciencia, Tecnología e Innovación – Colciencias, a través del Fondo Nacional deFinanciamiento para la Ciencia, Tecnología y la Innovación Francisco José de Caldas (ID Proyecto: FP44842-502-2015).

Referencias

Acs, Z. J., Anselin, L., & Varga, A. (2002). Patents and innovation counts as measures of regional production of new knowledge. *Research Policy*, *31*(7), 1069–1085. https://doi.org/10.1016/S0048-7333(01)00184-6

Borgatti, S. P., & Halgin, D. S. (2011). On network theory. *Organization Sicence*, *22*(5), 1168–1181.

Carvalho, M.M., Fleury, A., Lopes, A. . (2013). An overview of the literature on technology roadmapping (TRM): contributions and trends. *Technol. Forecast. Soc. Chang*, *80*(7), 1418–1437.

Cheng, H., Yang, L., & Liu, Z. (2015). A Survey on 3D Hand Gesture Recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, *PP*(99), 1. https://doi.org/10.1109/TCSVT.2015.2469551

Chiu, W. T., & Ho, Y. S. (2007). Bibliometric analysis of tsunami research. *Scientometrics*, *73*(1), 3–17. https://doi.org/10.1007/s11192-005-1523-1

Danthi, N., Wu, C. O., Shi, P., & Lauer, M. (2014). Percentile ranking and citation impact of a large cohort of national heart, lung, and blood institute-funded cardiovascular R01 grants. *Circulation Research*, *114*(4), 600–606. https://doi.org/10.1161/CIRCRESAHA.114.302656

Eck, N. J. Van, & Waltman, L. (2013). VOSviewer Manual. *1 January 2013*, (January), 1–28. https://doi.org/10.3402/jac.v8.30072

Fujita, K., Kajikawa, Y., Mori, J., & Sakata, I. (2014). Detecting research fronts using different types of weighted citation networks. *Journal of Engineering and Technology Management - JET-M*, *32*, 129–146. https://doi.org/10.1016/j.jengtecman.2013.07.002

Garcia Martinez, M., & Walton, B. (2014). The wisdom of crowds: The potential of online communities as a tool for data analysis. *Technovation*, *34*(4), 203–214. https://doi.org/10.1016/j.technovation.2014.01.011

Garfield, E. (1972). Bibliometric Analysis as a Tool in Journal Evaluation. *Science*, 471–479. https://doi.org/DOI: 10.1126/science.178.4060.471

González-Ortega, D., Díaz-Pernas, F. J., Martínez-Zarzuela, M., & Antón-Rodríguez, M. (2014). A Kinect-based system for cognitive rehabilitation exercises monitoring. *Computer Methods and Programs in Biomedicine*, *113*(2), 620–631. https://doi.org/10.1016/j.cmpb.2013.10.014

Herath, S., Harandi, M., & Porikli, F. (2017). Going deeper into action recognition: A survey. *Image and Vision Computing*, *60*, 4–21. https://doi.org/10.1016/j.imavis.2017.01.010

Jose M.Merigo, Fabio Blanco-Mesa, Anna M. Gil-Lafuente, R. R. Y. (2016). Thirty Years of the International Journal of Intelligent Systems: A Bibliometric Review. *International Journal of Intelligent Systems*, *00*(1), 1–34. https://doi.org/10.1002/int.21859

Koehler, W. (2001). Information science as "Little Science": The implications of a bibliometric analysis of the Journal of the American Society for Information Science. *Scientometrics*, *51*(1), 117–132. https://doi.org/10.1023/a:1010516712215

Lallemand, C., Gronier, G., & Koenig, V. (2015). User experience: A concept without consensus? Exploring practitioners' perspectives through an international survey. *Computers in Human Behavior*, *43*, 35–48. https://doi.org/10.1016/j.chb.2014.10.048

Liu, Z., Yin, Y., Liu, W., & Dunford, M. (2015). Visualizing the intellectual structure and evolution of innovation systems research: a bibliometric analysis. *Scientometrics*, *103*(1), 135–158. https://doi.org/10.1007/s11192-014-1517-y

Martínez-Torres, R. (2013). Analysis of open innovation communities from the perspective of social network analysis . *Intangible Capital*, *9*(1), 46–64. https://doi.org/10.3926/ic.352

Mengüç, Y., Park, Y. L., Pei, H., Vogt, D., Aubin, P. M., Winchell, E., … Walsh, C. J. (2014). Wearable soft sensing suit for human gait measurement. *International Journal of Robotics Research*, *33*(14), 1748–1764. https://doi.org/10.1177/0278364914543793

Naghizadeh, R., Elahi, S., Manteghi, M., Ghazinoory, S., & Ranga, M. (2015). Through the magnifying glass: an analysis of regional innovation models based on co-word and meta-synthesis methods. *Quality and Quantity*, *49*(6), 2481–2505. https://doi.org/10.1007/s11135-014-0123-7

Narin, F. (1987). Bibliometric techniques in the evaluation of research programs. *Science and Public Policy* , *14*(2), 99–106. https://doi.org/10.1093/spp/14.2.99

Pisharady, P. K., & Saerbeck, M. (2015). Recent methods and databases in vision-based hand gesture recognition: A review. *Computer Vision and Image Understanding*, *141*, 152–165. https://doi.org/10.1016/j.cviu.2015.08.004

Pritchard, A. (1969). Statistical Bibliography Or Bibliometrics? *Journal of Documentation*, *25*((4)), 348–349.

Pudovkin, A. I., & Garfield, E. (2002). Algorithmic procedure for finding semantically related journals. *Journal of the American Society for Information Science and Technology*, *53*(13), 1113–1119. https://doi.org/10.1002/asi.10153

Rautaray, S. S., & Agrawal, A. (2012). Vision based hand gesture recognition for human computer interaction: a survey. *Artificial Intelligence Review*, *43*(1), 1–54. https://doi.org/10.1007/s10462-012-9356-9

Saviotti, P.-P. (2005). On the co-evolution of technologies and institutions. In *Towards Environmental Innovation Systems* (pp. 9–31). INRA, Université Pierre Mendes-France, 38040 Grenoble Cedex 9, France. https://doi.org/10.1007/3-540-27298-4_2

Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action: A survey. *International Journal of Human Computer Studies*, *74*, 14–31. https://doi.org/10.1016/j.ijhcs.2014.09.006

Seufert, M., Egger, S., Slanina, M., Zinner, T., Hossfeld, T., & Tran-gia, P. (2015). A Survey on Quality of Experience of HTTP Adaptive Streaming. *Ieee Communication Surveys & Tutorials*, *17*(1), 469–492. https://doi.org/10.1109/COMST.2014.2360940

Sink, D. S., Tuttle, T. C., & Devries, S. J. (1984). Productivity measurement and evaluation: What is available? *National Productivity Review*, *3*(3), 265–287. https://doi.org/10.1002/npr.4040030305

Small, H. (1973). Co-citation in the Scientific Literature: A New Measure of the Relationship between Two Documents. *Journal of the American Society for Information Science and Technology*, *July-Augus*, 265–269.

Takey, S. M., & Carvalho, M. M. (2016). Fuzzy front end of systemic innovations: A conceptual framework based on a systematic literature review. *Technological Forecasting and Social Change*, *111*, 97–109. https://doi.org/10.1016/j.techfore.2016.06.011

Terry, J. L. (1994). Authorship in College & Research Libraries Revisited: Gender, Institutional Affiliation, Collaboration. *College & Research Libraries*, 377–384.

Tsarouchi, P., Makris, S., & Chryssolouris, G. (2016). Human – robot interaction review and challenges on task planning and programming. *International Journal of Computer Integrated Manufacturing*, *29*(8), 916–931. https://doi.org/10.1080/0951192X.2015.1130251

Turk, M. (2014). Multimodal interaction: A review. *Pattern Recognition Letters*, *36*(1), 189–195. https://doi.org/10.1016/j.patrec.2013.07.003

van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*(2), 523–538. https://doi.org/10.1007/s11192-009-0146-3

White, H. D., & Mccain, K. W. (1998). Visualizing a Discipline : An Author Co-Citation Analysis of Information Science , 1972 – 1995. *Society*, *49*(4), 327–355.

Yeo, H.-S., Lee, B.-G., & Lim, H. (2015). Hand tracking and gesture recognition system for human-computer interaction using low-cost hardware. *Multimedia Tools and Applications*, *74*(8), 2687–2715. https://doi.org/10.1007/s11042-013-1501-1

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 06/09/2018
Aceitação/Acceptance: 18/11/2018

App Móvil Desarrollada con Metodología Ágil para IoT Controlada desde una Red LAN/WAN con Placa de Desarrollo de Hardware Libre (Arduino)

Xavier Villamil[1], Teresa Guarda[1,2,3]

xaviervq@yahoo.com, tguarda@gmail.com

[1] Universidad de las Fuerzas Armadas ESPE, Centro de Postgrados, 170511, Quito, Ecuador.

[2] Universidad Estatal Península de Santa Elena – UPSE, La Libertad, Ecuador.

[3] Algoritmi Centre, Minho University, Guimarães, Portugal

Pages: 379–392

Resumen: La presente investigación está enfocada en el desarrollo de una App móvil en un ambiente amigable como es MIT app inventor, basada en el análisis de una metodología ágil de desarrollo de aplicaciones móviles, todo esto como parte del Internet de las cosas (IoT) que a su vez pueda ser controlada desde una red LAN o WAN interactuando con las placas de desarrollo de hardware libre Arduino Uno y Shield Arduino Ethernet, tomando como espacio físico un Datacenter en el cual se puedan monitorear por medio de sensores la temperatura, humedad, grabación y reproducción de sonidos, además el apagado o encendido de focos (led) y movimiento de motores paso a paso para simular la ejecución de alguna acción.

Palabras-clave: App; Arduino; Shield; Sensor; IoT.

Mobile App Developed with Agile Methodology for IoT Controlled from a LAN / WAN Network with Free Hardware Development Plate (Arduino)

Abstract: This research is focused on the development of a mobile App in a friendly environment such as MIT app inventor, based on the analysis of an agile methodology for the development of mobile applications, all of this as part of the Internet of Things (IOT) that It can be controlled from a LAN or WAN by interacting with the Arduino Uno and Shield Arduino Ethernet hardware development boards, taking as physical space a Datacenter in which temperature, humidity, recording and temperature can be monitored by means of sensors. reproduction of sounds, in addition the turning off or on of lights (led) and movement of motors step by step to simulate the execution of some action.

Keywords*:* APP; Arduino; Shield; Sensor; IOT.

1. Introducción

Dentro del ámbito tecnológico las aplicaciones móviles, así como el uso de teléfonos inteligentes se han vuelto parte de la vida cotidiana, de ahí surge la necesidad de la creación de aplicaciones móviles apoyado en IoT, que es lo que está al orden del día en la actualidad.

Varias formas de construcción existen a la hora de la creación de aplicaciones móviles, pero no facilitan su uso y la metodología para el desarrollo no es sencilla, por ello se pierde el interés y solo las personas que se desenvuelven en la programación en el área de TI son los llamados a realizar ciertas aplicaciones.

La presente investigación se basa en el uso de los siguientes componentes tecnológicos:

- Desarrollo de Software APP Móvil en Mit App Inventor (Inventor, 2018);
- Metodología de desarrollo de software ágil;
- IoT como medio una Red LAN/WAN plasmado en un Datacenter;
- Placa de Desarrollo de Hardware Libre (Arduino), Shield Ethernet, Sensor de Humedad y Temperatura, Modulo de Sonido, LEDs (Relés).

La presente investigación enfoca el desarrollo ágil, rápido y eficiente de una aplicación móvil utilizando la aplicación web Mit App Inventor que según (Villamil, Guarda, & Ninahualpa, 2018) es un proyecto educativo que cambia el esquema rutinario al desarrollar aplicaciones, con este software se pueden desarrollar proyectos muy interesantes y con un potencial muy grande a la hora de cristalizar un proyecto debido a la forma innovadora de crear Apps móviles en un entorno didáctico e intuitivo que despierta el interés en la creación de aplicaciones de cualquier tipo. En este entorno basado en un portal web lo único que debe tener el usuario es una cuenta de Gmail que a su vez no tiene costo alguno, para empezar a generar proyectos, mediante un entorno amigable programando en bloques como rompecabezas lo que hace muy agradable el ambiente a la hora de generar cualquier idea o proyecto que se quiera cristalizar.

Este desarrollo estará basado en una metodología ágil de desarrollo para dichas aplicaciones denominada DR. Móvil publicada por (Gasca Mantilla, 2014) basada en 5 fases muy claras que dan cuenta de todo lo necesario a la hora de desarrollar una APP móvil de forma ágil. Por otro lado, el objetivo de la App móvil será el de monitorear ciertos aspectos en un Datacenter, todo esto plasmado en IoT por medio del uso de una red LAN/WAN interactuando con la placa de desarrollo de hardware libre Arduino Uno, Shield Ethernet, sensor de temperatura y humedad, módulo de sonido (Grabación y reproducción de sonido), motor paso a paso (movimiento o acción de objetos) y un relé para simular el encendido o apagado de focos o leds.

En el ámbito de TI se conocen de forma general metodologías para el desarrollo de software clásico, así como lenguajes de programación que son los tradicionales y no aportan de forma adecuada el interés en el desarrollo o explotación de aplicaciones móviles que aporten a IoT en las Empresas u Organizaciones en sus procesos de acuerdo a la continuidad del negocio a nivel país, con esta investigación se dejará un precedente para que se puedan realizar a futuro nuevos proyectos tecnológicos que aporten a las áreas de TI con un bajo costo de desarrollo e implementación.

2. Materiales y Costos

En la presente investigación se utilizaron productos tecnológicos de hardware y software de última generación, a su vez los costos generales utilizados en el desarrollo dando cuenta del costo del proyecto que se detallan a continuación.

Descripción	Costo
Placa Arduino Uno	$15,00
Shield Arduino Ethernet	$15,00
Cables Dupont de conexiones	$2,00
Led de color azul y Foco de 100W	$1,10
Motor paso a paso de 5V	$12,00
Shield ULN200 para Arduino	$1,00
Módulo Relé para 2 canales para Arduino	$3,00
Sensor de Sonido con grabación ISD1820 para Arduino	$6,00
Sensor de Temperatura y Humedad DHT22 para Arduino	$6,00
TOTAL	**$61,10**

Tabla 1 – Hardware utilizado con sus costos

Descripción	Costo
Arduino IDE 1.8.6	$0,00
Fritzing 0.9.3	$0,00
Mit APP Inventor 2	$0,00
Gantt Project 2.8.9	$0,00
TOTAL	**$0,00**

Tabla 2 – Software utilizado con sus costos

Descripción	Costo
Desarrollo en Arduino y App Inventor según cronograma por 2 meses costo aprox, por hora $8 utilizando 2 horas diarias 5 días a la semana	$640,00
Telefonía	$20,00
Electricidad	$20,00
Material de oficina, impresiones	$50,00
Publicación en PlayStore	$25,00
TOTAL	**$755,00**

Tabla 3 – Costos generales de la investigación

En la Tabla 1. se detalla el presupuesto del hardware utilizado para el proyecto de la presente investigación que como se aprecia es muy bajo, los materiales son útiles para determinar varios aspectos que se pueden monitorear en un Datacenter, cabe señalar que todos los productos detallados y utilizados son nuevos. El software utilizado como se evidencia en la Tabla 2., tienen licencia GNU/GPL alineado a software libre razón por la cual no existen costos en cuanto a licencias cabe señalar que el software descrito que se ha utilizado da cuenta de las últimas versiones provistas por sus fabricantes.

La Tabla 3. Detalla los costos indirectos que se producen como efecto de la investigación del proyecto que no necesariamente refleja el costo real del proyecto, pero si evidencia ciertos costos que son fundamentales a la hora de realizar este tipo de investigaciones desde cero.

3. Metodología Ágil Aplicada al Desarrollo de Apps Móviles

Como parte de la presente investigación se realiza una APP Móvil desarrollada en la Herramienta Web Mit App Inventor denominada Ardumonitor, para ello se analizaron varias metodologías como Delfroid según (Ávila Domenech & Meneses Abad, 2013) misma que está más apegada a una comparativa asociada entre metodologías como XP y SCRUM, así como un método hibrido planteado según (Ignacio Leiva Mundaca, 2015) que explora el funcionamiento de diferentes métodos ágiles para el desarrollo de aplicaciones móviles según las necesidades que la industria de software lo amerite.

Con este precedente y en el contexto del desarrollo realizado se ha tomado como referencia la Metodología de Desarrollo de aplicaciones móviles Dr. Móvil denominada así según (Gasca Mantilla, 2014), dicha metodología basa su estructura en cinco fases muy claras como se muestra en la Figura 1.

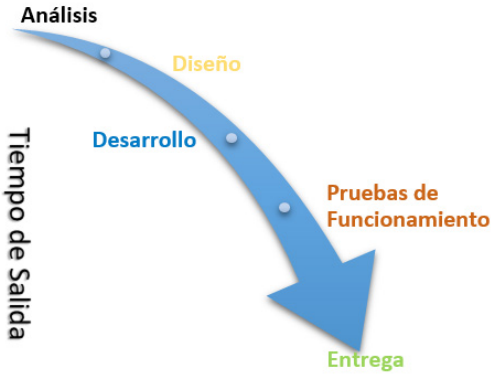


Figura 1 – Fases de la Metodología de desarrollo de aplicaciones móviles Dr. Móvil

De acuerdo a lo anteriormente expuesto se detalla cada fase de la Metodología de desarrollo de aplicaciones móviles en las próximas subsecciones.

3.1. Fase de Análisis

Como parte de la fase de análisis se detallan las siguientes tareas: (1) obtención de requerimientos; (2) clasificación de requerimientos; y (3) personalización del servicio.

1. Obtención de requerimientos: El objeto de estudio en el cual está enfocada la presente investigación es el desarrollo de una APP móvil por medio de una herramienta de software que facilite la creación, además de ser intuitiva y que cumpla con las expectativas del desarrollador, a su vez hoy en día el internet de las cosas (IoT) según (CISCO, 2018) vincula los objetos a Internet, habilitando datos y perspectivas que nunca antes habían estado disponibles. Se crean nuevas oportunidades de ingresos y se disminuyen los costos operativos a través de la administración de los dispositivos conectados, que para este proyecto se ve plasmado con el uso de la APP móvil a través de un Smartphone conectado a una Red LAN/WAN dependiendo del caso para Monitorear varios aspectos esenciales (Temperatura, humedad, sonidos, encendido o apagado de focos, leds) dentro de un Datacenter, con esto se dejará sentado un precedente para realizar futuros proyectos que beneficien al área de TI de las empresas.
2. Clasificación de requerimientos: El ámbito en el que se puede desarrollar la presente investigación es diverso ya que por medio de la utilización de la APP móvil a través de una Red LAN o WAN enfocado en el internet de las cosas (IoT) no clasifica solo el uso de las necesidades de monitoreo de un Datacenter, sino que da cuenta que lo utilizado como herramientas de hardware (Arduino uno, Shield ethernet, sensores, módulos, motores, etc.) amplían el espectro y pueden ser utilizados en innumerables proyectos que pueden ser plasmados a futuro no solo en lo tecnológico sino también en el ámbito del agro o la domótica por poder citar unos ejemplos.
3. Personalización del Servicio: El servicio de continuidad fundamental de la aplicación móvil será el monitoreo de varios aspectos cotidianos que pueden ocurrir en un Datacenter, con esto se puede garantizar la continuidad y disponibilidad de los servicios tan críticos que se administran dentro como son: Blades, UPS, Centrales telefónicas, SAN, NAS, Cableado estructurado, Aire Acondicionado, etc. Por medio del hardware utilizado a través de Arduino y sus complementos (Shield ethernet, sensores, módulos, motores paso a paso, relés) se podrá determinar la temperatura, humedad, grabación de sonidos, encendido o reinicio físico de servidores, apagado o encendido de Focos o LEDS.

3.2. Fase de Diseño

Como parte de la fase de diseño se realizan las siguientes actividades: (1) definir el escenario; (2) estructura del software; (3) definición de tiempos; y (4) asignación de recursos.

1. Definir el escenario: La investigación se desenvuelve en un escenario remoto por medio de una APP móvil que utiliza el protocolo HTTP puerto 80 para este caso instalada en un Smartphone que se conecta al hardware de desarrollo libre Arduino Uno empotrado al Shield Arduino Ethernet que a su vez se conectará por medio de un patchcord (cable de conexión) a la Red corporativa

de una empresa u organización. Con base en lo anterior se derivan dos posibles escenarios principales que son:

1. Conexión a una Red LAN por medio de WIFI
2. Conexión a una Red WAN por medio de datos móviles de las operadoras de telefonía celular que se dispongan en el medio.

El circuito electrónico diseñado para lograr el escenario remoto propuesto se lo aprecia en la Figura 2., donde se puede determinar todos los elementos de interconexión de una forma clara.

2. Estructura del software: El software en su totalidad está desarrollado en Mit APP inventor que ayuda al análisis remoto de la aplicación emulando en cada parte de la construcción los componentes de software utilizados como lo menciona (Riego, 2017) en su sitio Web. Para el desarrollo de la APP móvil se ha utilizado el modelo de implementación que agilita el proceso de construcción de la APP, para ello se detallan los componentes de software utilizados que se los puede apreciar en la Figura 3 y el escenario físico en el que se desenvuelve la App móvil en la Figura 4.

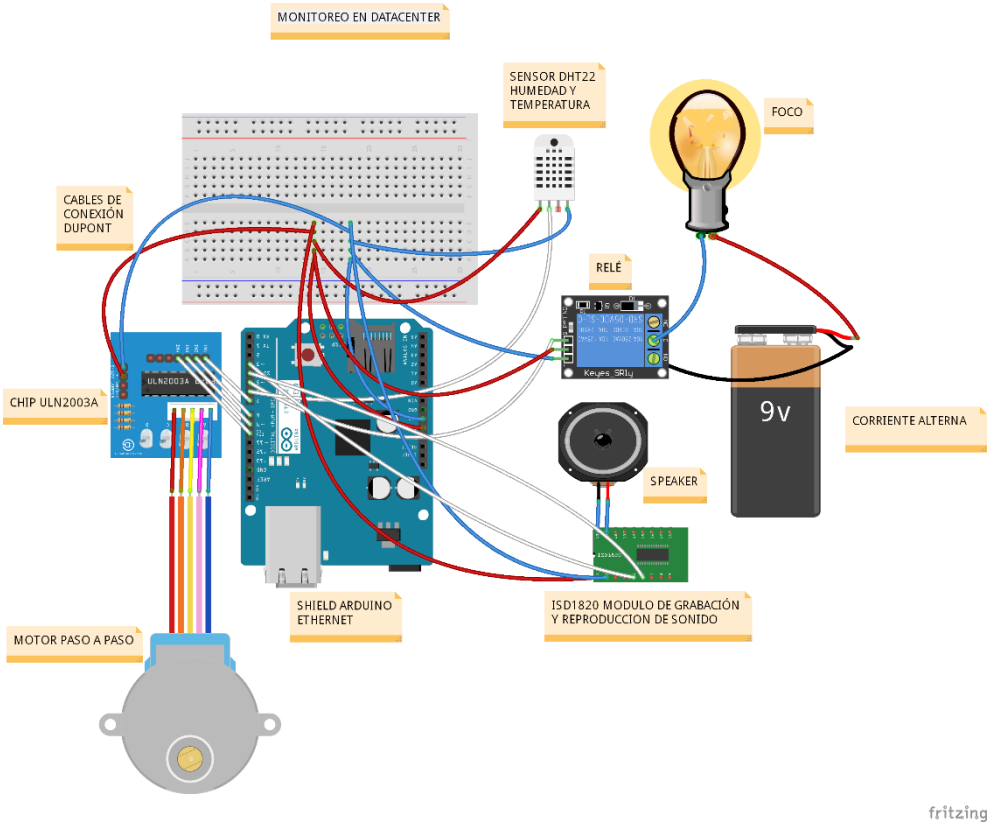


Figura 2 – Circuito electrónico ArduMonitor

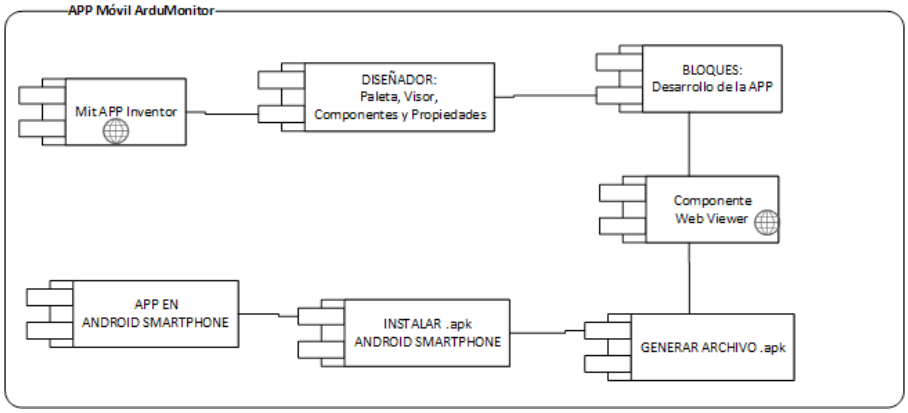


Figura 3 – Diagrama de Componentes

3. Estructura del software: El software en su totalidad está desarrollado en Mit APP inventor que ayuda al análisis remoto de la aplicación emulando en cada parte de la construcción los componentes de software utilizados como lo menciona (Riego, 2017) en su sitio Web. Para el desarrollo de la APP móvil se ha utilizado el modelo de implementación que agilita el proceso de construcción de la APP, para ello se detallan los componentes de software utilizados que se los puede apreciar en la Figura 3 y el escenario físico en el que se desenvuelve la App móvil en la Figura 4.

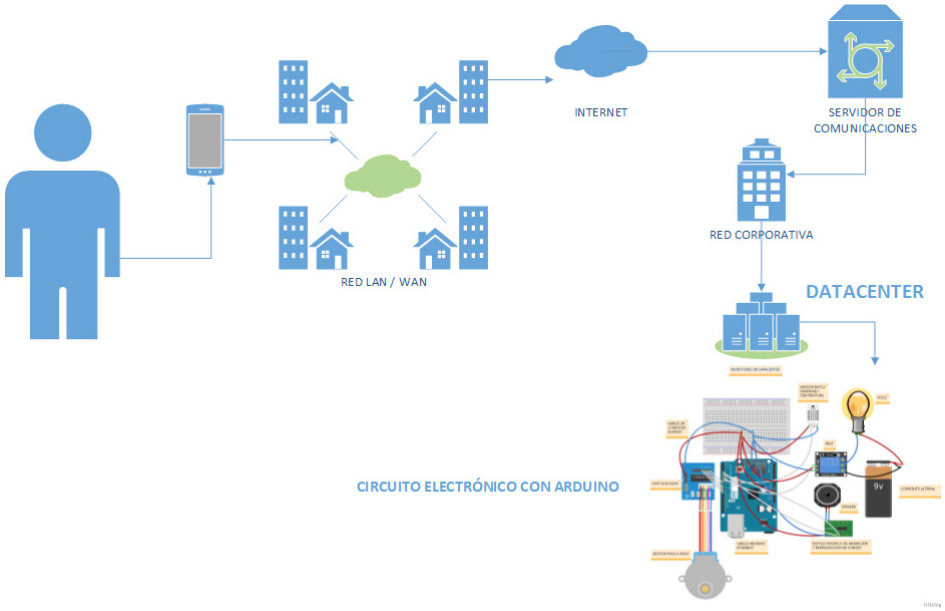
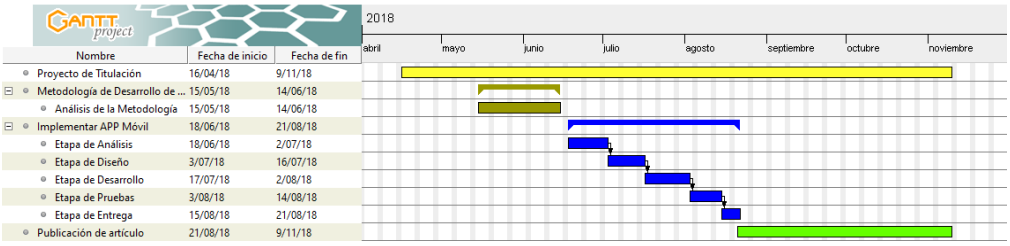


Figura 4 – Diagrama de Despliegue APP Móvil ArduMonitor

4. Definición de tiempos: El tiempo previsto para el desarrollo del presente proyecto de investigación se lo puede apreciar en el Diagrama de Gantt de la Figura 5.



Nombre	Fecha de inicio	Fecha de fin
Proyecto de Titulación	16/04/18	9/11/18
Metodología de Desarrollo de ...	15/05/18	14/06/18
Análisis de la Metodología	15/05/18	14/06/18
Implementar APP Móvil	18/06/18	21/08/18
Etapa de Análisis	18/06/18	2/07/18
Etapa de Diseño	3/07/18	16/07/18
Etapa de Desarrollo	17/07/18	2/08/18
Etapa de Pruebas	3/08/18	14/08/18
Etapa de Entrega	15/08/18	21/08/18
Publicación de artículo	21/08/18	9/11/18

Figura 5 – Diagrama de Gantt ArduMonitor

5. Asignación de recursos: Los recursos utilizados en el desarrollo del presente proyecto de investigación se los puede apreciar en el Capítulo 2. Materiales y Costos del presente artículo.

3.3. Fase de Desarrollo

Como parte de la fase de desarrollo se realizan la codificación y las pruebas unitarias. Para codificar la App Móvil en Mit App Inventor básicamente se tienen dos ambientes: Diseñador y Bloques, de acuerdo a lo expuesto en la Figura 6., se puede apreciar el ambiente Diseñador fundamental para conseguir la interfaz gráfica del usuario (GUI).

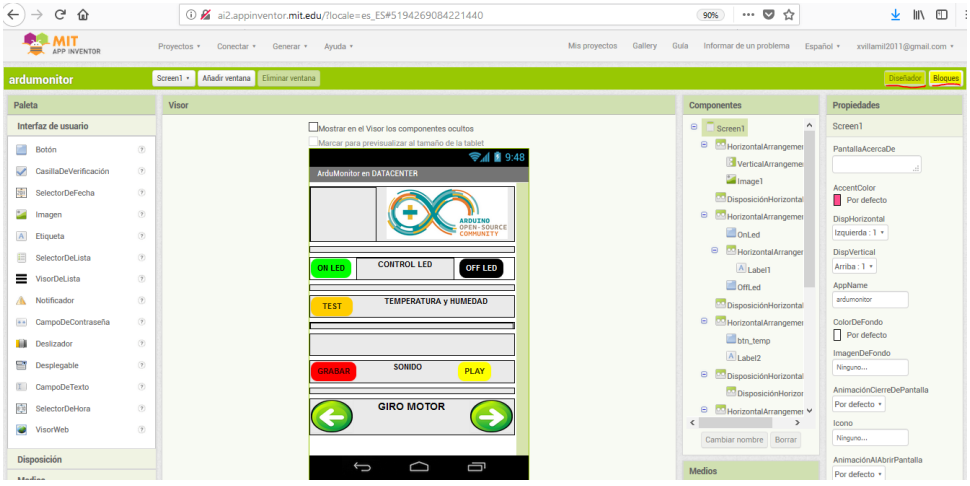


Figura 6 – Ambiente Diseñador APP Ardumonitor

En el ambiente bloques se emula la codificación y las acciones o eventos que se quiere que ejecute la aplicación diseñada para este caso en la Figura 7. Se muestran los bloques utilizados conforme al diseño expuesto.

Una vez realizado el proceso de diseño y codificación por bloques, se tiene lista la App Móvil para las pruebas.

Hay que tener en cuenta que como parte del desarrollo de la APP móvil ArduMonitor, se tiene la codificación realizada en el IDE de Arduino, documentado y se lo puede descargar en la siguiente dirección web:

https://www.dropbox.com/s/97mvs10ef8qd62g/ardueth1.ino?dl=0

Como parte de las pruebas de la APP móvil se realiza la conexión al emulador de Android que tiene como herramienta Mit App Inventor como lo muestra la Figura 8.

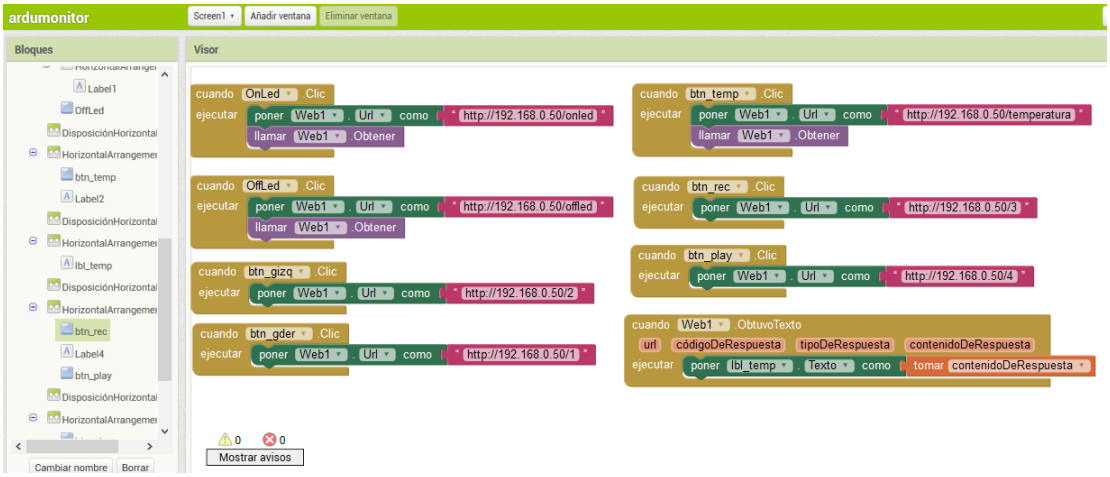


Figura 7 – Ambiente Bloques APP Ardumonitor

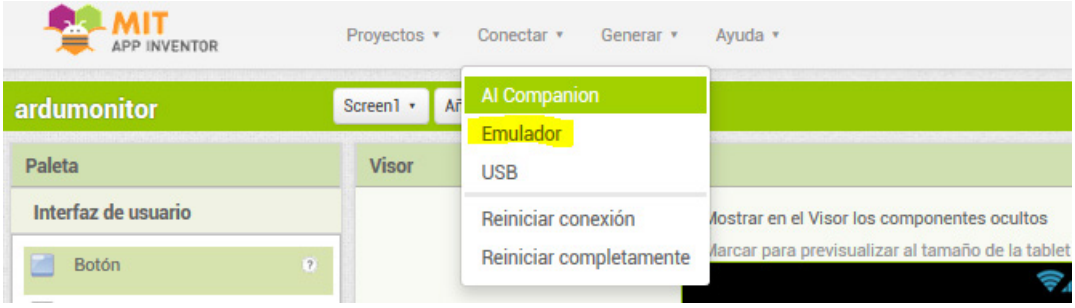


Figura 8 – Conexión al emulador de Android

Como resultado de la emulación se presenta la APP en un ambiente real como lo muestra la Figura 9.

3.4. Fase de Pruebas de Funcionamiento

El funcionamiento de la APP Móvil se la detalla en el Capítulo 4 Resultados del presente proyecto de investigación.

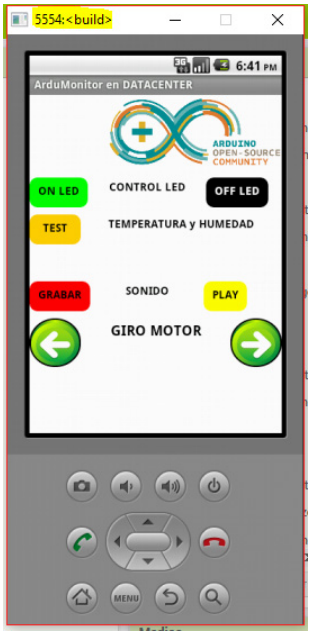


Figura 9 – Emulador de Android APP Ardumonitor

3.5. Fase de Entrega

La descarga de la APP Móvil ArduMonitor estará disponible a partir del 10 de octubre del 2018 en Google Play para su respectiva descarga, para todo público interesado en el presente proyecto que puede hacer uso como herramienta base para ser utilizado en el monitoreo de varios aspectos tecnológicos importantes en un Datacenter. A su vez en cualquier momento la APP Móvil se la puede descargar en la siguiente dirección web:

https://www.dropbox.com/s/gby3wujvq5eix2c/ardumonitor.apk?dl=0

4. Resultados

Como resultado del desarrollo de la APP móvil basada en la Metodología ágil de desarrollo de aplicaciones móviles Dr. Móvil utilizando el diseño y la programación en bloques por medio del portal web Mit App inventor (Inventor, 2018) de la presente investigación es la APP ArduMonitor, la cual se muestra en su fase final en la Figura 10.

4.1. Funcionamiento de la Aplicación

La aplicación funciona por medio de la instalación de la App Móvil en un smartphone con sistema operativo Android el nombre del paquete a instalar es ardumonitor.apk, que se lo puede descargar en el literal 3.5 del presente artículo.

La aplicación funciona mediante la llamada a un componente web utilizado en Mit App Inventor que a su vez se comunica por medio del protocolo http puerto 80 con la

Figura 10 – APP Ardumonitor

dirección IP LAN/WAN en la que esté conectada la placa de desarrollo de hardware libre Arduino Uno empotrado al Shield ethernet, seguida de la dirección IP se envía el parámetro (variable) que recibirá el cliente ethernet de Arduino para ejecutar la acción que le corresponda de acuerdo al monitoreo o acción que se desee realizar, como ejemplo de lo citado se muestra la Figura 11.

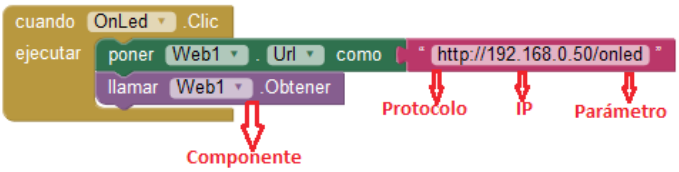


Figura 11 – Componente Web App Móvil

La aplicación se muestra en forma gráfica en 5 secciones como se detalla a continuación:

1. Control Led - Esta sección cuenta con dos botones: ON LED para invocar al sketch en Arduino enviando el parámetro “onled” para encender un foco o led dependiendo el caso y OFF LED parámetro “offled” para apagar el foco o led que se encuentre encendido, esta sección se la puede apreciar en la figura 12.

Figura 12 – Sección 1 Control Led App ArduMonitor

2. Temperatura y Humedad - Esta sección cuenta con un botón denominado TEST que invoca al sketch en Arduino enviando el parámetro “temperatura” para que a su vez devuelva los datos de la Temperatura y la Humedad presentes en ese momento en el Datacenter, esta sección se la puede apreciar en la figura 13.
3. Mensajes - En esta sección se cargan todos los mensajes que muestran los resultados al presionar un botón en la App Móvil ArduMonitor, está sección se la puede apreciar en la figura 13.

Figura 13 – Sección 2: Temperatura y Humedad App ArduMonitor, Sección 3: Mensajes App ArduMonitor

4. Sonido - Esta sección cuenta con dos botones: GRABAR para invocar el sketch en Arduino enviando el parámetro “3” para que empiece a grabar los sonidos que se produzcan en el Datacenter durante 10 segundos y PLAY parámetro “4” para reproducir el sonido grabado, está sección se la puede apreciar en la figura 14.

Figura 14 – Sección 4 Sonidos App ArduMonitor

5. Giro motor - Esta sección cuenta con dos botones: Flecha Izquierda para invocar el sketch en Arduino enviando el parámetro “2” para que gire el motor 60° a la izquierda emulando alguna acción en el Datacenter y Flecha derecha parámetro “1” para que gire el motor 60° a la derecha de igual forma emulando alguna acción, está sección se la puede apreciar en la figura 15.

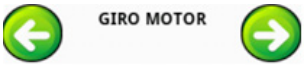


Figura 15 – Sección 5 Sonidos App ArduMonitor

Todas las acciones ejecutadas en los botones presentes en la App Móvil ArduMonitor se las puede apreciar en la codificación del sketch programado en Arduino y que se encuentra disponible para descargar en el literal 3.3 Fase de Desarrollo del presente artículo.

5. Conclusiones y Recomendaciones

Se concluye que el proyecto de software de la presente investigación tuvo cierta complejidad ya que el entorno en la que se desenvuelve la App móvil tiene un enfoque hacia IoT, por esta razón durante el desarrollo no se tenían ciertos conocimientos en la forma de interconectar dicha App vía LAN/WAN para que recepte los datos enviados desde el hardware libre Arduino conectado a sus sensores, módulos, leds y motor paso a paso. Cabe señalar que fue de gran ayuda la Metodología ágil aplicada al desarrollo de App Móviles "DR. Móvil", ya que cada una de sus fases son muy claras y enriquecen la forma en la que se avanza con el proyecto y la documentación es concreta. Se concluye que la APP Móvil "ArduMonitor" interactúa perfectamente utilizando como medio una Red LAN/WAN con la placa de desarrollo de hardware libre Arduino y el Shield ethernet, a su vez todos los factores para monitorear un Datacenter fueron muy prácticos que arrojan datos reales y confiables, ya que en las pruebas se consiguió el fin para lo cual fue creado el proyecto de desarrollo en la presente investigación y se los puede replicar en cualquier proyecto que tenga un fin similar. Una vez analizados todos los factores tecnológicos utilizados en el presente proyecto a través del desarrollo de una App móvil resulta muy motivador, ágil e innovador ya que la forma en la que se genera un proyecto de cualquier tipo sea compleja o no, mediante el entorno de desarrollo web Mit app Inventor (Inventor, 2018) se lo puede lograr consiguiendo un producto de calidad, apoyándose en fuentes de consulta que aporten al proyecto planteado.

Se recomienda aplicar lo expuesto en la presente investigación para que sirva como referente de futuros proyectos relacionados a IOT, que puedan aportar a las áreas de tecnología de las organizaciones o a cualquier persona que guste de los sistemas informáticos o la electrónica a explotar sus ideas y transformarlos en emprendimientos reales en beneficio de la sociedad de la información.

Referencias

Arduino. (2018). *Arduino Introduction*. (Arduino community) Retrieved from https://www.arduino.cc/en/Guide/Introduction

Ávila Domenech, E., & Meneses Abad, A. (2013). Delfdroid y su comparación evaluativa con XP y Scrum mediante el método 4-DAT. *Revista Cubana de Ciencias Informáticas*, 8.

CISCO. (2018). *Internet de las cosas (IoT)*. Retrieved from https://www.cisco.com/c/es_ec/solutions/internet-of-things/overview.html#~stickynav=8

García, P. (2017). *El Profe García Robótica y Electrónica Libre*. Retrieved noviembre 12, 2017, from http://elprofegarcia.com/

Gasca Mantilla, M. &. (2014). Metodología para el desarrollo de aplicaciones móviles. *18*(40), 20–35.

Gonzalez, M. R. (2014). AAprender a Programar 'Apps' como Enriquecimiento Curricular en Alumnado de Alta Capacidad. *Bordón Revista de Pedagogía*, 135–155.

Ignacio Leiva Mundaca, M. V. (2015). Método ágil híbrido para desarrollar software en dispositivos móviles. *Ingeniare. Revista chilena de ingeniería*, 473–488.

Inventor, M. A. (2018). *Mit APP Inventor 2*. Retrieved noviembre 12, 2017, from http://appinventor.mit.edu/explore/library

Riego, A. R. (2017). *Instalar y ejecutar el emulador en AI2* . (APP Inventor en Español) Retrieved noviembre 12, 2017, from https://sites.google.com/site/appinventormegusta/instalacion/instalar-y-ejecutar-el-emulador-en-ai2

RincónIngenieril. (2018). *RinconIngenieril*. Retrieved from https://www.rinconingenieril.es/rele-arduino-ethernet/

Villamil, X., Guarda, T., & Ninahualpa, G. (2018). Agile software development for mobile applications and wireless interaction with hardware development board (Arduino). *2018 13th Iberian Conference on Information Systems and Technologies (CISTI), Caceres*, 1–5.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 07/10/2018
Aceitação/Acceptance: 17/12/2018

Seguridad de la Información en WSN aplicada a Redes de Medición Inteligentes basado en técnicas de criptografía

Liliana Valencia[1], Teresa Guarda[1,2,3,4], Gabriel Patricio Luna Arias[1], Geovanni Ninahualpa Quiña[1]

llvalencia@espe.edu.ec, tguarda@gmail.com, gabriel.luna@itdospuntocero.net, gninahualpa@espe.edu.ec

[1] Universidad de las Fuerzas Armadas ESPE, Centro de Postgrados, 170511, Quito, Ecuador

[2] Universidad Estatal Península de Santa Elena – UPSE, La Libertad, Ecuador

[3] Universidad de las Fuerzas Armadas-ESPE, Unidad Académica Especial ESSUNA, Salinas, Ecuador

[4] Algoritmi Centre, Minho University, Guimarães, Portugal.

Pages: 393–406

Resumen: La seguridad de la información en una WSN (red de sensores inalámbricos) es primordial para garantizar la integridad y disponibilidad de los datos, más aún considerando un sistema de medición inteligente donde la información recolectada es utilizada para monitoreo y facturación de servicios como agua, luz o gas, un ataque a esta red perjudicaría tanto a los usuarios como a las entidades de cobro de los servicios, con cuentas y datos erróneos o no disponibles. Por ello este documento propone la implementación de técnicas de encriptación tanto en topología estrella como MESH, con la finalidad de encontrar la mejor combinación que cumpla con los requerimientos de un Smart Metering System. Adicional se propone la adición de un nodo autenticador que evitará ataques de Blackhole y Wormhole, garantizando con ello la integridad y disponibilidad de los datos transmitidos al Gateway.

Palabras-clave: Seguridad, Encriptación, Topología, WSN.

WSN Security applied to Smart Metering Systems based on cryptography techniques

Abstract: The information security in WSN (Wireless Sensor Network) is essential to guarantee the integrity and availability of data, even more considering a smart measurement system where the information collected is used for monitoring and billing of services such as water, electricity or gas, an attack over this network would harm both the users and the official entities of these services with erroneous or unavailable accounts and data. Therefore, this document proposes the implementation of encryption techniques in nodes of WSN, applying in a star and a Mesh topology, in order to find the best combination that meets the Smart

Metering System requirements. Additionally, this paper proposes the addition of an authenticator node, which will prevent Blackhole and Wormhole attacks, guaranteeing the integrity and availability of the data transmitted to the Gateway of these systems.

***Keywords**:* Security, Encryption, Topology, WSN.

1. Introducción

El desarrollo de las redes de sensores inalámbricos WSN (Wireless Sensor Networks) han tomado un papel primordial en la evolución tecnológica debido a su relevancia dentro del conocido Internet de las Cosas (IOT) (Sánchez, Santidrián & Martínez, 2015), (Gu, Chen, Martínez & Zhuang, 2010), (Guarda, Bustos, Torres, & Villao, 2018). Dentro de una WSN se utiliza nodos los cuales son desplegados en cualquier ambiente para realizar un monitoreo en campo, como es el caso de los sistemas de medición inteligente (Smart Metering System) en los cuales a través de sensores monitorea el consumo de luz, gas o agua en tiempo real, manejando interfaces amigables con gráficos y datos procesados, brindando confort tanto al usuario como a las empresas dispensadoras de los servicios, incluso esta información es utilizada para la facturación de servicios (Geetha & Jamuna, 2013). Como se observa en la figura 1, cada uno de estos nodos tiene un módulo de batería y una pequeña y limitada unidad de procesamiento, la misma que permite la auto-organización y toma de decisiones en cuanto a la data transmitida dentro de la red (Sharma & Ghose, 2010), adicional a ello cuenta con una unidad de comunicación la cual permite la transferencia de información y un módulo de sensamiento que proporciona la data requerida para monitorear (Salinas & Duchén, 2010).

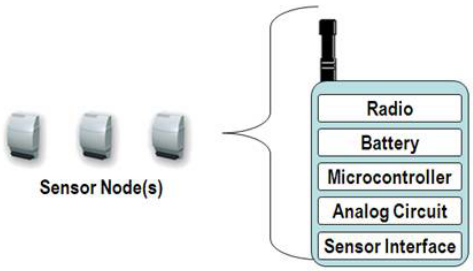


Figura 1 – Partes de un nodo - Sensor inalámbrico. Fuente: (NI, 2009)

Cada uno de estos nodos toman un rol dentro de la red, el nodo final (End-Device), con bajos recursos de procesamiento, toma la información sensada y la envía a un nodo enrutador o nodo Gateway (García, Manotas, Acosta & Romero,2014) el mismo que procesa la información la almacena o la envía a otro nodo enrutador, estos roles depende de la aplicación y topología de la red implementada, incluso existen sensores como es el caso de Waspmote - Libelium los cuales permiten tomar dos roles a la vez como Nodo final y nodo enrutador adaptables a la topología de la red (Libelium, 2016). La principal aplicación de una WSN es realizar el monitoreo de un lugar o de un evento con la finalidad de entregar información útil y verídica a la sociedad. En base a ello se

pueden destacar los siguientes ámbitos de uso (Sharma & Ghose, 2010): aplicaciones militares (utilizados principalmente con sensores de detección de armas y guía para ataques); aplicaciones médicas (sensores que miden actividad vital del cuerpo humano, actualmente es conocida como WBAN (Wireless Body Area Network); monitoreo de desastres (uso de sensores como acelerómetro, detección de gases, temperatura, presión se puede realizar el estudio o detección temprana de desastres); aplicaciones home (en esencia el manejo y comunicación de dispositivos electrónicos a través de WSN, dando lugar a una interfaz entre el usuario y la tecnología en otras palabras IOT).

Dentro de una aplicación de Smart Metering System se debe considerar que el tiempo de respuesta de la red sea lo bastante bajo considerando que la data presentada es en tiempo real, es por ello que sus aplicaciones de procesamiento no deben ser demasiado exigentes. Adicional la información que debe proporcionar esta red de sensores debe contar con los siguientes aspectos basados en los objetivos de seguridad (CIA): integridad (la información alterada en estos casos llegará a perjudicar tanto a los usuarios, al mostrar una información errónea, o en el caso de facturación con valores distintos al real, lo que podría provocar que la empresa que brinda el servicio tenga pérdidas en las ganancias, o incluso que la misma pierda credibilidad de parte de sus usuarios rechazando el sistema impuesto); disponibilidad (si la información no se encuentra disponible para poder observar el consumo o facturación, los usuarios levantarán quejas frente a las empresas dispensadoras de este servicio). La confidencialidad, en estos sistemas, no es considerada como un aspecto relevante de la seguridad debido a que la información proporcionada por los nodos sensores no es sensible ni privada, por ello incluso la desencriptación es realizada con una clave pública (Castelluccia, Mykletun & Tsudik, 2005), detallado en una sección posterior.

El objetivo de este artículo es plantear una solución que garantice la seguridad de información compartida en una red de sensores inalámbricos aplicada a sistemas de medición inteligentes, haciendo uso de técnicas de encriptación y hash sin afectar el desempeño requerido de la red en cuanto al tiempo de entrega de información. Se planteará escenarios con topología estrella y malla, y los aspectos de seguridad basados en HASH y encriptación asimétrica, recomendados debido a la limitación computacional de los nodos dentro de una WSN. Adicional se propone la inclusión de un nodo autenticador el mismo que verifique nodos maliciosos dentro de la WSN.

2. Trabajos Relacionados

En nuestro planeta donde se ha hecho indispensable el uso de energía, es necesario contar con un artefacto conocido como medidor, el cual nos muestra el consumo de energía en nuestros hogares actualmente estos medidores están evolucionando con la tecnología dando lugar a los Smart Metering System como se menciona en (Geetha & Jamuna, 2013), en los cuales permiten mejorar el factor potencia del consumo de energía, estos sistemas brindan predicciones de uso que inclusive puede llegar a modelar comportamientos, con la finalidad de optimizar el consumo, adicional los usuarios son beneficiados ya que disminuyen en un 15% el consumo de energía como se indica en (Chinnow, Bsufka, Schmidt & Bye, 2011). En (Libano, Paranhos, Melchiors & Mano, 2006), (Corral, Coronado, De Castro & Ludwig, 2012) se plantea el desarrollo de sistemas de medición inteligente, en los cuales se destaca como ventajas de estos

sistemas, el mayor control y optimización de la facturación y la reducción de número de empleados, adicional en (Bennett & Wicker, 2010), (Chinnow, Bsufka, Schmidt & Bye, 2011) muestran la aplicación de seguridad en estas redes considerando protocolos de enrutamiento. En cuanto a WSN, se han planteado varias soluciones de encriptación como los siguientes trabajos (Valencia & Quiroga, 2014) , (Kaschel & Chinchay, 2014) , (Cakiroglu & Cetin, 2010) , (Castelluccia, Mykletun & Tsudik, 2005), en los cuales proponen sobre todo soluciones basadas en encriptación asimétrica haciendo uso de llaves tanto públicas como privadas, adicional recomiendan el uso de la técnica ECC(Elliptic curve cryptography) debido al tamaño corto de sus claves, adecuados a la limitada capacidad de los nodos en una WSN a diferencia de técnicas como AES (Advanced Encryption Standard) o incluso RSA.

Se ha realizado incluso variación de parámetros dentro de estas técnicas de encriptación como en (Contreras & Mayol, 2016), donde se realiza la aplicación de ECDSA (Elliptic Curve Digital Signature Algorithm), el mismo que combina RSA junto con curvas elípticas, en el artículo se concluye que su uso es recomendado debido a sus claves cortas, pero que su principal limitación se da en el uso de la memoria. Considerando parámetros de desempeño en la criptografía como uso de memoria, tiempo de encriptación, nivel de seguridad planteado, se han realizado trabajos comparativos de diferentes técnicas de criptografía aplicada a WSN como en (Passing, 2007) , en el cual aplica las técnicas de criptografía MD5, SHA-1, AES a sensores físicos BTNode desarrollados por ETH Zurich, en los cuales se muestra el desempeño en tiempo de cada una de las técnicas, y planteando un modelo de calibración para distintos sistemas de seguridad en WSN. En (Salinas, Duchén, 2010) se propone seguridad para la autenticidad basado en un método de criptografía respecto a la identidad (IBS-Identity Based Cryptography), el mismo que es una variante de la encriptación asimétrica propuesta por Shamir, en esta técnica ya no es requerida la verificación de la autenticidad de las llaves pública ya que existe una autoridad de confianza (TA - Trust Autority) que se encarga de proporcionar llaves públicas basadas en algún identificador del poseedor, en base a estas llaves se da la llave privada, como también es el caso de una encriptación asimétrica común.

Los trabajos mencionados han planteado soluciones para redes de sensores inalámbricos en general o propuestas en base a los protocolos de enrutamiento de los sistemas de medición inteligente, Por ello en este trabajo se procederá a plantear una solución de seguridad frente a una WSN utilizando técnicas de encriptación.

3. Ataques y Técnicas en WSN

Dentro de los trabajos relacionados se han planteado diversos tipos de ataques en WSN y métodos de encriptación, para el caso de sistemas de medición inteligente se ha destacado los puntos más relevantes en base al requerimiento principal de integridad y disponibilidad de la información.

3.1. Ataques Comunes WSN

DoS - Denied of Service, es el envío de información innecesaria a cada uno de los nodos, disminuyendo la posibilidad de proveer el servicio adecuadamente en incluso destruyendo la conectividad en la red. En un WSN aplicado a Smart Metering System el

tiempo de muestreo de cada paquete es primordial considerando que su funcionalidad trabaja en tiempo real, debido a ello un ataque de negación de servicio incrementaría este tiempo y afectaría directamente a la disponibilidad de la data recolectada.

Ataques en Transmisión de Información, considerando que las redes de sensores inalámbricos son susceptibles a ser atacadas por Eavesdropping o husmeo de la información podría el atacante tomar la información alterar la misma y modificarla, si la capacidad computacional del atacante es alta podría incluso llegar a atacar a la información transmitida de varios nodos [21], en el caso de los sistemas de medición inteligentes esta alteración de información podría provocar una facturación errónea en el consumo de los usuarios, por lo tanto este ataque afectaría directamente a la integridad de la data recolectada.

Ataque Wormhole, el atacante logra incluir o convertir a un nodo en malicioso, el cual al requerir actualizar vecindades o enviar paquetes provoca que el nodo receptor tenga otra dirección de nodo transmisor (Fig. 2), el mismo que no es alcanzable a nivel de enrutamiento provocando que los mensajes no lleguen y provoque perdida de comunicación en la red (Sharma & Ghose, 2010) (Pathan, Lee & Hong, 2006). Este ataque afecta a la disponibilidad de la información, y a la integridad de enrutamiento de la red.

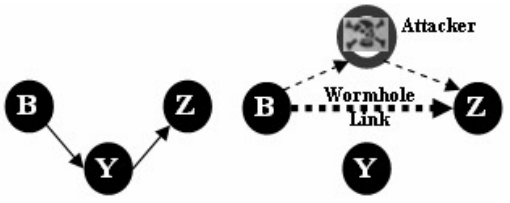


Figura 2 – Ataque Wormhole dentro de una WSN. Fuente: (Pathan, Lee & Hong, 2006)

Ataque Blackhole, en este tipo de ataque el nodo malicioso abarca todo el tráfico de la red de sensores sin dejar que los mismos lleguen al Gateway (Fig. 3), por lo tanto, la data la data recolectada no puede ser transmitida al usuario. Dentro de aplicaciones respecto a sistemas de medición se relaciona directamente con la disponibilidad de la información e incluso si el atacante desea podría afectar la integridad de la data recolectada y enviarla al Gateway el mismo que transmitiría una falsa información.

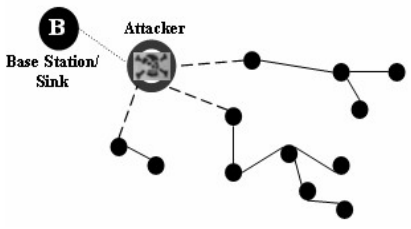


Figura 3 – Ataque Blackhole dentro de una WSN. Fuente: (Pathan, Lee & Hong, 2006)

Dentro de una red de sensores inalámbricos por el mismo hecho que se encuentran desplegados en ambientes públicos, cualquier persona - atacante puede acceder a la red, no solo provocando un ataque sino realizando la combinación de varios de ellos, puede aplicar un Blackhole y al mismo tiempo incluir un ataque de denegación de servicio. En una red utilizada que incluso es utilizada para facturación como es el caso de los sistemas de medición inteligente, esto puede provocar grandes pérdidas financieras y de credibilidad a las empresas y a los usuarios desconcierto y desconfianza respecto al sistema que están utilizando para sus servicios.

3.2. Técnicas Aplicadas en WSN

La encriptación ECC es un método de encriptación asimétrico, el cual hace uso de llaves públicas y privadas. Utiliza el logaritmo discreto de curva elíptica conocido como ECDLP, el cual proporciona el mismo nivel de seguridad que el método RSA, también encriptación asimétrica, con la diferencia que ECC cuenta con longitud de claves más pequeñas (Passing, 2007). ECC determina sus operaciones matemáticas sobre la curva elíptica haciendo uso de la ecuación 1. Donde cada valor de *a* y *b* representan una curva elíptica distinta. Los puntos *(x, y)* dentro de la ecuación se encuentran dentro de la curva elíptica planteada (Passing, 2007). ECC hace uso del algoritmo ECDSA (Contreras & Mayol, 2016), para este algoritmo es necesario obtener la llave privada y la pública $y^2 = x^3 + ax + b$. La llave privada *d* es un valor aleatorio obtenido del intervalo *[1, n-1]* donde n es el orden del punto base *P (a, b)*. La clave pública *Q* es igual a la clave privada multiplicada por el punto generador $Q = dG$. Una vez obtenidas las llaves se procede a la firma digital, haciendo relación de las llaves *d,Q*. En primer lugar, se determina la relación *kP= (a1, b1)*. El primer parámetro a calcular es r, en que $r = a1 \ mod \ (n)$. El segundo parámetro de la firma digital es calculado en base a la ecuación 4, donde *H(m)* corresponde al Hash del mensaje a firmar m, el algoritmo que se debe hacer uso es SHA-1, $s = k^{-1}[H(m) + d.r] mod \ n$.

HASH MD5/SHA-1, como se observa en la figura 4 Hash es un código hexadecimal de longitud fija dependiendo de la técnica de HASH utilizada, con la finalidad de corroborar la integridad de la información recibida (Kumar, 2016). MD5 trabaja con 128 bits de salida mientras que SHA-1 cuenta con 160 bits de salida, SHA-1 es lógicamente más robusto que MD5 en cuanto a colisiones, pero ello tiene un costo de tiempo y procesamiento, por lo cual convierte a MD5 en un HASH rápido, lo cual cumple con lo requerido en cuanto a aplicaciones en tiempo real

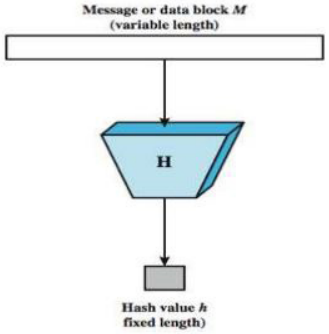


Figura 4 – Funcionamiento HASH. Fuente: (Kumar, 2016)

4. Modelo de Criptografía Propuesto

Para el modelo planteado se considerará el uso de nodos programables, como es el caso de los sensores Waspmote - Libelium o nodos construidos en base a sistemas Arduino u otros relacionados Kaschel, H. Chinchay, I. (2014) con módulos de transmisión ZigBee, se propone este requerimiento ya que como parte del planteamiento de la solución se requiere tener acceso a la trama que será enviada por estos nodos sensores (Trama API). Para el modelo planteado se detallará en primer lugar las topologías a utilizar en los modelos. Luego de ello se procederá a plantear soluciones en base a la literatura y requerimientos de tiempo real.

4.1. Topologías WSN

Para el despliegue de una red de sensores inalámbricos se plantea dos escenarios: Topología Estrella y Tipo Mesh. En la topología Estrella todos los nodos enviarán la información a un nodo Gateway, es decir los roles dentro de esta topología serán o nodos finales o nodo Gateway conectado a un servidor o computador con la finalidad de procesar la información recolectada. Como se observa en la figura 5 la información es transmitida en una sola dirección. Es decir, los nodos finales solamente actúan como transmisores de la información sensada. Los ataques más destructivos para este tipo de red serian Blackhole y Wormhole, debido a que todos los nodos enviarán su información hacia un solo nodo que podría ser un nodo malicioso.

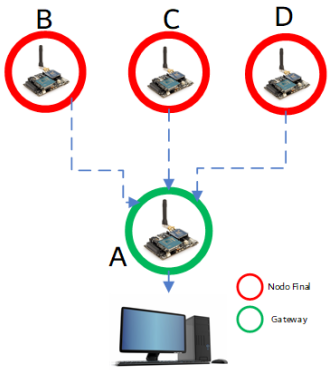


Figura 5 – Distribución de nodos Topología Estrella.

La Topología Mesh se caracteriza por contar tres tipos de nodos en la red, nodos finales, enrutadores y el Gateway (Fig. 6), los nodos enrutadores tienen un nivel de procesamiento limitado y pueden recibir la información y transmitir la misma con capacidad de auto organización, en este tipo de topología se puede prevenir los Wormhole debido a que su enrutamiento permite tener redundancia de caminos para poder llegar a su destino utilizando a sus vecinos. Este tipo de topología a pesar de contar con una estructura más robusta en cuanto a seguridad y paquetes perdidos, puede llegar a incrementar el tiempo requerido para la transmisión de los datos, aspecto negativo en sistemas tiempo real, por ello es necesario buscar técnicas de encriptación con bajo costo computacional.

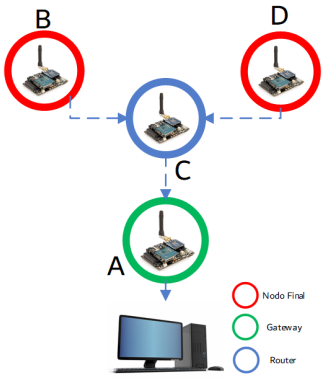


Figura 6 – Distribución de nodos Topología Mesh.

4.2. Técnica de criptografía propuesta

La criptografía en topología tipo estrella trabaja con nodos de bajos recursos en procesamiento, para garantizar la integridad y disponibilidad de la información se plantea el uso de los siguientes procedimientos, considerar que no se enfatiza en la confidencialidad, debido a que los datos sensados no son datos sensibles a la comunidad: Función HASH de información (cada uno de los nodos finales incluirá en su trama de transmisión el valor HASH de la data a ser enviada (Kaschel & Chinchay, 2014), el Gateway al recibir la información sensada corroborará que la información se encuentre acorde al valor HASH, caso contrario descartará el paquete y se encontrará a la espera de una nueva muestra como se muestra en la figura 7 es decir no procesará la información basura. Considerar que cada nodo tiene la capacidad de configurar su periodo de muestreo. La función HASH que se propone aplicar es MD5 el cual devuelve un código de 32 caracteres hexadecimales, se aplica este método considerando que su tiempo de ejecución es menor frente a SHA-1 como se muestra en el artículo (Passing, 2007) con tiempos menores a 100ms incluyendo el tiempo de demora de transmisión del nodo final al Gateway conveniente para los sistemas de medición inteligente. Con este método se busca que el Gateway pueda garantizar la integridad de la información transmitida por cada uno de los nodos, evitando el ataque en paquetes enviados); Encriptación ECC (para la encriptación de la data enviada se propone utilizar ECC considerando que debido a sus claves cortas su costo en red no es alto, como se menciona en (Chelli, 2015) (Kaschel & Chinchay, 2014), se trabaja con este método considerando que la capacidad de procesamiento es bastante limitada en los nodos finales. Para el manejo de las llaves tanto públicas como privadas, se recomendaría su distribución por parte de entidades especializadas en el caso de Ecuador, el Ministerio de Obras Públicas, el mismo que será considerada como Autoridad de confianza (TA) (Fig.8). Cada nodo final encriptará su información con la llave privada única por cada nodo, de igual manera su objetivo es complementar la integridad de la información sin dejar que ningún atacante externo a la red pueda incluir información no acorde a lo sensado); Nodo Autenticador (se propone incluir un nodo que solicite a todos los nodos que se encuentren en la red, nodos finales y Gateway, información de su dirección IP + ID. Para no interferir con la memoria utilizada del nodo, se procederá a obtener la función HASH en conjunto de los

datos enviados. Esta información será solicitada regularmente, si el hash de un segundo requerimiento no es acorde con el hash anterior, el nodo autenticador procederá a notificar a la autoridad de confianza, la misma que procederá a verificar la configuración y nuevos nodos que se encuentren en la red (Fig. 9). Con esta solución se propone evitar la aparición de nodos maliciosos que busquen afectar la disponibilidad de la red).

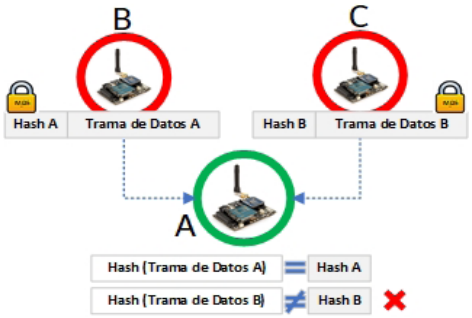


Figura 7 – Aplicación HASH en información transmitida por nodos finales.

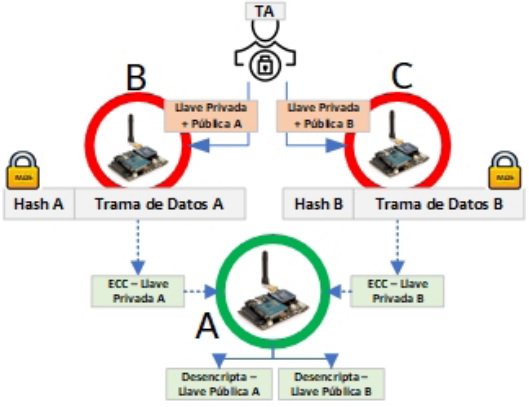


Figura 8 – Encriptación Asimétrica ECC en información transmitida por nodos finales.

4.2.1. Criptografía en topología tipo Mesh.

Como se mencionó en secciones anteriores la topología Mesh cuenta con nodos con capacidad de procesamiento mayor a los nodos finales los cuales se los conoce como nodos Router, en base a ello se procederá a plantear metodologías de encriptación que mantengan el equilibrio entre la seguridad y los requerimientos de una red de sensores inalámbrica para sistemas en tiempo real, como es el caso de los de medición inteligente.

Árbol con función HASH: Se considera el principio del Árbol de Merkle (Montiel, Hernández, Lizama & Simancas, 2017), en este caso las hojas serán los nodos finales los cuales enviarán HASH de su información sensada a los nodos enrutadores, los mismos

que deben comprobar que la información recibida sea igual a la información de HASH, una vez verificada la información tomará su propia información sensada y la concatenará con la información proporcionada por el primer nodo, aplicará HASH y lo enviará al siguiente nodo en su lista de enrutamiento, este proceso se realizará hasta llegar a la raíz o nodo Gateway de la red. Este proceso difiere del aplicado originalmente en el árbol de Merkle debido a que no se realiza HASH sobre la concatenación de HASH, sino sobre la concatenación de información transmitida. Se considera utilizar la metodología MD5, debido a que mientras incremente la información concatenada de cada nodo, más será el tiempo de ejecución de HASH como se muestra en (Passing, 2007). Los otros métodos como por ejemplo SHA-1 muestran incrementos mayores de tiempo a mayor información. De igual manera al aplicar este control HASH, garantizamos que al nodo Gateway le llegue la información verdadera que fue sensada como se observa en la figura 10.

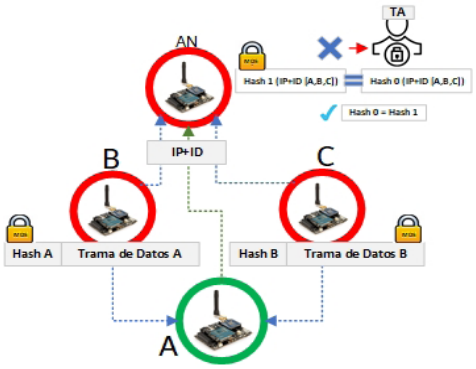


Figura 9 – Funcionalidad Nodo Autenticador.

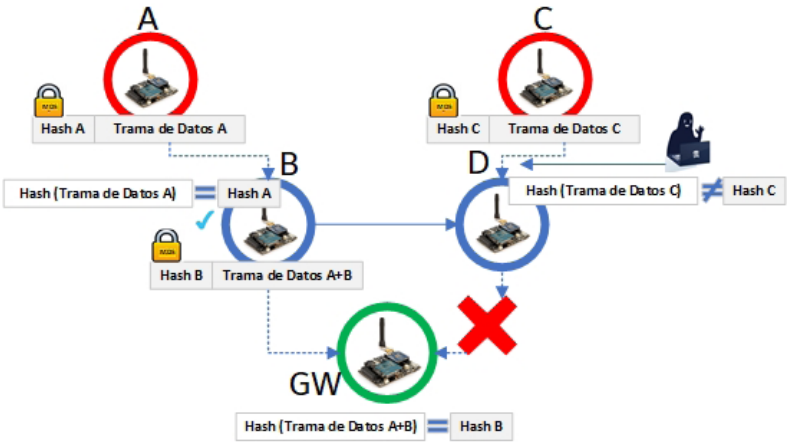


Figura 10 – Árbol de HASH aplicado en topología tipo MESH.

Encriptación: En este artículo se planteará nuevamente la encriptación asimétrica ECC. A pesar que los nodos enrutadores cuentan con mayor capacidad de procesamiento, técnicas como encriptación simétrica, requieren mayor costo computacional (Contreras & Mayol, 2016), como consecuencia de ello se incrementaría el tiempo de envío de la información caso que no es recomendable ya que la información que se mostraría a los usuarios no trabajaría a tiempo real.

Nodo Autenticador: Este nodo cumple la misma función del explicado en la sección de la topología tipo estrella, como se observa en la figura 9.

5. Análisis Comparativo

Una vez planteadas las técnicas de criptografía a ser aplicadas en cada una de las topologías se analiza las ventajas y desventajas de cada una de ellas, resumidas en la tabla 1. Tiempo Óptimo**:** en este caso la topología tipo estrella cuenta con menor tiempo de ejecución y transmisión de información considerando que todos los nodos enviarían la información al mismo tiempo. Y el tiempo empleado en la ejecución de HASH sería similar entre todos los nodos, en cambio en topología tipo Mesh, cada nodo incrementaría su tiempo de envío y ejecución de HASH considerando que por cada salto de nodo la información incrementará de tamaño; procesamiento: considerando que en la topología tipo Mesh cada nodo debe realizar HASH de paquetes de información de mayor tamaño, su procesamiento debe ser mayor conforme se acerque a la raíz del árbol o Gateway; seguridad: ambos modelos cumplen con los objetivos de seguridad planteados, integridad y disponibilidad, pero al considerar que en la topología tipo estrella existe mayor probabilidad de pérdida de paquetes, es decir información sensada, la disponibilidad de la misma se verá afectada en un grado mínimo, por ello se concluye que la topología tipo MESH es más segura que la topología Estrella; riesgo DoS: ambas topologías están expuestas a DoS, debido a que todos los nodos se encuentran en ambientes de libre acceso, por ello el atacante puede infectar a algún nodo y en el caso de topología tipo estrella no dejarle llegar al nodo Gateway, y en el caso de topología tipo Mesh cambiar sus listas de enrutamiento; integridad: al aplicar Criptografía HASH y el nodo autenticador, se garantiza que la información entregada sea la información sensada tanto en topología tipo estrella como en tipo Mesh. Por ello incluso se busca evitar los ataques Blackhole y Wormhole con la implementación de un nodo dedicado a la autenticación de la red de sensores inalámbricos; y capacidad de nodos: en una topología tipo estrella la capacidad de nodos es limitada considerando las colisiones producidas en el Gateway es por ello que la topología tipo Mesh es escalable en cuanto a despliegue de sensores o nodos.

Características	**Topología Estrella**	**Topología Mesh**
Tiempo óptimo	SI	NO
Mayor Procesamiento	NO	SI
Mayor Seguridad	NO	SI
Riesgo DoS	SI	SI
Integridad	SI	SI
Mayor Capacidad de Nodos	NO	SI

Tabla 1 – Comparativo de topologías y técnicas de criptografía propuestas.

6. Discusión y Conclusiones

Los principales ataques en una red de sensores inalámbricos (WSN) afectan principalmente a la integridad y disponibilidad de la información recolectada como la denegación de servicio, los Blackhole, Wormhole, ataque a la información que está siendo transmitida, la confidencialidad no es un objetivo de seguridad relevante debido a que la información captada por los sensores no suele ser sensible ni secreta a las personas. Una red de sensores inalámbricos aplicado a sistemas de medición inteligente puede estar lo suficientemente robusta en cuanto a automatización y calibración de los datos que son recolectados, pero si la red no cuenta con seguridad suficiente, puede verse afectado tanto el usuario como la empresa que presta el servicio, debido a ataques en la integridad y disponibilidad de la información transmitida, su seguridad va dirigida propiamente a un bienestar financiero.

La seguridad implementada a redes de sensores inalámbricos aplicados en Smart Metering System debe estar en equilibrio y acorde a los requerimientos de tiempo real, en base a ello se ha planteado el uso de técnicas de criptografía como MD5 y ECC para la encriptación con la finalidad que el incremento de procesamiento y por consecuencia el aumento de tiempo de transmisión de la información no sea contraproducentes para estos sistemas. La propuesta del nodo autenticador se da principalmente para evitar que nodos maliciosos se incluyan en la red y cambien listas de enrutamiento o aíslen a nodos de su transmisión de información en otras palabras busca evitar el ataque de Blackhole y Wormhole. Dentro de un trabajo futuro se podría plantear implementar un firewall dentro de este nodo con la finalidad que los ataques se queden en este nodo y no ingresen a la red. En el caso de la topología tipo MESH, se podría dejar de lado las técnicas de encriptación y pensar en aplicar protocolos de firewall en cada uno de los nodos enrutadores buscando un protocolo con requerimiento computacional bajo el mismo que garantice el tiempo de transmisión acorde al tiempo real requerido.

En este artículo se ha tratado la seguridad dentro de la red de sensores inalámbricos pero que pasa luego que la información llega al nodo Gateway sea cual sea la topología, cual es la mejor manera de garantizar la seguridad en la transmisión del Gateway a la nube, se puede pensar en encriptación asimétrica de igual manera o en la aplicación de Blockchain donde los mineros sean instituciones como la propuesta autoridad de confianza, el análisis de factibilidad de estos sistemas de seguridad puede ser un trabajo futuro que garantice la seguridad de todo el flujo de información de una red WSN.

Referencias

Sánchez, J. Santidrián, L. Martínez J. (2015). Solución para garantizar la privacidad en Internet de las Cosas. *El profesional de la información*, vol. 24, no. 1, España.

Gu, J. Chen, C. Martínez J. Zhuang, Y. (2010). Wireless sensor network-based topology structures for the Internet of Things localization. *Chinese Journal of Computers,* vol. 33, no. 9, pp. 1548–1556, 2010, Anhui, China.

Guarda, T., Bustos, S., Torres, W., & Villao, F. (2018). Botnets the Cat-Mouse Hunting. The 2018 International Conference on Digital Science (pp. 408-416). Montenegro: Springer, Cham.

Sharma, K. Ghose, M. (2010). Wireless sensor networks: An overview on its security threats. *In Proceedings of IJCA Special Issue on Mobile Ad-hoc Networks*, DOI: 10.5120/1008-44, 42–45.

Salinas, M. Duchén, G. (2010). Protocolo de autenticación para redes inalámbricas de sensores basado en identidad. *Revista Facultad de Ingeniería Universidad de Antioquia*, vol. 52, pp. 196–205.

García, J. Manotas, R. Acosta, R. Romero, A. (2014), *Revisión del estado del arte de redes Zigbee en WSN, Revista I+D en TIC*, Universidad Simón Bolívar, Vol. 5 Núm. 1, Barranquilla-Colombia.

Libelium (2016), Waspmote Documentation, *Libelium, 2016. [Online]. Available:* http://www.libelium.com/development/waspmote/documentation/

Geetha, A. Jamuna, K. (2013). Smart metering system. *2013 International Conference on Information Communication and Embedded Systems (ICICES),* DOI: 10.1109/ICICES.2013.6508368, IEEE, Chennai, India.

Corral, P. Coronado, B. De Castro A. Ludwig, O. (2012),"Design of Automatic Meter Reading based on Zigbee," *in IEEE Latin America Transactions*, vol. 10, no. 1, pp. 1150-1155, Jan. 2012. DOI: 10.1109/TLA.2012.6142452

Bennett, C. Wicker, B. (2010), "Decreased time delay and security enhancement recommendations for AMI smart meter networks," *2010 Innovative Smart Grid Technologies (ISGT)*, Gothenburg, 2010, DOI: 10.1109/ISGT.2010.5434780.

Chelli, K. (2015), "Security issues in wireless sensor networks: Attacks and countermeasures," *In Proceedings of the World Congress on Engineering 2015*, vol. 1, 2015.

Valencia, C. Quiroga, H. (2014), Seguridad en la capa física y teoría de juegos en redes de sensores inalámbricos: Actualidad y perspectiva, *In Proceedings of XV Congreso Internacional de Telecomunicaciones SENACITEL*, vol. 8, no. 9.

Kaschel, H. Chinchay, I. (2014), Implementación de una curva elíptica aplicada a una WSN limitada en hardware, *In Proceedings of XVI Congreso Latinoamericano de Control Automático*.

Cakiroglu, M. Cetin, O. (2010), Performance evaluation of scalable encryption algorithm for wireless sensor networks, *Scientific Research and Essays,* vol. 5, pp. 856–861.

Castelluccia, C. Mykletun, E. Tsudik, G. (2005), Efficient aggregation of encrypted data in wireless sensor networks, *The Second Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services*, DOI: 10.1109/MOBIQUITOUS.2005.25, pp. 109-117, USA.

Pathan, A. Lee, H. Hong, C. (2006), Security in wireless sensor networks: issues and challenges, *Advanced Communication Technology*, vol. 8.

Passing, M. (2007), Experimental performance evaluation of cryptographic algorithms on sensor nodes, *In Proceedings of Mobile Adhoc and Sensor Systems (MASS)*.

Contreras, J. Mayol, R. (2016), Variación de parámetros de criptografía con curvas elípticas usados en la firma digital de datos sobre una red de sensores inalámbricos, *In Proceedings of Mobile Adhoc and Sensor System (MASS)*.

Kumar, M. (2016), A light weight cryptographic Hash algorithm for wireless sensor network, *International Journal of Engineering Research and General Science,* vol. 4, no. 1.

Montiel, L. Hernández, F. Lizama, L. Simancas, E. (2017), Firma Digital móvil basada en criptografía Hash, *Ciencias Multidisciplinarias Proceedings ECORFAN*.

Chinnow, J. Bsufka, K. Schmidt, A. Bye, R. (2011), A simulation framework for smart meter security evaluation, *2011 IEEE International Conference on Smart Measurements of Future Grids (SMFG) Proceedings*. DOI: 10.1109/SMFG.2011.6125758, Italia.

Libano, F. Paranhos, I. Melchiors, J. Mano, O. (2006), Medidor de energía de potencia en un hardware / software de bajo costo, *IEEE Conference Proceedings of ISIE*, Montreal, Quebec, Canadá, 2006, pp.1-5.

National Instruments (2009), ¿Qué es una red de sensores inalámbricos?, *National Instruments – Notas Técnicas. [Online]. Available:* http://www.ni.com/white-paper/7142/es/

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 06/09/2018
Aceitação/Acceptance: 20/11/2018

Sistema de Recuperación de Información para Apoyar la Revitalización del Nasa Yuwe

Luz Marina Sierra Martínez[1], Carlos Alberto Cobos Lozada[1], Juan Carlos Corrales Muñoz[1], Tulio Rojas Curieux[1], Luis Carlos Gómez Flórez[2]

lsierra@unicauca.edu.co, ccobos@unicauca.edu.co, jcorral@unicauca.edu.co, trojas@unicauca.edu.co, lcgomezf@uis.edu.co

[1] Universidad del Cauca, Calle 5 #4-70, 190003, Popayán, Colombia.

[2] Universidad Industrial de Santander, Cra 27 Calle 9 Ciudad Universitaria, 680002, Bucaramanga, Colombia.

Pages: 407–422

Resumen: En este artículo se presenta un Sistema de Recuperación de Información (SRI) propuesto con el fin de ayudar al pueblo indígena Nasa (Colombia) a interactuar en su lengua, el Nasa Yuwe, lengua que se encuentra en peligro de extinción. El proceso para la construcción del SRI involucró la adaptación de un analizador léxico estándar disponible en Lucene a las características propias del lenguaje y la correspondiente definición de palabras vacías. La evaluación de este sistema se realizó basado en una colección cerrada de prueba construida específicamente para Nasa Yuwe y tomando como medida principal de evaluación la curva de Precisión-Recuerdo. El SRI propuesto mostró resultados prometedores en el procesamiento de los textos, corrigiendo los errores en la tokenización de los textos escritos en Nasa Yuwe y obteniendo mejoras en la precisión de la recuperación de los documentos, que en algunos casos supera el 35% en relación con la línea base. El grupo de investigación espera que el uso del SRI por parte de la comunidad indígena Nasa coadyuve en el proceso de revitalización de su lengua.

Palabras-clave: Analizador Léxico, Comunidad Indígena Nasa, Palabras Vacías del Nasa Yuwe, Sistema de Recuperación de información.

Information Retrieval System for Nasa Yuwe

Abstract: In this paper, an Information Retrieval System (IRS) is proposed with the aim of helping the Nasa indigenous people of Colombia to interact in their language, Nasa Yuwe, currently is in danger of extinction. The process for constructing the IRS involved adapting a standard lexical analyzer available in Lucene to the characteristics pertaining to the Nasa Yuwe language and the corresponding definition of empty words. Evaluations of the system were conducted using a closed test collection built specifically for Nasa Yuwe and the precision-recall curve as the principal evaluation measure. The proposed IRS showed promising results in the processing of texts, correcting the errors in the tokenization of texts written in Nasa Yuwe, and obtaining improvements in the Precision of documents retrieved,

which in some cases exceeds 35% in relation with the base line. The research group expects the use of the IRS by the Nasa indigenous community to contribute to the process of revitalization of their language.

***Keywords**:* Lexical Analysis, Nasa Indigenous Community, Stop Words for Nasa Yuwe, Information Retrieval System.

1. Introducción

La Comunidad Nasa está estimada en 186.178 personas, representando el 13,37% de la población indígena de Colombia (Acción Social; Unión Europea; Cordepaz, 2010). Su lengua, el Nasa Yuwe, es una de las 65 lenguas vernáculas que aún sobreviven en Colombia, y que a pesar de ser hablada por una importante parte de la población Nasa, se evidencia el proceso de debilitamiento en su valoración y uso social (CRIC-Programa de Educación Bilingüe e Intercultural-Comisión General de Lenguas, 2008). Es por esto que las organizaciones indígenas ((Consejo Regional Indígena del Cauca - CRIC) (Consejo Regional Indígena del Cauca - CRIC & Programa de Educación Bilingüe e Intercultural -, 2011)(Organización Nacional Indígena de Colombia - ONIC, 2012)) y las entidades gubernamentales (Ministerio de Educación, el Ministerio de Tecnologías y las Comunicaciones, el Ministerio de la Cultura, el Departamento Administrativo de Ciencia, Tecnología e Innovación – COLCIENCIAS, Universidades Nacionales e Internacionales) han considerado el uso de las Tecnologías de la Información (TI), como una oportunidad estratégica para que su adecuación, apropiación y desarrollo al entorno social y cultural del pueblo Nasa favorezca la revitalización de su lengua.

En este trabajo se consideró adecuado utilizar técnicas computacionales, específicamente los Sistemas de Recuperación de Información (SRI) (Manning, Raghavan, & Shütze, 2009), para favorecer las posibilidades de que la comunidad Nasa interactúe directamente con documentos en Nasa Yuwe, motivando la lectura, escritura y recopilación de documentos escritos en esta lengua sobre los cuales se pueda operar un modelo de recuperación de información que le facilite a sus usuarios la búsqueda de los documentos. Se buscó que en el sistema se haga una consulta en Nasa Yuwe y que los documentos recuperados también escritos en Nasa Yuwe se ordenen de acuerdo con las necesidades del usuario y de las características propias de la lengua.

Con base en el anterior contexto, al inicio de la investigación se planteó la siguiente pregunta de investigación: ¿Cómo adaptar las fases del proceso de recuperación de información a textos escritos en Nasa Yuwe?, y para su desarrollo se abordó la adaptación de un analizador léxico estándar (tokenizer) y la definición de una lista de palabras vacías (stop words removal list) para Nasa Yuwe. En el desarrollo de este trabajo se destacan los siguientes aspectos:

1. El modelamiento y construcción de un SRI que soporta la búsqueda y recuperación de documentos en Nasa Yuwe (Carterette & Voorhees, 2011);
2. La adaptación de las tareas de preprocesamiento (análisis léxico, filtrado, conversion de caracteres, remoción de palabras vacías, entre otras) que son altamente dependientes de las características de la lengua en la que se encuentran escritos los documentos a recuperar, como se aprecia en varios trabajos del estado del arte (Hossein Jadidinejad, Mahmoud, & Dehdari, 2010) (Agosti, Bacchin, Ferro, & Melucci, 2003, pág. 281);

3. El uso del modelo de espacio vectorial para la representación de los documentos, modelo que obtiene muy buenos resultados en diferentes contextos y que es el más utilizado en la actualidad (Ribeiro-Neto & Baeza-Yates, 2011); y
4. El uso de la primera colección de prueba para Nasa Yuwe con el objetivo de evaluar el SRI desarrollado (Sierra Martínez, Cobos Lozada, Corrales Muñoz, & Rojas Curieux, 2015).

A continuación, en la sección 2 se presentan los referentes más importantes para el desarrollo del presente trabajo. Luego en la sección 3 se describen los pasos seguidos para hacer la adaptación de las tareas de Recuperación de Información (RI) al SRI propuesto. Después en la sección 4, se presentan los resultados obtenidos y en la última sección se presentan las conclusiones del desarrollo del trabajo y lo que el grupo de investigación espera desarrollar en el futuro cercano.

2. Estado del arte

Los trabajos previos analizados presentaron una amplia variedad de enfoques para realizar la adaptación de las tareas del proceso de recuperación de información, a continuación, se describen los más relevantes. Luego se resume la colección de prueba que se usó para la evaluación del SRI desarrollado.

Virmani & Taneja (2018), proponen una Técnica de Preprocesamiento de Texto (TPIR), para mejorar la eficiencia de las tareas de preprocesamiento, la cual consta de 4 fases: División del texto en párrafos u oraciones y extracción de la información sintáctica a través de tokenización; Combinación de tokens similares; Separación de palabras vacías y tokens; y aplicación de un algoritmo de stemming basado en corrección ortográfica. La evaluación se realizó sobre The International Monetary Fund, obteniendo mejores resultados en comparación con el preprocesamiento tradicional, al reducir tokens y espacios de almacenamiento. Gupta, Patel, & Bramhe (2018), proponen un enfoque de recuperación de documentos que utiliza un índice Hash teniendo en cuenta características de los términos (term features). La metodología propuesta utiliza 5 pasos: Preprocesamiento, donde cada documento se convierte en un vector de características, y se eliminan las palabras vacías, se compara con un vector de palabras clave para encontrar el vector de características del documento, luego, un conjunto de características positivas y negativas se utiliza para asignar el ID del término, se genera un número Hash para el documento y el valor de índice asignado da la identificación del documento, finalmente, las consultas del usuario reciben el mismo procesamiento, de forma que los id de los términos son la clave para la función hash, la cual recupera los documentos que coinciden. Los experimentos se realizaron con datasets reales y artificiales.

Hammo, Yagi, y otros (2015), describen el proceso de recuperación de información para el Arábigo, mostrando los cambios semánticos de la lengua en relación con la frecuencia de uso de una palabra a través del tiempo y cómo cambia su significado en el tiempo. El tokenizador implementado para este procesamiento incluye dos salidas: tokens que corresponden a unidades léxicas cuyos caracteres son reconocibles y tokens que necesitan más análisis morfológico. Utilizaron una lista de palabras vacías (palabras que aportan poco al significado y tienen alta frecuencia) de tal forma, que se mejora la velocidad de

búsqueda. Adicionalmente, se pudo apreciar que esta lista de palabras difícilmente ha cambiado durante los pasados 16 siglos en los documentos escritos en árabe.

Guan y Zhang (2013), presentan aspectos relevantes sobre el procesamiento de textos en chino, proponen un algoritmo que mejora el método de segmentación de texto aplicado al modelo espacio vectorial, que logra una mejora significativa en las tasas de similitud sobre textos de noticias en chino, mostrando que sus resultados son casi cercanos a la segmentación humana y puede remplazarla, ahorrando tiempo y dinero, y mejorando la eficiencia de la similitud de textos de noticias.

Jiang y Zhai (2007), presentan la generalización de varias estrategias de tokenización en un conjunto de heurísticas organizadas para ser aplicadas en la recuperación de información en biomedicina, concluyendo que remover caracteres no funcionales es seguro, y que se pueden aplicar diferentes tipos de tokenización, lo cual presenta mejoras en el desempeño de hasta un 8% en la medida de evaluación utilizada, el índice MAP. También evaluaron empíricamente el impacto de utilizar o no una lista de palabras vacías sobre el desempeño de un SRI, encontrando que no mejora el desempeño o ligeramente lo mejora, en el experimento se utilizaron 132 palabras vacías de inglés tomadas del PUBMED.

Klatt y Bohn (2004), presentan un proceso de tokenización para una colección de prueba en Alemán, demostrando las diferentes opciones y pasos que se pueden realizar en el análisis léxico permitiendo: 1) La separación de los respectivos tokens, y 2) Complementar la información de los tokens con anotaciones de estructuras lingüísticas.

Los trabajos previos fueron significativos para la investigación ya que mostraron por ejemplo, la definición de reglas heurísticas y de varios pasos de análisis para el proceso de tokenización; el uso de Lucene que permitió visualizar su aplicabilidad en la construcción de SRI para textos escritos en diferentes lenguas, por tanto, su uso fue incorporado en el desarrollo de esta propuesta; adicionalmente, se pudo apreciar que aunque teóricamente la definición de una lista de palabras vacías puede llevar a incrementar el desempeño de un SRI (Baeza-Yates, 2004), en la literatura se aprecian diferentes casos, unos donde se logran mejoras significativas con el uso de la lista de palabras vacías, pero otros donde se observan ligeras mejoras o no existen dichas mejoras, eso sí, en todos los casos se reduce el espacio requerido para el almacenamiento.

La Comunidad Nasa y su lengua, el Nasa Yuwe

La comunidad Nasa es la más numerosa población indígena del departamento del Cauca (Colombia) (Pachón C., 1996). El Nasa Yuwe es una de las lenguas oficiales de la República de Colombia, es hablada por la comunidad Nasa (Instituto Colombiano de Cultura Hispánica, 2000). Ha sido una lengua de tradición oral, y solo en el siglo XX es cuando se define su alfabeto y más tarde se da la unificación de este alfabeto, lo cual representó un importante paso para encontrar caminos que coadyuven en su desarrollo y revitalización (Rojas Curieux T., 2006). La lengua Nasa Yuwe, ha sido considerada recientemente como lengua independiente (no clasificada) (Rojas, 2012). El alfabeto unificado del Nasa Yuwe consta de 32 vocales y 61 consonantes (Farfán Martínez & Rojas Curieux, 2010). Para efectos del desarrollo de este trabajo se han utilizado textos escritos con el alfabeto unificado.

Colección de prueba para textos escritos en Nasa Yuwe

La primera colección de prueba de textos escritos en Nasa Yuwe (Sierra Martínez, Cobos Lozada, Corrales Muñoz, & Rojas Curieux, 2015) fue construida como una herramienta para la evaluación del prototipo de SRI para textos escritos en dicha lengua. En el proceso de conformación de los elementos de la colección participaron profesores Nasa hablantes de Nasa Yuwe y un lingüista experto en la lengua, quienes hicieron revisión y edición de los documentos seleccionados para conformar la colección, emitieron sus conceptos sobre las necesidades de información o consultas a realizar sobre esta colección, y emitieron los juicios de relevancia sobre cada par documento-consulta. La colección de prueba quedó conformada como se muestra en la Tabla 1.

Características	**Valor**
Tamaño de la colección	113 KB
Formatos de los documentos	(UTF-8) text
Número de documentos	97
Promedio de terminos por documento	103
Tamaño promedio de documentos	1.5 KB
Número de consultas	8

Tabla 1 – Características de la colección de prueba de Nasa Yuwe

3. Metodología

El proceso de adaptación de las tareas de RI se enmarcó en el Patrón de Investigación Iterativa (Iterative Research Pattern) (Pratt, 2009). Se dio prioridad a dos tareas de RI: la adaptación del analizador léxico (tokenizer) y la definición de una lista de palabras vacías (stop words removal list) para Nasa Yuwe. Además, como medida de evaluación se utilizó la curva Precisión-Recuerdo, comúnmente usada por la comunidad académica y científica de RI.

Primero se identificaron tres casos para realizar la evaluación de desempeño tanto de la línea base como del SRI propuesto en este trabajo, estos casos fueron: Caso 1, Evaluar con todas las consultas (1 a 8) de la colección, con lo cual se obtiene el desempeño del sistema teniendo en cuenta las diferentes condiciones particulares de la lengua Nasa Yuwe; Caso 2, las consultas 1 a 7, de tal forma que se pudiese revisar el impacto de la consulta 8 sobre el desempeño del sistema, dado que esta consulta solo tiene un documento relevante; y Caso 3, las consultas 2 a 4, que permitieron evaluar el comportamiento del sistema con pocas consultas y con las consultas que más documentos relevantes tienen.

Luego, se definió como punto de partida las librerías provistas por Lucene para .NET versión 2.9.4 (Lucene Apache, 2009), de tal forma que se definiera la línea base del proceso de adaptación del analizador léxico para Nasa Yuwe. Los pasos seguidos fueron: 1) Definición de una línea base utilizando el tokenizador estándar de Lucene, obteniendo los tokens de los documentos y las consultas de la colección de prueba construida para Nasa Yuwe. 2) Se revisaron, identificaron y documentaron los problemas en la tokenización para los documentos de la colección. 3) Se aplicó un filtro para remover

terminaciones de palabras (postfijos), con el cual no se obtuvo ninguna mejora, por tanto, se decidió no tenerlo en cuenta en los procesamientos siguientes. 4) Se adicionó al tokenizador estándar un convertidor de textos a letras minúsculas, obteniendo mejoras significativas en los valores de la curva Precisión-Recuerdo. 5) Finalmente, se repitió el paso 2, para los resultados obtenidos con el procesamiento descrito.

Después, se analizaron cada uno de los problemas identificados en la tokenización, con el fin de encontrar las posibles causas y definir las mejoras a realizar en el tokenizador estándar de Lucene, para adaptarlo al procesamiento de documentos escritos en Nasa Yuwe.

Después se tomó como línea base el desempeño del tokenizador Nasa adaptado con la lista de palabras vacías del español para los tres casos y se empezó a evaluar el desempeño con las palabras candidatas definidas por el Lingüista experto en Nasa Yuwe, y así se definieron las palabras que conformaron la lista de palabras vacías definitiva para Nasa Yuwe.

Finalmente, se utilizaron otras medidas de evaluación, como la Medida F1 at K, Precisión at K, Recuerdo at K, entre otras, con el fin de realizar análisis adicionales.

4. Resultados de la adaptación

El modelo usado para la representación de los documentos en el SRI para textos escritos en Nasa Yuwe, fue el espacio vectorial, además se utilizó la métrica de similitud propuesta por Lucene (Lucene Apache, 2009), que es una mejora a la similitud de cosenos.

4.1. Errores detectados en el procesamiento de los textos usando el tokenizador estándar de Lucene

Entre los errores detectados en los tokens al utilizar el tokenizador estándar de Lucene (ST, Standard Tokenizer) + el convertidor de textos a minúsculas (LC, Lowercase Converter) se encuentran: Error 1: La separación inadecuada de palabras. Error 2: el cambio de una vocal interrupta por una oral al final de una palabra. Error 3: En algunos casos las vocales Nasales generaban también una división y un reemplazo por la correspondiente vocal oral; en la Tabla 2, se presentan algunos ejemplos de estos errores.

Palabra Nasa	Procesada	Tokens detectados	Tokens correctos	Error encontrado
u'pu'	u pu	2	1	Error 1
je'z	je z	2	1	Error 1
ẽente'	ẽente	1	1	Error 2
sayu'	Sayu	1	1	Error 2
ẽente'	e ente	2	1	Error 3
kũhku'ka'wã'	ku hku'ka'wa	2	1	Error 3

Tabla 2 – Ejemplos de errores

El total de ocurrencias de errores encontradas en el procesamiento de los 97 documentos de la colección fue: Error 1: 3936, Error 2: 2531, Error 3: 141, con un máximo de ocurrencias por documento de 218, 182 y 46 respectivamente.

4.2. Analizador léxico (tokenizador) adaptado para Nasa Yuwe

Los problemas del tokenizador estándar de Lucene que fueron detectados se corrigieron para textos escritos en Nasa Yuwe de la siguiente forma:

Error 1 relacionado con la separación inadecuada de palabras. Se deshabilitaron las tildes de las vocales orales y Nasales interruptas como elemento separador de palabras (ver Tabla 2), de la siguiente forma: 1) Se identificaron diferentes tipos de tilde en la forma de escribir documentos escritos en Nasa Yuwe como son: ´, `, ‘,’, es decir, que no sólo se podrían encontrar en los documentos de la colección sino en las consultas del usuario, por tanto, fue necesario unificar estos tipos de tildes por una sola (’) antes de iniciar el procesamiento de textos y consultas en el tokenizador Nasa. 2) Se eliminó la tilde unificada como elemento separador de palabras en un texto u oración.

Error 2 relacionado con el cambio de una vocal interrupta por una oral al final de una palabra. Esto se solucionó manteniendo la tilde (’) de las vocales orales y Nasales interruptas al final de cada palabra, para lograrlo se procedió de la misma manera que en el ítem anterior, es decir, se unificaron los diferentes tipos de tildes y se deshabilitaron como elemento separador de tokens.

Error 3 relacionado con evitar el cambio de las vocales Nasales por vocales orales y la división de la palabra. Esta situación fue un tanto extraña, dado que no se presentaba con todos los documentos. Con este error se procedió de la siguiente forma: 1) Identificar porqué se presentaba esta situación en unas palabras y en otras no, situación que se presentaba porque la vocal Nasal estaba dividida en dos, es decir, al ubicar el cursor sobre la letra se apreció que el carácter estaba formado por dos de ellos, aunque en apariencia era igual al carácter de la vocal nasal. Por tanto, fue necesario identificar si era un error en la escritura o una situación que se podía presentar en la escritura normal de la lengua Nasa Yuwe. 2) Revisar si era un error en la digitalización de los textos, (por ejemplo, al aplicar el OCR al momento de escanear los documentos), un error al aplicar el tipo de fuente (Charis Sil) o un error al utilizar el software de escritura de Nasa Yuwe (Narváez Burbano & García, 2014). Como resultado se encontró que, con estos métodos de escritura no se presentaba esta situación con las vocales Nasales. Finalmente, se revisaron las distintas formas en que los profesores Nasa pueden escribir textos, una de las formas que generó esta situación fue cuando se utilizaba una herramienta de inserción de caracteres especiales como el uso de la página del alfabeto fonético internacional (http://westonruter.github.io/ipa-chart/keyboard), por lo tanto, esta situación se corrigió en el tokenizador para Nasa, dado que se podría presentar al momento de introducir el texto en un documento o en una consulta. 3) La situación se corrigió al incluir un paso adicional en el procesamiento de textos Nasa que permitiera unificar las vocales Nasales tanto en los documentos como en las consultas antes de iniciar el procesamiento.

Adicionalmente a la corrección de los errores detectados, se identificaron algunas palabras que varían en la escritura por ejemplo: Khutx, Kuthx y Kutxh utilizadas para

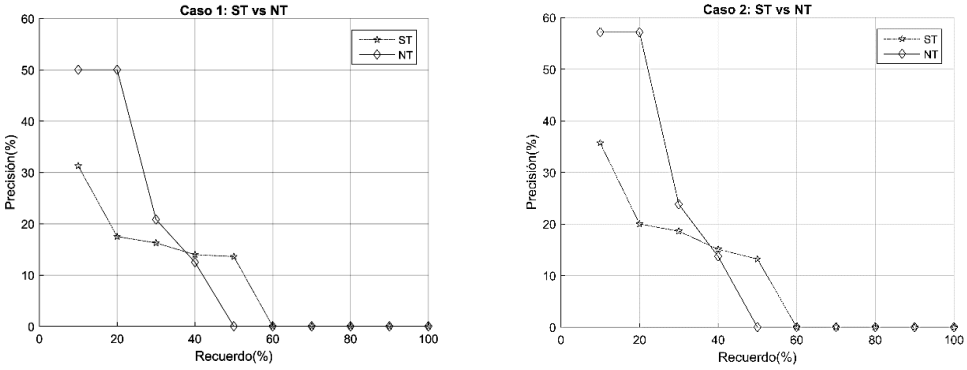
escribir maíz, a'te y a'te' para época o luna dependiendo de la variante del idioma, lo cual es un problema dado que si el usuario para escribir maíz lo escribe de manera diferente a como esta en los documentos de la colección, el desempeño del sistema se ve muy afectado, porque no encontraría documentos relevantes relacionados con la consulta, por esta razón se decidió incluir su unificación en el tokenizador Nasa (manejo de sinónimos, homónimos, entre otros).

Una vez se modificó el tokenizador estándar de Lucene para resolver estas situaciones, se obtuvo el tokenizador Nasa.

4.3. Evaluación de desempeño del tokenizador Nasa

Primero, se presentan algunas curvas de Precisión – Recuerdo con el fin de visualizar el desempeño del tokenizador adaptado para Nasa Yuwe en contraste con el tokenizador estándar de Lucene.

En la Figura 1a, se presentan los resultados del tokenizador estándar (ST) vs el tokenizador Nasa (NT) para el caso 1. Se puede apreciar que para los valores del 10% al 40% de recuerdo el tokenizador Nasa muestra similar o mejor desempeño (respuestas más precisas en los primeros niveles de recuerdo, que es lo mejor cuando se comparan dos o más SRI), para valores del 40% al 60% el ST muestra un mejor desempeño y posteriormente el comportamiento es similar.



(a) (b)

Figura 1 – Desempeño del tokenizador nasa adaptado Caso 1 y Caso 2

En los resultados del tokenizador estándar (ST) y el tokenizador Nasa (NT) para el caso 3, se pudo apreciar: 1) un incremento sustancial en la evaluación para un valor del 10% de recuerdo cuando se aplica el tokenizador Nasa, lo cual se entiende como consecuencia de que para el caso 3, las consultas tienen una cantidad considerable de documentos relevantes; 2) el comportamiento de la curva de Precisión-Recuerdo es muy similar al del Caso 2 que se muestra en la Figura 1b.

A continuación, en la Figura 2 se presenta el desempeño del tokenizador Nasa (NT) + el convertidor de texto a minúsculas (LC) comparado con el tokenizador estándar de

Lucene (ST) + el convertidor de texto a minúsculas (LC) para el caso 2, en donde se puede apreciar: 1) Para el primer valor de recuerdo (10%) el ST + LC tiene mejor desempeño, lo cual es entendible, dado que como se está tokenizando incorrectamente la probabilidad de encontrar similitudes entre documentos y consultas se puede incrementar (basado en los subcomponentes de las palabras). 2) Para valores de recuerdo entre 20 y 50% el desempeño de NT + LC, es mejor, demostrando así que las mejoras incluidas en la adaptación del tokenizador son adecuadas a las características de los textos escritos en lengua Nasa.

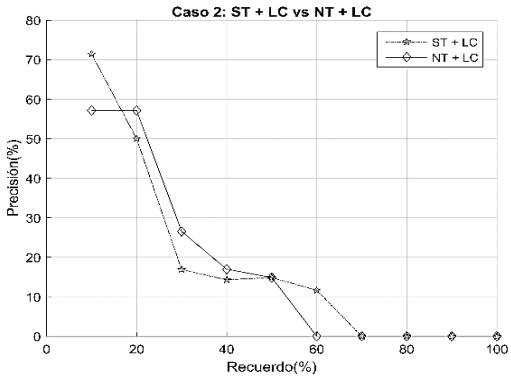


Figura 2 – Desempeño del tokenizador nasa adaptado Caso 2

Durante la revisión de porque se presentan bajas en los valores iniciales de la curva Precisión-Recuerdo, se obtuvo que cuando se utiliza ST + LC, se encuentran errores en la tokenización, por tanto al tokenizar mal las palabras se cuenta con una mayor probabilidad de encontrar palabras comunes en más documentos, aunque no sea en realidad la palabra buscada, por ejemplo: si para consultar se usa la palabra U'JWE, el tokenizador estándar saca dos tokens de ella U y JWE, en el texto de los documentos se puede encontrar la palabra U'JWE' la cual saca dos tokens U y JWE; a pesar de que son palabras de escritura similar son diferentes, pero ST + LC puede tener mayores valores de precisión si estos tokens coinciden en documentos que son relevantes a la consulta de usuario. Mientras que el tokenizador Nasa las trata como términos diferentes, reduciendo su aparición solo a los tokens correctos en cada documento.

Para el caso 3, se observó un comportamiento similar al descrito para la Figura 2, sin embargo, se apreció un mejor desempeño de NT + LC en los primeros valores de recuerdo, y para valores de recuerdo del 40% en adelante los dos tokenizadores tienen un comportamiento similar, de tal forma, que para este caso de estudio (consultas con más documentos relevantes), NT + LC presenta un mejor desempeño para el primer valor de recuerdo y una disminución de precisión para los siguientes valores, disminución que es más notoria que la que se aprecia en el caso 2.

En la evaluación del sistema utilizando NT vs NT + LC para los casos 2 y 3, se observó un comportamiento similar en los primeros valores de recuerdo y mejoras en el desempeño al utilizar el NT + LC, lo cual es entendible dado que al unificar términos se puede incrementar las coincidencias en los documentos.

Finalmente, es preciso mencionar que las palabras Nasa de los documentos de la colección de prueba quedaron correctamente tokenizadas, por tanto, la adaptación realizada al tokenizador estándar de Lucene para el procesamiento de textos escritos en Nasa Yuwe fue exitosa, a pesar de que, para algunos valores de recuerdo, la precisión es levemente inferior en relación con la aplicación del tokenizador estándar (que presenta errores). La cantidad de tokens para la colección de textos escritos en Nasa Yuwe en cada procesamiento resultó así: 1) ST + LC, 4149 términos; 2) NT, 5020 términos; y 3) NT + LC, 4801 términos.

4.4. Definición de una lista de palabras vacías para Nasa Yuwe

Inicialmente se tomó como línea base la lista de palabras vacías usadas para español, lo que no presentó ninguna variación (conforme se esperaba) en el desempeño del sistema (NT + LC), por tanto, se hizo necesario definir una lista de palabras vacías específica para Nasa Yuwe.

Luego, para la definición de la lista de palabras vacías para Nasa Yuwe se tomaron las 100 palabras con mayor frecuencia en la colección de prueba, y con la ayuda del Lingüista experto en Nasa Yuwe se definieron 34 palabras candidatas a ser palabras vacías. Una vez definidas las palabras candidatas se procedió a realizar las evaluaciones de desempeño requeridas utilizando el NT + LC para cada caso de estudio. Como resultado (al utilizar las 34 palabras candidatas), se obtuvo una disminución en el desempeño, por tanto, fue necesario revisar el impacto de cada palabra candidata sobre el desempeño del sistema en contraste con lo logrado por el sistema usando todas las palabras.

Finalmente, se obtuvo la lista definitiva de palabras vacías para Nasa Yuwe, de aquellas palabras que son comunes para todos los casos quedando conformada la lista por 16 palabras así: Txãa, naa, teeçx, seena, wẽt, mẽh, ma'w, pa'ka, txa'w, txajuyu', ũskan, u'jn, meeçxa, meeçxa', txa's, txajx. En la Figura 3, se observa un mejor desempeño en la curva al utilizar NT + LC + las palabras vacías definidas para Nasa Yuwe, para el caso 1.

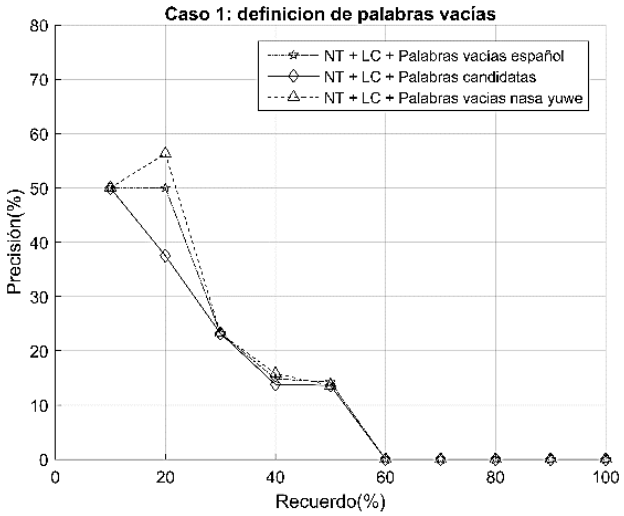


Figura 3 – Desempeño utilizando lista palabras vacías Caso 1

Con esta lista de palabras vacías para Nasa Yuwe, se procedió a definir su impacto en el tamaño de la estructura del índice, el cual implicó una disminución inferior al 6%. En la Tabla 3, se presentan los diferentes procesamientos considerados para cada caso.

Caso	Procesamiento	Tamaño de la estructura del índice	% de decremento
Caso 1	NT + LC + STW(SPANISH)	224 KB	---
	NT + LC + Empty word list (34 candidates)	211 KB	5.80%
	NT + LC + Empty word list (16)	218 KB	2.68%
Caso 2	NT + LC + STW(SPANISH)	224 KB	-----
	NT + LC + Empty word list (34 candidates)	211 KB	5.80%
	NT + LC + Empty word list (16)	218 KB	2.68%
Caso 3	NT + LC + STW(SPANISH)	224 KB	-----
	NT + LC + Empty word list (34 candidates)	211 KB	5.80%
	NT + LC + Empty word list (16)	218 KB	2.68%

Tabla 3 – Impacto de la lista de palabras vacías en el índice

En la Tabla 4, se presentan las medidas de Precisión at K, Recuerdo at K y Medida F1 at K resumidas para todas las consultas. Allí se puede observar que la evaluación para el NT + LC + Lista de palabras vacías, es mejor en la mayoría de los resultados que las evaluaciones para ST + LC+ Lista de palabras vacías, lo que implica que NT + LC+ Lista de palabras vacías recupera más documentos relevantes (recuerdo), pero además los resultados son más precisos a diferentes cantidades de resultados (at K significa a cierta cantidad K de resultados presentados al usuario).

Resultado	ST + LC + Nasa Yuwe empty word list			NT + LC + Nasa Yuwe empty word list		
	Recall at K	Precision at K	F1 Measure at K	Recall at K	Precision at K	F1 Measure at K
1	**8.06%**	**62.50%**	**14.28%**	6.45%	50.00%	11.43%
2	11.29%	43.75%	17.95%	**14.52%**	56.25%	23.08%
3	11.29%	29.17%	16.28%	**17.74%**	**45.83%**	**25.58%**
4	11.29%	21.88%	14.89%	**17.74%**	**34.38%**	**23.40%**
5	11.29%	17.50%	13.73%	**17.74%**	**27.50%**	**21.57%**

Tabla 4 – Medidas de desempeño (mejores resultados en negrita)

4.5. Prototipo de Sistema de Recuperación de Información para Nasa Yuwe

Para la construcción del prototipo de SRI para textos escritos en Nasa Yuwe, se utilizó la metodología XP (Tangient LLC., 2013). El objetivo de la construcción del prototipo fue la evaluación del desempeño del SRI y la recuperación de los textos escritos en Nasa Yuwe de la colección de prueba. En total se definieron 4 historias de usuario distribuidas

en 3 iteraciones. El prototipo se desarrolló con una arquitectura de 3 capas (Microsoft, 2010) y la tecnología de desarrollo utilizada fue VS.NET 2010 usando C# y una base de datos SQL Server 2008.

La Figura 4 presenta una interfaz del prototipo de SRI para Nasa Yuwe. En esta figura se puede apreciar que en la zona 1 se muestra el lugar donde se encuentran almacenados los índices de búsqueda de Lucene, sobre los cuales se realizan las consultas, seguidamente se encuentra un espacio en donde se digitan las palabras de la consulta y el botón para iniciar el proceso de búsqueda. En la zona 2, que se divide en dos partes, presenta los resultados de la búsqueda; en la parte izquierda se encuentra la lista de documentos recuperados mostrada mediante el nombre del archivo y las primeras líneas que contiene el documento de texto de la colección de prueba recuperado, allí se puede seleccionar el documento que sea de interés y en la parte derecha de la zona 2, se encuentra un botón mediante el cual se puede visualizar el documento seleccionado en la lista de la parte izquierda.

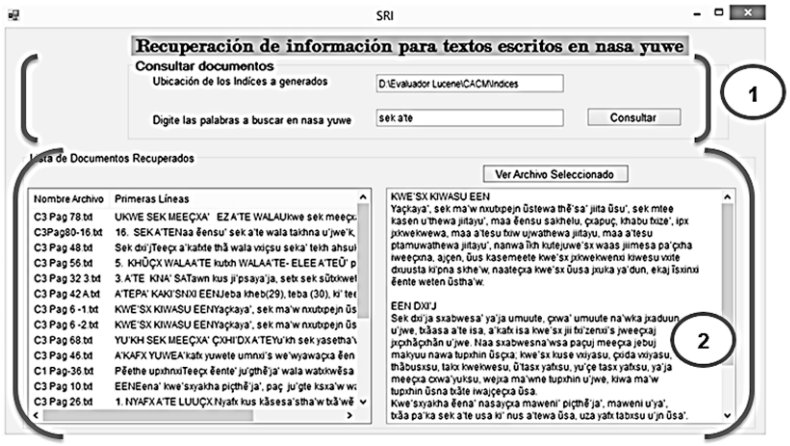


Figura 4 – Interfaz prototipo de Recuperación de Información para textos escritos en Nasa Yuwe

5. Conclusiones

En este trabajo se realizó la adaptación del tokenizador estándar de Lucene para textos escritos en Nasa Yuwe, de manera exitosa, corrigiendo los errores del tokenizador estándar.

A pesar de que el Nasa Yuwe cumple con la Ley de Zipf, al momento de definir la lista de palabras vacías identificada para esta lengua, no se obtuvo el resultado esperado a nivel de mejorar el desempeño en los valores de la curva precisión-recuerdo y a nivel de la reducción del tamaño de la estructura del índice, es decir, mientras en la mayoría de las lenguas se pueden obtener reducciones hasta del 40% en el tamaño del índice (Baeza-Yates & Ribeiro-Neto, Modern Information Retrieval, 1999), en Nasa Yuwe fue cercana tan solo al 6%. Se requiere ampliar el tamaño de la colección para llegar a una conclusión definitiva a este respecto.

El prototipo permitió la evaluación del desempeño del SRI para Nasa Yuwe utilizando diferentes parámetros, incluido el tokenizador a usar (estándar o Nasa Yuwe), la aplicación o no del convertidor de textos a minúsculas (usándolo o no) y el uso o no de la lista de palabras vacías definida para Nasa Yuwe.

Se logró proponer un esquema de representación y búsqueda para textos escritos en Nasa Yuwe, utilizando el modelo de espacio vectorial y haciendo la adaptación de un tokenizador para Nasa y la definición de una lista de palabras vacías para Nasa Yuwe. Este esquema fue materializado en un prototipo de sistema de recuperación de información para textos escritos en dicha lengua.

La experiencia obtenida en la realización de este trabajo es muy valiosa, dado que no existen antecedentes similares para Nasa Yuwe tanto en la adaptación de las tareas de análisis léxico (tokenizer) como en la elaboración de la lista de palabras vacías (stop words removal list), por tanto, se convierte en un referente para futuros trabajos en esta lengua o en otra de similar origen o características.

El grupo de investigación espera en futuras investigaciones soportar otras actividades de procesamiento de textos en Nasa Yuwe, como la identificación de partes del discurso, la unificación sintáctica y semántica de palabras, y el uso de vocabularios que ayuden no sólo en el proceso de búsqueda y recuperación, sino también en la digitación (opción de completar) de los textos de los documentos y las consultas. Además, se espera ofrecer el SRI a través de una plataforma web que permita su uso desde cualquier lugar, tiempo y dispositivo.

Agradecimientos

Los autores agradecen a los Departamentos de Antropología, Sistemas y Telemática y a la Vicerrectoría de Investigaciones de la Universidad del Cauca, además a la Escuela de Ing. de Sistemas y Computación de la Universidad Industrial de Santander, por toda su colaboración para realizar el presente trabajo.

Referencias

Hossein Jadidinejad, A., Mahmoud, F., & Dehdari, J. (2010). Evaluation of Perstem: A Simple and Efficient Stemming Algorithm for Persian. In C. Peters, G. M. Di Nunzio, M. Kurimo, & Others (Hrsg.), *Multilingual Information Access Evaluation I. Text Retrieval Experiments* (Bd. Volume 6241, S. 98-101). Springer Berlin Heidelberg.

Acción Social; Unión Europea; Cordepaz. (2010). *Informacion Sobre Acciones y Procesos Institucionales para los Pueblos Indigenas de Colombia*. Bogota.

Agosti, M., Bacchin, M., Ferro, N., & Melucci, M. (2003). Improving the Automatic Retrieval of Text Documents. In C. Peters, M. Braschler , & J. Gonzalo (Hrsg.), *Advances in Cross-Language Information Retrieval* (Bd. Volume 2785, S. 279–290). Springer Berlin Heidelberg.

Baeza-Yates, R. (2004). Challenges in the Interaction of Information Retrieval and Natural Language Processing. In C. 2.-2. 5th International Conference, *Computational Linguistics and Intelligent Text Processing* (Bd. Volume 2945, S. 445–456). Springer Berlin Heidelberg.

Baeza-Yates, R., & Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. New York: Pearson - Addison Wesley.

Carterette, B., & Voorhees, E. (2011). Overview of Information Retrieval Evaluation. In *Current Challenges in Patent Information Retrieval* (S. 69–85). Springer Berlin Heidelberg.

Consejo Regional Indígena del Cauca - CRIC. (kein Datum). *Consejo Regional Indígena del Cauca - CRIC*. Abgerufen am 18. Octubre 2012 von http://www.cric-colombia.org/portal/estructura-organizativa/plataforma-de-lucha/

Consejo Regional Indígena del Cauca - CRIC, & Programa de Educación Bilingüe e Intercultural -, P. (2011). *Sistema Educativo Indígena Propio -SEIP. Primer Documento de Trabajo*. CRIC .

CRIC-Programa de Educacion Bilingüe e Intercultural-Comisión General de Lenguas. (Junio 2008). *Estudio Sociolingüisticos Fase preliminar. Base de datos - CRIC 01/2007 Lengua Nasa Yuwe y Namtrik*. Popayàn, Cauca, Colombia: CRIC-Universidad del Cauca-Ministerio de Cultura.

Cui, Y., Chen, Y., & Li, J. (2011). Research of Information Search Engine in Forestry Based on the Lucene. In *Advances in Automation and Robotics, Vol. 2. Series Lecture Notes in Electrical Engineering* (S. 603–609). Volume 123 of the s pp: Springer Berlin Heidelberg. doi:10.1007/978-3-642-25646-2_78

Farfán Martínez, M., & Rojas Curieux, T. (2010). *Zuy Luuçxkwe kwe'kwe'sx ipx kwetuy piyaaka. Cartilla de aprendizaje de Nasa Yuwe como segunda lengua*. Buenos Aires.

Guan, W., & Zhang, P. (2013). Research and application of news-text similarity algorithm based on Chinese word segmentation. *3rd International Conference on Consumer Electronics, Communications and Networks (CECNet)* (S. 484–487). Xianning: IEEE Xplore Digital Library.

Gupta, R., Patel, D., & Bramhe, A. (2018). A Hash-Based Approach for Document Retrieval by Utilizing Term Features. En Advances in Intelligent Systems and Computing 4th International Conference on Computational Intelligence in Data Mining, ICCIDM 2017 (págs. 617-627). Sambalpur, India: Springer Verlag. doi:10.1007/978-981-10-8055-5_55

Hammo , B., Yagi, S., Ismail, O., & AbuShariah, M. (2015). Exploring and exploiting a historical corpus for Arabic. *Language Resources and Evaluation*, 1–23. doi:10.1007/s10579-015-9304-9

Instituto Colombiano de Cultura Hispánica. (2000). *Geografía Humana de Colombia, Región Andina Central Tomo IV Volumen II*. Bogotá: Banco de la República.

Jiang, J., & Zhai, C. (October 2007). An empirical study of tokenization strategies. *Information Retrieval, 10*(4), 341–363. doi:10.1007/s10791-007-9027-7

Klatt, S., & Bohnet, B. (2004). You Don't Have to Think Twice if You Carefully Tokenize. In *Natural Language Processing – IJCNLP 2004 - Lecture Notes in Computer Science* (Bd. 3248, S. 299–309). Hainan Island, China: Springer Berlin Heidelberg. doi:10.1007/978-3-540-30211-7_32

Kumar Pandey, A., & Siddiqui, T. (2009). Evaluating Effect of Stemming and Stop-word Removal on Hindi Text Retrieval. In *Proceedings of the First International Conference on Intelligent Human Computer Interaction* (S. 316–326). Allahabad, India: Springer India.

Li, H., Li, W., Wang, G., & Peng , X. (2012). Information Retrieval Services Based on Lucene Architecture. In *Information Computing and Applications Volume 307 of the series Communications in Computer and Information Science* (S. 638–645). China: Springer Berlin Heidelberg. doi:10.1007/978-3-642-34038-3_88

Lucene Apache. (2009). *Lucene Apache*. Abgerufen am Julio 2009 von https://lucene.apache.org/core/3_6_2/api/core/org/apache/lucene/search/Similarity.html

Manning, C., Raghavan, P., & Shütze, H. (2009). *An Introduction to Information Retrieval*. Cambridge University Press.

Microsoft. (2010). *Microsoft*. Abgerufen am Octubre 2013 von https://www.google.com.co/search? q=arquitectura+de+tres+capas+microsoft & oq=a & aqs=chrome.5.69i60l4j69i57j69i59.11903j0j4 & sourceid=chrome&espv=210 & es_sm=93 & ie=UTF-8

Narváez Burbano, E., & García, M. (2014). *Monografía y Anexos: Alternativa para la entrada de caracteres en lenga Nasa Yuwe aplicada a la producción de materiales tipo texto*. Popayán: Tesis Ingeniería de Sistemas Universidad del Cauca.

Organización Nacional Indígena de Colombia - ONIC. (Abril 2012). *Organización Nacional Indígena de Colombia - ONIC*. Abgerufen am 21. Octubre 2012 von http://cms.onic.org.co/2012/04/es-urgente-aprobar-declaracion-americana-sobre-los-derechos-de-los-pueblos-indigenas/

Pachón C., X. (1996). Los Nasa o la Gente Páez. In Instituto Colombiano de Cultura Hispánica, *Geografía Humana de Colombia. Región Andina Central Tomo IV* (Bd. 2). Bogotá.

Pratt, K. S. (2009). *Design Patterns for Research Methods: Iterative Field Research*. Recuperado el Octubre de 2013, de http://www.kpratt.net/wp-content/uploads/2009/01/research_methods.pdf

Ribeiro-Neto, B., & Baeza-Yates, R. (2011). *Modern Information Retrieval -the concepts and technology behind search* (Second Ausg.). Harlow: Addison Wesley.

Rojas Curieux, T. (2006). *Por los caminos de la recuperación de la lengua Paéz (Nasa Yuwe)*. Popayán: Letrarte editores.

Rojas, T. E. (2012). Esbozo Gramatical de la lengua nasa (lengua Paéz). In UNICEF (Hrsg.), *El Lenguaje en Colombia. Tomo I: Realidad Lingüístidca de Colombia*. Bogotá: Academia Colombiana de la Lengua e Instituto Caro y Cuervo.

Sierra Martínez, L. M., Cobos Lozada, C. A., Corrales Muñoz, J. C., & Rojas Curieux, T. (2015). Building a Nasa Yuwe Test Collection. In *Computational Linguistics and Intelligent Text Processing Volume 9041 of the series Lecture Notes in Computer Science* (S. 112–123). El Cairo, Egipto: Springer International Publishing. doi:10.1007/978-3-319-18111-0_9

Tangient LLC. (2013). *Entorno virtual de aprendizaje*. Abgerufen am 27. Enero 2013 von Entorno virtual de aprendizaje: http://programacion-extrema.wikispaces.com/

Virmani, D., & Taneja, S. (2018). A text preprocessing approach for efficacious information retrieval. En Advances in Intelligent Systems and Computing nternational Conference on Smart Innovations in Communications and Computational Sciences, ICSICCS 2017 (Vol. Volume 669, págs. 13-22). Moga- India: Springer Verlag. doi:10.1007/978-981-10-8968-8_2

Zhang, C., & Zhan, S. (2012). Research and Implementation of Full-Text Retrieval. In *Proceedings of the 2012 International Conference on Communication, Electronics and Automation Engineering* (Bd. 181, S. of the series Advances in Intelligent Systems and Computing pp 349-356). Springer Berlin Heidelberg. doi:10.1007/978-3-642-31698-2_50

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 22/10/2018
Aceitação/Acceptance: 29/12/2018

Sobre el Éxito de Proyectos y el Framework PSA: Un estudio Fenomenológico en la industria chilena

Manuel Cáceres[1], Aldo Quelopana[1]

mcp018@alumnos.ucn.cl, aldo.quelopana@ucn.cl

[1] Universidad Católica del Norte, Avenida Angamos 0610, 1270709, Antofagasta, Chile.

Pages: 423–432

Resumen: Pese a la gran cantidad de literatura relacionada con el éxito de los proyectos de Sistemas de Información, todavía no se ha llegado a un consenso que permita a investigadores y profesionales trabajar en conjunto para reducir las bajas tasas de éxito reportadas por este tipo de proyectos. El Framework PSA (Project Success Analysis) es un esfuerzo por solucionar esta situación, debido a que presenta una serie de conceptos y sus correspondientes relaciones que permite este análisis. La presente investigación realiza una validación de los componentes de este framework en empresas chilenas, en particular en la región de Antofagasta. Se condujo un estudio fenomenológico (cualitativo) que entregó una serie de resultados, siendo el principal de ellos la importancia de considerar adecuadamente a todos los involucrados (stakeholders) del proyecto, en el proceso de acordar lo que se entenderá por el éxito de un proyecto en particular.

Palabras-clave: éxito de proyecto; sistemas de información; criterios; factores; framework.

About Project Success and the PSA Framework: A Phenomenological study on the Chilean industry

Abstract: Despite the great quantity of literature related to the project success in the Information Systems, there is still no consensus that allows researchers and professionals to work together to reduce the low success rates reported for this type of projects. The PSA (Project Success Analysis) framework is an effort to resolve this situation, since it presents a variety of concepts and their relations, which allows the success analysis. The current research validates the framework components for Chilean companies, particularly in the Antofagasta Region. A phenomenological study (qualitative) was conducted which delivered a series of results, the main one was the importance of properly considering all the stakeholders involved in the project in the process of defining what will be considered as the success of a specific project.

Keywords*:* project success; information systems; criteria; factors, framework.

1. Introducción

El "éxito" y su correspondiente obtención dentro de un proyecto, ha sido un tema de discusión entre varios investigadores a lo largo de los años en el área de los Sistemas de Información (SI) (Thomas & Fernández, 2008) (Ika, 2009) (Howsawi, Eager, & Bagia, 2011) (Irvine & Hall, 2015). Las altas tasas de empresas que catalogan sus proyectos como fallidos (Chen, Law, & Yang, 2009) (The Standish Group International, 2014), han generado una atención permanente en el tema, que se ha manifestado por la gran cantidad de publicaciones al respecto (Bloch, Blumberg, & Laartz, 2012) (Drury-Grogan, 2014) (Cecez-Kecmanovic, Kautz, & Abrahall, 2014) (Doherty, Ashurst, & Peppard, 2012). Sin embargo, después de todo este tiempo, lejos de obtener resultados definitivos o consenso frente al tema, podemos indicar con seguridad que se cumple la paradoja de (Cobb, 1996): "conocemos por qué los proyectos [de SI] fallan y sabemos cómo prevenir sus fallas, entonces ¿Por qué todavían fallan?".

Tal como existe una gran probabilidad de perderse en un lugar desconocido cuando no se dispone de un mapa o GPS, nuestro análisis nos ha llevado a considerar que mucho del esfuerzo realizado por los investigadores y profesionales a lo largo de los años, para entender y solucionar los problemas que evitan el éxito de los proyectos, se ha diluido al no contar con una visión holística que enmarque las discusiones y alcance de las investigaciones, logrando una colaboración que permita solucionar problemas y ambigüedades. Es por esta razón que en (Quelopana & Flores, 2018) se hace la presentación de un framework denominado PSA (Project Success Analysis) con el fin de aportar con la visión mencionada. Este framework fue propuesto para ser analizado por la comunidad, por medio de su utilización y mejoras, de manera de contribuir a una solución en conjunto.

Con el objetivo de avanzar con la intención del framework PSA, el presente trabajo busca validar que los conceptos y sus correspondientes relaciones mencionadas, son de conocimiento y uso por parte de las empresas en Chile. Se presenta entonces un estudio inicial localizado en la región de Antofagasta el cual tiene un carácter de índole cualitativo (fenomenológico) (Leedy & Ormrod, 2012), debido a que estará basado en las percepciones que los distintos involucrados (stakeholders) han tenido en sus proyectos.

A fin de lograr el objetivo, el trabajo ha sido dividido de la siguiente forma, primero se presentarán las definiciones que involucran el framework PSA, luego de lo cual se presentará la metodología utilizada en la investigación. Posteriormente se evidenciarán los resultados encontrados con este trabajo, para finalizar con las conclusiones.

2. Framework PSA

La base de este estudio exploratorio fue el framework PSA (Quelopana & Flores, 2018). Este framework fue diseñado para consensuar la concepción del éxito para un proyecto en particular, a fin de que todos los stakeholders pudieran trabajar sobre acuerdos, minimizando o incluso eliminando de esta manera, ambiguedades que afecten lo que se desea lograr con el proyecto. Este framework se compone de cuatro niveles de definiciones, los cuáles son mostrados en la Figura I.

Cabe destacar que este framework se enfoca en organizar aspectos y relaciones de alto nivel asociadas al éxito del proyecto, no asi como el de otros autores como el propuesto por de Witt (1998), que se enfoca principalmente en el cumplimiento de objetivos, u otros que se enfocan en obtener la informacion de los resultados de los proyectos a fin soportar su correspondiente gestion del conocimiento (Todorović et al., 2015).

2.1. Nivel I: Éxito del proyecto

Es el nivel más alto y se utiliza como una indicación de que el éxito del proyecto se logra mediante el éxito del producto y de la gestión del proyecto a la misma vez. Si alguno de ellos falla, entonces el éxito del proyecto en general se verá afectado negativamente.

2.2. Nivel II: Criterios

Una vez que se tiene claridad en cuanto a cómo se compone el éxito, el segundo nivel busca la definición de principios o estándares que permitan juzgar a un proyecto como exitoso o no. Estos criterios deben estar relacionados al proyecto particular que se llevará a cabo y deben estar asignados tanto a la gestión del proyecto como al producto.

2.3. Nivel III: Factores

Teniendo los criterios definidos, es necesario analizar los factores que les impactan tanto positiva como negativamente. Es importante mencionar que estos factores pudieran influenciar a criterios asignados tanto al producto como a la gestión del proyecto.

2.4. Nivel IV: Ciclo de vida

Este último nivel busca identificar cómo el ciclo de vida elegido para llevar a cabo el proyecto, influencia a los factores dependiendo en qué etapa se encuentran.

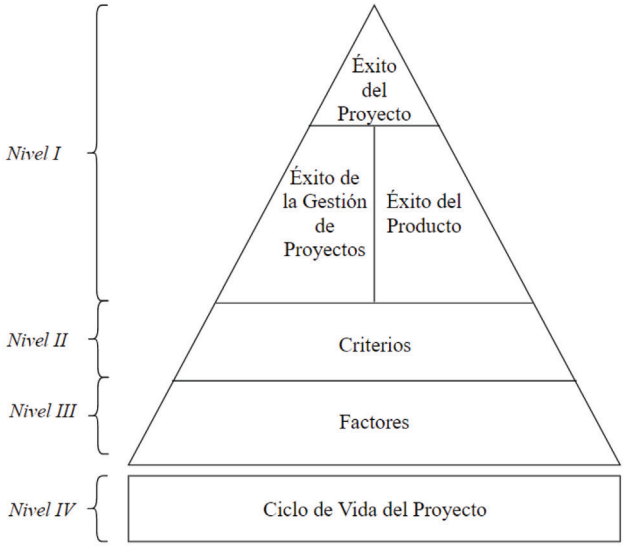


Figura 1 – Framework PSA (Quelopana & Flores, 2018)

3. Metodología

Se llevó a cabo un estudio fenomenológico de tipo cualitativo, el cual intenta comprender las percepciones de las personas, sus perspectivas y entendimientos de una situación en particular (Leedy & Ormrod, 2012). El estudio consideró a cinco empresas de la región obteniendo un total de trece proyectos TI. Dos de estas empresas se desempeñan como prestadoras de servicio a la minería, otras dos son organizaciones de educación superior (universitaria), y una ese encuentra en el ámbito del transporte ferroviario. Es importante señalar que se hizo una diferenciación entre la unidad (departamento, área o sección) que prestó el servicio de desarrollo (proveedor) y quién lo recibió (cliente), independientemente si este cliente se encontraba de manera externa a la organización proveedora, como es el caso de las dos empresas prestadoras de servicio a la minería.

El estudio consideró las siguientes etapas (Figura 2): (i) se identificó las declaraciones que se relacionaban con el tema, (ii) se agruparon las declaraciones a fin de obtener un significado relevante, (iii) se buscaron perspectivas divergentes y (iv) se construyó un significado general.

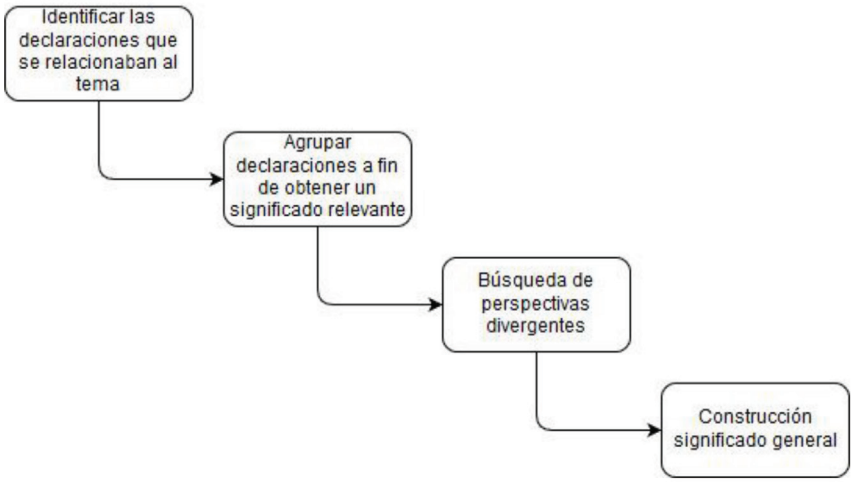


Figura 2 – Etapas Estudio

Como método se utilizaron entrevistas semiestructuradas. Estas entrevistas se realizaron de esta manera a fin de tener una estructura general inicial, pero a la misma vez obtener flexibilidad al momento de indagar en ciertos puntos por medio de preguntas adicionales formuladas en el momento. El conjunto inicial de preguntas realizadas para guiar las entrevistas se encuentra en la Figura 3, en donde se puede apreciar que no todas las preguntas se encontraban directamente dirigidas al framework, dado que tenían por objetivo proporcionar un contexto para el entendimiento del proyecto.

Las entrevistas se llevaron a cabo en primera instancia con los proveedores del proyecto (tal como se menciona, a quienes se les pidió el producto / servicio), luego de lo cual se entrevistó a los clientes de dichos proyectos (quienes se beneficiaron del producto /

servicio). Las entrevistas con los clientes presentaron el mayor desafío debido a la falta de disponibilidad, lográndose abarcar un 80% del total de ellos.

Respecto a la población entrevistada, los representantes de los proveedores se desempeñaban como líderes de sus proyectos o expertos en el desarrollo e implementación de SI. Mientras que, los representantes de los clientes abarcaron un espectro más amplio, debido a que en algunos casos se pudo contactar a usuarios finales, y en otros, se entrevistaron a quienes recibieron el proyecto, usualmente los jefes de unidades.

Los proyectos analizados se encuentran relacionados con el Desarrollo de Software (proyectos a, b, f, g, h, i, j), Inteligencia de Negocio (c), Mejoramiento de Procesos (d), Automatización (e, l), Implementación de Software (k) e Infraestructura (m).

1. Encargado del área (Proveedor)
    a. ¿Cuál es la estructura organizacional que soporta a los proyectos?
    b. Tipos de proyectos trabajados.
    c. Tamaño de los proyectos.
    d. Tamaño de los equipos de trabajo.
    e. ¿Cómo se considera que un proyecto es exitoso (o no)?
    f. ¿Cuál es el porcentaje de proyectos que usted ha tenido que hayan sido catalogado como exitosos?
    g. ¿Qué factores consideraría usted como relevantes para el éxito (o no) del proyecto?
    h. ¿Cuál es la madurez que considera usted que tiene su organización en cuanto a la cultura de proyectos?
2. Jefes de proyectos (Proveedor)
    a. Proyecto X
        i. Tamaño de proyecto
        ii. Tamaño equipo trabajo
    b. ¿Cómo se considera el éxito (o no) del proyecto?
    c. ¿Cuál es el porcentaje de proyectos que usted ha tenido que hayan sido catalogado como exitosos?
    d. ¿Qué factores condicionaron el resultado del proyecto?
    e. ¿A quién iba dirigido el proyecto (clientes)?
3. Clientes
    a. Etapas del proyecto
    b. ¿Cómo consideraría el éxito (o no) del proyecto?
    c. ¿Qué factores condicionaron el resultado del proyecto?

Figura 3 – Instrumento entrevista Semi-Estructurada (elaboración propia)

4. Resultados

De manera general se puede indicar que, de las empresas encuestadas, dos de ellas ocupaban fuertemente metodologías para la gestión de sus proyectos, dos ocupaban algunas practica de gestión de proyectos, y una reconoció no ocupar ninguna metodología para la gestión de proyectos.

El estudio entregó los siguientes resultados:

4.1. Nivel 1

Se determinó por medio de las entrevistas que tanto los proveedores como los clientes identifican los conceptos del éxito propuestos por el framework PSA, pero no cómo éstos se encuentran relacionados (Éxito del Proyecto = Éxito del Producto + Éxito de la Gestión de Proyectos). A nivel general los proveedores relacionan el éxito de la Gestión de Proyectos al tratar obtener rendimiento de distintos parámetros como por ejemplo costo. Por otro lado, los clientes relacionan el éxito al Producto al preocuparse por el desempeño correspondiente y como continuaba a lo largo del tiempo.

Estas percepciones llevan a destacar las diferencias existentes entre proveedores y clientes, debido a que mientras el 60% de los proyectos desarrollados por los proveedores son considerados exitosos, un porcentaje menor de percepción del éxito (40%) es considerado por quienes recibieron el producto / servicio respectivo (clientes), estas diferencias porcentuales pueden apreciarse en la figura 4.

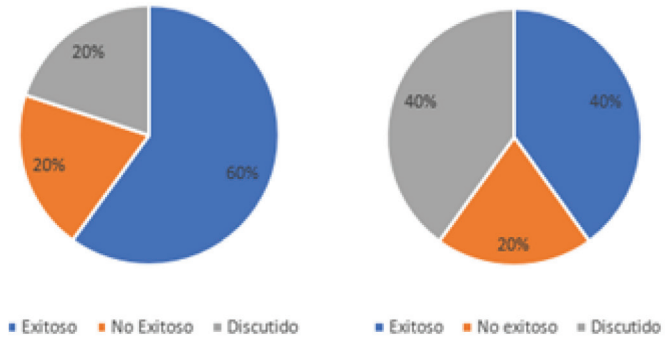


Figura 4 – Porcentaje de éxito en proyectos, proveedores (izquierda) y clientes (derecha)

Tal como se indicó en el apartado anterior, se reconoce que sólo el 80% de los clientes pudieron ser entrevistados, sin embargo, el hecho que el análisis sea considerado a nivel porcentual permite claramente mostrar una tendencia independiente la cantidad.

4.2. Nivel II y Nivel III

En relación con los niveles II (criterios) y III (factores) del framework PSA, podemos decir que las organizaciones encuestadas utilizan criterios, pero no los definen como concepto. Por otro lado, manejan de buena manera la definición de un factor de éxito, pero no establecen mayores relaciones entre los factores que se identifican y los criterios que impactan para lograr el éxito general del proyecto.

Las siguientes tablas presentan la información recopilada desde trece proyectos de las empresas analizadas (etiquetados con una letra cada uno, según se indica en el último

párrafo de la Metodología). Se les pidió a los proveedores y clientes que indicaran cuáles eran los principios o estándares que más consideraban a la hora de catalogar a un proyecto como exitoso o no (concepto de criterio). Los resultados son presentados en la Tabla I en donde se puede apreciar que los criterios más utilizados son el Costo y Alcance, seguidos por el Tiempo y la Calidad. Por otro lado, se puede identificar que los proyectos (c) y (d) presentan pocos criterios al momento de percibir el éxito, mientras que el proyecto (i) es el que tiene la mayor cantidad.

	Criterios	**a**	**b**	**c**	**d**	**e**	**f**	**g**	**h**	**i**	**j**	**k**	**l**	**m**
1	Costo			X	X					X	X	X	X	X
2	Tiempo		X			X				X	X	X	X	
3	Calidad		X			X	X		X			X	X	
4	Alcance	X	X				X	X	X	X	X	X		
5	Usabilidad		X			X			X	X			X	
6	Restricciones legales	X				X								
7	Acuerdos usuario					X	X	X		X				X

Tabla 1 – Criterios más comunes

Con respecto a los factores, la Tabla 2 nos muestran aquellos que fueron los más mencionados. Como se indicó anteriormente, no se establecieron relaciones de cómo los factores impactaban a los criterios, sin embargo, los proyectos c y d nuevamente tienen la menor cantidad, lo cual podría estar relacionado con los pocos criterios identificados en la Tabla 1. Los criterios más nombrados fueron: Involucramiento y Concientización de los Stakeholders, luego de lo cual la Gestión de los Riesgos es considerado importante

4.3. Nivel IV

Para el nivel IV (ciclo de vida), a pesar de que las organizaciones encuestadas no contemplaban formalmente el impacto de los factores de éxito, sobre una determinada etapa del ciclo de vida del proyecto, si tenían noción de que distintos factores de éxito podían ser más relevantes dependiendo de la etapa del proyecto.

5. Conclusiones

El presente trabajo tuvo como objetivo validar que los conceptos que presenta el Framework PSA, junto a sus correspondientes relaciones, son de conocimiento y uso por parte de las empresas en Chile. Se condujo un estudio fenomenológico con cinco organizaciones a fin de llevar a cabo el análisis.

	Factores	a	b	c	d	e	F	g	h	i	j	k	l	m
1	Buena definición del alcance	X	X							X		X		
2	Gestión de recursos		X	X	X						X			X
3	Metodología		X				X	X	X					X
4	Involucramiento SH	X			X	X	X	X	X	X		X	X	X
5	Tiempo de prueba	X					X		X		X	X	X	
6	Trabajo en equipo	X	X			X	X				X	X		X
7	Manejo de expectativas					X	X					X		
8	Project Champion		X				X			X	X	X	X	
9	Concientización SH	X	X			X		X	X	X		X	X	X
10	Mejoramiento de procesos		X			X								
11	Reportabilidad							X				X	X	X
12	Gestión de riesgos	X	X		X			X	X	X			X	X

Tabla 2 – Factores más comunes

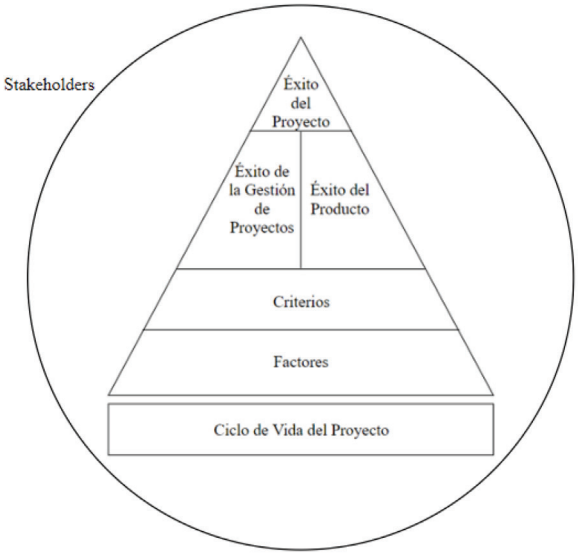


Figura 5 – Representación de los Stakeholders en el Framework PSA.

En base a los resultados obtenidos, se concluye que el éxito de los proyectos es un aspecto que debe seguir siendo desarrollado en las organizaciones encuestadas. Conceptualmente existe noción de los términos que involucran el framework PSA, sin embargo, las relaciones de los conceptos resultaron ser más complejas de visualizar y comprender. A pesar de que al momento de mencionar el éxito las respuestas aludían a métricas representacionales, no se aprecia una declaración formal al momento de definirlas. Al no existir una definición compartida sobre el éxito, las unidades proveedoras difieren de los clientes en su percepción sobre los resultados de los proyectos, lo cual claramente genera cierto grado de insatisfacción por parte de quienes reciben los servicios, pudiendo conducir a la pérdida de confianza a la labor que desempeñan las unidades TI dentro de la organización.

Debido a la importancia de consensuar las distintas visiones que tienen los stakeholders con respecto al éxito, evidenciado en este estudio por medio de los proveedores y clientes, se propone incorporar de manera explícita dicha representación en el framework PSA tal como aparece en la Figura 5, dejando su validación a un trabajo futuro. En la medida que más stakeholders conozcan y se comprometan con los consensos alcanzados en el desarrollo de un proyecto, más cerca se vislumbra el éxito.

Se propone como otro trabajo futuro, enseñar y aplicar el framework PSA en organizaciones y posteriormente analizar su desempeño, primero de una manera cualitativa y después de una manera cuantitativa, con el objetivo de medir el real impacto de este framework en el éxito de los proyectos de SI.

Referencias

Bloch, M., Blumberg, S., & Laartz, J. (2012). Delivering large-scale IT projects on time, on budget, and on value. *McKinsey on Business Technology*(27).

Cecez-Kecmanovic, D., Kautz, K., & Abrahall, R. (2014). Reframing success and failure of information systems: A performative perspective. *MIS Quarterly, 38*(2), 561–588.

Chen, C., Law, C., & Yang, S. (2009). Managing ERP Implementation Failure: A Project Management Perspective. *IEEE Transaction on Engineering Management, 56*(1), 157–170.

Cobb, M. (1996). *Unfinished Voyages - A follow-Up to the CHAOS Report*. The Standish Group International. Obtenido de http://sgraham745.net/uni/Semester%201/Project%20Management/S%20Drive/ProjMan/Stuff/CHAOS%20Unfinished%20Voyages%2096.doc

de Wit, A. (1998). Measurement of project success. *International Journal of Project Management, 6*(3), 164–170.

Doherty, N., Ashurst, C., & Peppard, J. (2012). Factors affecting the successful realisation of benefits from systems development projects: findings from three case studies. *Journal of Information Technology, 27*(1), 1–16.

Drury-Grogan, M. (2014). Performance on agile teams: Relating iteration objectives and critical decisions to project management success factors. *Information and Software Technology, 27*(9), 506–515.

Hastie, S., & Wojewoda, S. (4 de October de 2015). *Standish Group 2015 Chaos Report - Q&A with Jennifer Lynch*. Recuperado el 11 de December de 2017, de https://www.infoq.com/articles/standish-chaos-2015

Howsawi, E., Eager, D., & Bagia, R. (2011). Understanding project success: The four-level project success framework. *2011 IEEE International Conference on Industrial Engineering and Engineering Management* (págs. 620-624). Singapore: IEEE.

Ika, L. (2009). Project success as a topic in project management journals. *Project Management Journal, 40*(4), 6-19.

Irvine, R., & Hall, H. (2015). Factors, frameworks and theory: a review of the information systems literature on success factors in project management. *Information Research, 20*(3).

Leedy, P. D., & Ormrod, J. E. (2012). *Practical Research: Planning and Design*. Pearson.

Mahmud, I., Ramayah, T., & Kurnia, S. (2017). To use or not to use: Modelling end user grumbling as user resistance in pre-implementation stage of enterprise resource planning system. *Information Systems, 69*, 164–179.

Project Management Institute. (2017). *A guide to the project management body of knowledge - PMBOK Guide*. Newtown Square, Pennsylvania: Project Management Institute Inc.

Quelopana, A., & Flores, V. (2018). A framework for IS project success analysis: A proposal. *Iberian Conference on Information Systems and Technologies, CISTI*, (págs. 1-5).

The Standish Group International. (2014). *Project Smart*. Recuperado el 16 de May de 2017, de https://www.projectsmart.co.uk/white-papers/chaos-report.pdf

Thomas, G., & Fernández, W. (2008). Success in IT projects: A matter of definition? *International Journal of Project Management, 26*, 733–742.

Todorović, M. L., Petrović, D. Č., Mihić, M. M., Obradović, V. L., & Sergey D. Bushuyev. (2015). Project success analysis framework: A knowledge-based approach in project management. *International Journal of Project Management*, 772–783.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 22/09/2018
Aceitação/Acceptance: 25/11/2018

Esquema para la actualización de Ontologías de Competencias en base al Procesamiento del Lenguaje Natural y la Minería Semántica

Alexandra González-Eras[1], Jose Aguilar[2]

acgonzalez@utpl.edu.ec, aguilar@ula.ve

[1] Departamento de Ciencias de la Computación y Electrónica - Universidad Técnica Particular de Loja, Cdla. Universitaria San Cayetano Alto, 1101608, Loja, Ecuador.

[2] Centro de Estudios en Microelectrónica y Sistemas Distribuidos - Universidad de Los Andes, 5101, Mérida, Venezuela.

Pages: 433–447

Resumen: Este trabajo presenta un esquema para la actualización de competencias profesionales y académicas desde perfiles obtenidos de la Web, basado en ontologías. Primero se reconocen los términos mediante patrones lingüísticos. Después, los términos se alinean con tesauros, mediante medidas de relevancia y similitud léxica y semántica. Finalmente, los términos alineados pueblan la ontología. El caso de estudio es Ciencias de la Computación, para lo cual se usan los tesauros DISCO y BLOOM. La calidad de la ontología poblada se evalúa usando las medidas de completitud y robustez, las cuales reflejan altos valores en los experimentos realizados.

Palabras-clave: ontologías de competencias, similitud léxica, similitud semántica, relevancia textual, tesauros, DISCO II, Bloom.

A Scheme for updating the Competency Ontologies based on Natural Language Processing and Semantic Mining

Abstract: This work presents a scheme for updating the professional and academic competences from profiles obtained from the Web, based on ontologies. First are recognized the terms through linguistic patterns. Then, the terms are aligned with thesauri, through measures of relevance and lexical and semantic similarity. Finally, the aligned terms populate the ontology. The case study is in Computer Science, using the DISCO and BLOOM thesauri. The quality of the ontology is evaluated using the measures of completeness and robustness, which reflect high values in the experiments carried out.

Keywords*:* competency ontologies, lexical similarity, semantic similarity, textual relevance, thesauri, DISC II, Bloom.

1. Introducción

Actualmente, la gestión de competencias tiene mucho por hacer en cuanto al análisis comparativo automático de las competencias entre las ofertas académicas y las ofertas de trabajo. Por otro lado, el proceso de enriquecimiento de las competencias en los perfiles académicos y profesionales con información de la Web, conlleva a la definición de patrones y al alineamiento de tesauros, taxonomías y ontologías, entre otras cosas. A su vez, el uso de ontologías de competencias contribuye al proceso de comparación de competencias de ofertas de empleos y académicas de diferentes contextos (Hassan, 2012), como vocabulario para la extracción de competencias desde diferentes fuentes (Sateli, 2017), (Paquette, 2012), (Gluga, 2013) y, además, facilita sus procesos de actualización (Montuschi, 2015).

Este trabajo presenta un esquema para la actualización/enriquecimiento de competencias en los perfiles laborales y académicos, basado en la combinación de técnicas de Procesamiento del Lenguaje Natural (PLN), Minería Semántica y Aprendizaje Ontológico, las cuales permiten crear e integrar modelos semánticos, para representar y usar las competencias. En los siguientes apartados se presenta la propuesta, que se basa en la extracción y clasificación de instancias de competencias basada en patrones, mediante técnicas de recuperación de información, tesauros y algoritmos de aprendizaje; y, finalmente, la población de ontologías de competencias con las instancias extraídas. Para ello, se propone una arquitectura que, entre otras cosas, permite la actualización de ontologías de competencias, la cual resuelve el problema de ambigüedad lingüística y semántica que se presenta en la búsqueda de competencias desde información en español.

2. Arquitectura

La Fig. 1 presenta la arquitectura propuesta, basada en (González-Eras, 2015).

2.1. Caracterización de competencias.

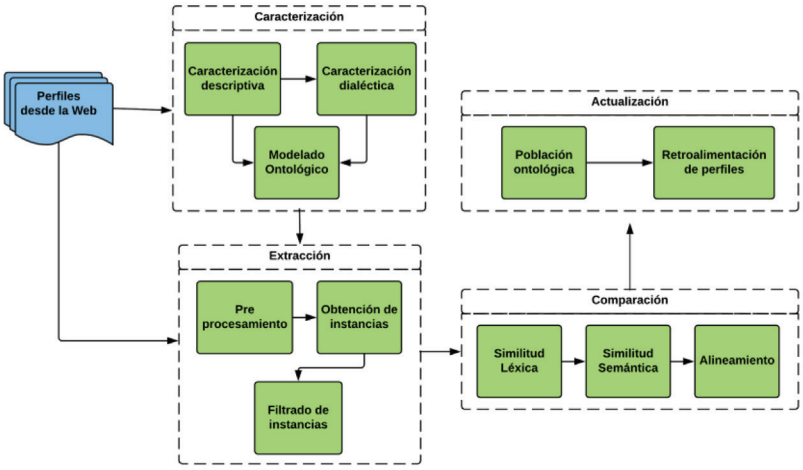


Figura 1 – Arquitectura del esquema de actualización de ontologías de competencias

Esta fase comienza con un proceso de revisión de los perfiles profesionales a nivel lingüístico, para lo cual se usan técnicas de PLN, para determinar la estructura lingüística de las competencias y sus componentes (conocimiento y habilidad). Con estas estructuras, se definen patrones que se usan para extraer términos de los documentos recolectados desde la Web. La Tabla 1 presenta los patrones de competencia, conocimiento y habilidad en lógica descriptiva (González-Eras, 2015). Con estos patrones, se extraen desde los documentos de la web, los posibles términos para poblar la ontología de competencias (Rodríguez, 2014).

Descripción	Lógica de Primer Orden	Ejemplo
La frase nominal (NP) puede ser una competencia (Co), un tópico de conocimiento (C) o una habilidad (H)	1. $\forall NP, NC \ \ NP \equiv NC$ 2. $\forall NP, SP, NC \ \ NP \equiv NC \sqcap SP \sqcap NC$ 3. $\forall NC, SP, AQ, NP \ \ NP \equiv NC \sqcap SP \sqcap NC \sqcap AQ$ 4. $\forall NC, SP, NP \ \ NP \equiv NC \sqcap SP \sqcap NC \sqcap SP \sqcap NC$	Administración. Experto en Java. Desarrollo de software. Conocimiento sistemas operativos. Conocimiento base datos. Administración base datos.
La frase verbal (VP) puede ser una competencia o una Habilidad	5. $\forall VMN, NC, SP, VP \ VP \equiv VMN \sqcap NC \sqcap SP \sqcap NC$	Administrar bases datos.

Tabla 1 – Patrones para el reconocimiento de términos de competencia y sus componentes

2.2. Extracción de competencias

A. Obtención de términos. Detecta los términos de conocimiento y habilidad en los documentos recolectados de la Web. Esta fase pre-procesa los documentos usando técnicas de PLN como Tokenización y Análisis Morfo-léxico (ver Tabla 2). El algoritmo usa técnicas de PLN para el análisis morfo-léxico (Faria, 2014); y, después, aplica los patrones propuestos en la Tabla 1 para extraer los términos de conocimiento, habilidad o competencia (ver los siguientes enunciados).

Inicio
Entrada Colección de perfiles profesionales P Lista de Patrones descriptivos PD según la Tabla 1 Procedimiento 1. Para toda la colección de perfiles P 1.1. Realizar el pre-procesamiento del perfil(id_i) según el Enunciado 1 1.2. Realizar el Análisis Morfo-léxico del perfil(id_i) según el Enunciado 2 1.3. Para todo el perfil(id_i) 1.3.1. Para toda la lista de patrones Pd 1.3.1.1. Obtener el término C_j según el patrón(Pd_k) 1.3.1.2. Agregar C_j al dataset de términos DI según el Enunciado 3

Tabla 2 – Pseudocódigo para la obtención de términos

Enunciado 1. El pre-procesamiento consiste en el tratamiento inicial de los documentos, es decir, en realizar una limpieza de etiquetas HTML como encabezados, números, fechas, metadata, dejando solamente aquellos textos que se encuentran dentro de etiquetas tipo . Luego, los extractos resultantes se dividen en oraciones y, después, en palabras (tokens); las que son normalizadas (manejo de mayúsculas, lematización, etc.) (Manning, 2014).

Enunciado 2. El análisis morfo-léxico consiste en la asignación de etiquetas lingüísticas a cada uno de los tokens, para establecer su categoría lingüística. Con esta entrada, se aplica el conjunto de patrones de la Tabla 1, para encontrar los términos de competencia, conocimiento y habilidad (González-Eras, 2015).

Enunciado 3. Sea DI={(c,id_1,C_1,Pd_1),⋯, (c,id_n,C_n,Pd_n)} un dataset de términos válidos, donde c indica el tipo de perfil (1: académico, 2: laboral), id_i el identificador del perfil, C_j el término que fue extraído, y Pd_k el patrón descriptivo del término.

B. Filtrado de términos. realiza el análisis de la relevancia de cada uno de los términos reconocidos, en el conjunto de documentos (ver siguientes conceptos):

Definición 1. El valor de relevancia de un término consiste en la posición del mismo dentro de la colección de términos analizados. En concreto, el valor de relevancia de un término C_j del Perfil id_i, está dado por la ecuación (1).

$$Score(id_i, C_j) = \sum_{j=1}^{n} IDF(C_j) \cdot \frac{f(C_j, id_i) \cdot (k_1 + 1)}{f(C_j, id_i) + k_1 \cdot (1 - b + b \cdot \frac{|D|}{avgdl}} \quad (1)$$

En donde, $f(C_j,id_i)$ es la frecuencia de aparición del término C_j en el Perfil id_i; |D| es el número de términos en el Perfil id_i); *avgdl* es longitud media de los perfiles que conforman la colección; k_1 y b son parámetros de ajuste; $IDF(C_j)$ es el peso dado a la instancia C_j en la colección de perfiles; n es el número de perfiles de la colección.

Definición 2. La frecuencia de aparición de un término consiste en la cantidad de veces que este término se presenta en el perfil, considerando palabras iguales o semejantes, según la medida de distancia de la Definición 7, y el umbral del Enunciado 4. La ecuación (2) explica el cálculo de $f(C_j,id_i)$:

$$f(C_j, id_i) = \sum_{j=1}^{m} C_j \quad (2)$$

En donde, m representa el total de términos C encontrados en el Perfil id_i.

Definición 3. El peso de un término C_j $IDF(C_j)$ está dado por la frecuencia inversa del mismo en relación a la colección de perfiles (ver ecuación (3)):

$$IDF(C_j) = log \frac{n - n'(C_j) + \delta}{n'(C_j) + \delta} \quad (3)$$

Donde, $n'(C_j)$ es el número de perfiles que contienen el término C_j, y δ es un parámetro de ajuste al peso dado a un término.

Enunciado 4. Se establece un rango U_R entre 0,3 y 1, que corresponde al valor de relevancia de C_j para considerar que es relevante en la colección de perfiles. El valor de U_R se definió en base al trabajo realizado en (Dijkman, 2011).

Enunciado 5. Sea DI'={(c,id_1,C_1,Pd_1,r),⋯, (c,id_n,C_n,Pd_n,r)} un dataset de términos válidos, donde c indica el tipo de perfil (1: académico, 2: laboral), id_i el identificador del perfil, C_n el término C_j que fue extraído, Pd_i el patrón descriptivo del término y r el valor de relevancia del término en la colección de perfiles, según el Enunciado 4.

Con esas definiciones, se define el siguiente macro algoritmo (ver tabla 3).

Inicio
Entrada Dataset de términos DI según el Enunciado 3 Procedimiento 1. Para todo DI 1.1. Obtener el término (C_j) 1.2. Calcular la relevancia del término C_j *Score(id_i,C_j)* según la Definición 1 1.3. Si *Score(id_i,C_j)* es mayor que el umbral de relevancia U_R según el Enunciado 4 1.3.1. Descartar C_j. 1.4. Si no 1.4.1. Agregar C_j al dataset de términos relevantes DI', según el Enunciado 5

Tabla 3 – Pseudocódigo del Filtrado de términos

Es así que el proceso de filtrado de términos usa la frecuencia de términos iguales y semejantes, para establecer la relevancia de cada uno en el conjunto de perfiles, siendo seleccionados para la fase de alineamiento con tesauros, los términos cuya relevancia se encuentra en el rango del umbral U_R.

2.3. Comparación con tesauros de habilidades y conocimiento

La comparación se realiza alineando los términos relevantes contra los términos de tesauros. Para ello, se usan dos medidas de similitud (González-Eras, 2018):

A. Comparación léxica con los tesauros DISCO y BLOOM

Similitud Léxica. se establece una similitud léxica entre los términos relevantes de los perfiles, y cada término perteneciente a los tesauros DISCO y BLOOM (ver Tabla 4). Para ello, se consideran dos medidas, la primera llamada *Dislex(C_j,C')*, que usa la medida Levenshtein para determinar la distancia de edición entre el término C_j y el término C' (Levenshtein, 1966); y la segunda llamada *Simlex(C_j,C')*, que usa el coeficiente de Sorensen para determinar cercanías según la similitud de sus pares de caracteres (Alqadah, 2011).

Enunciado 6. Sea D={(C'_1,n_1),⋯, (C'_n,n_n)} un conjunto de términos de conocimiento organizados en jerarquía, tal que un término C' pertenece a un nivel n_k del tesauro.

Definición 4. Un término C' tiene asociado un conjunto de frases F' descritas según ec. (4); una frase F'_i es una colección de términos descritos según ec. (5), un término C'_i

puede tener cero o más frases de competencia, y cada frase de competencia F'_i contiene uno o más términos de conocimiento C'_i o habilidad H'_i.

$$C'_i = \{F'_i \dots F'_n\} \tag{4}$$

$$F' = \{(H'_i, C'_i) \dots (H'_n, C'_n)\} \tag{5}$$

Enunciado 7. Sea $B=\{(TH_1,n_1),\cdots,(TH_n,n_n)\}$ un conjunto de términos de habilidad organizados jerárquicamente, en donde un término TH pertenece a un nivel n_k dado del tesauro, y tiene asociado un conjunto de verbos V relacionados, descritos según ec. (6).

$$TH_i = \{V_i \dots V_n\} \tag{6}$$

Definición 5. Un verbo V_i tiene una colección de sinónimos S (ver ec. (7)).

$$V'_i = \{S_i, \dots S_n\} \tag{7}$$

Definición 6. La distancia de edición entre dos términos C_j y C'_i viene dada por el número de cambios de caracteres que deben realizarse para que el término C_j se convierta en el término C'_i (ver ec. (8)) (Levenshtein, 1966). El valor de esa medida es máximo cuando el número de cambios es cero (C_j y C'_i son iguales), y es min en el caso contrario.

$$Dis_{lex}(C_j, C'_i) = \begin{cases} \max(C_j, C'_i) & Si\ Dis_{lex}(C_j, C'_i) = 0 \\ \min(C_j, C'_i) & Si\ Dis_{lex}(C_j, C'_i) > 0 \end{cases} \tag{8}$$

Definición 7. La similitud léxica entre dos términos C_i e C'_i es dos veces el número de pares de caracteres que son comunes a ambos términos, divididos por la suma del número de pares de caracteres en los dos términos (9).

$$Sim_{lex}(C_j, C'_i) = \frac{2 \times |pares\ (C_j) \cap pares\ (C'_i)|}{|pares\ (C_j) + pares\ (C'_i)|} \tag{9}$$

Enunciado 8. Se establece un umbral de distancia U_D igual a 4, como la mínima distancia de edición que puede existir entre C_j y C'_i para considerar que tienen una similitud. El valor de U_D de 4 está basado en el trabajo de (Dijkman 2011).

Enunciado 9. Se establece un umbral U_L igual a 0.4, como la mínima similitud léxica que puede existir entre C_j y C'_i para considerar que tienen una similitud. El valor de U_L de 0.4 está basado en el trabajo de (Dijkman 2011).

Enunciado 10. Sea $P'=\{(c,id_1,C_1,H_1,n_1),\cdots,(c,id_n,C_n,H_n,n_n)\}$ un dataset de términos válidos, en el que c indica el tipo de perfil (1: académico, 2: laboral), id_i es el identificador

del perfil, C_n representa el término de conocimiento C_j de los perfiles, H_i es la habilidad relacionada al término, y n_k es el nivel k del árbol D donde se halló el máximo valor de similitud del término, que cumple con los umbrales definidos en los enunciados 8 y 9.

Con esas definiciones, se define el siguiente macro algoritmo (ver tabla 4).

Inicio

Entrada
Dataset de términos relevantes DI' según el Enunciado 5
Árbol del tesauro DISCO D con términos de conocimiento según el Enunciado 6
Árbol del tesauro BLOOM B con términos de habilidad según el Enunciado 7
Procedimiento
1. Para todo el Dataset DI'
1.1. Obtener el término de conocimiento (C_j)
1.2. Para todo el Árbol de términos de conocimiento D
1.2.1. Obtener el término de conocimiento del Árbol D (C'_i)
1.2.2. Calcular la distancia de edición entre C_j y C'_i *Dislex*(C_j,C'_i) según Definición 6
1.2.3. Si *Dislex*(C_j,C'_i) es mayor al umbral de distancia U_D según Enunciado 8
1.2.3.1. Calcular la similitud léxica entre los tópicos de conocimiento C_j y C'_i *Simlex*(C_j,C'_i) según Definición 7
1.2.3.2. Si *Simlex*(C_j,C'_i) es menor al umbral léxico U_L según Enunciado 9
1.2.3.2.1. Descartar C_j
1.2.3.3. Si no
1.2.3.3.1. Agregar el término C_j al Dataset de términos P' según Enunciado 10
1.2.4. Sino
1.2.4.1. Descartar C_j

Tabla 4 – Cálculo de la Similitud Léxica

B. Comparación estructural con el tesauro DISCO

Similitud semántica. Se establece una similitud semántica entre los términos relevantes mediante la alineación de los pares de términos seleccionados por la similitud léxica contra el Tesauro DISCO II, basado en el esquema propuesto en (González-Eras 2015), (ver Tabla 5), en el cual, para cada par de términos C_j y C'_i se verifica la similitud de sus ancestros, hermanos e hijos en el tesauro (Mendonca 2015),(Aguilar, 2001). Luego, para los términos con una alta similitud semántica, se obtiene la similitud de sus términos de habilidad contra el tesauro BLOOM.

Definición 8. La similitud semántica de dos términos C_j y C'_i está dada por la sumatoria de las similitudes de ancestros, hermanos e hijos del término C_j, divididas por 3. Entonces, para cada par de términos se obtiene un valor de similitud que se encuentra en el rango de 0 a 1, donde 0 representa no similitud y 1 representa alta similitud (ver ec. (10, 11, 12, 13)).

$$Sim_{sem}(C_j, C'_i) = \frac{SA(C_j, C'_i) + SS(C_j, C'_i) + SD(C_j, C'_i)}{3} \quad (10)$$

$$SA(C_j, C'_i) = \frac{1}{n}\sum_{i=1}^{n} \max(Sim\left(Anc_i(C_j), Anc_1(C')\right), \ldots, Sim\left(Anc_i(C_j), Anc_n(C')\right) \quad (11)$$

$$SS(C_j, C'_i) = \frac{1}{n}\sum_{i=1}^{n} \max(Sim(Sin_i, Sin'_1), \ldots, Sim(Sin_i, Sin'_n) \quad (12)$$

$$SD(C_j, C'_i) = \frac{1}{n}\sum_{i=1}^{n} \max(Sim(Des_i, Des'_1), \ldots, Sim(Des_i, Des'_n) \quad (13)$$

En donde: $Anc_i(C_j)$: ancestro i del término C_j; $Anc_j(C')$: ancestro j del término C'; $Sim(Anc_i(C_j),Anc_j(C'))$: medida de similitud entre ancestros de los términos C_j y C'; Sin_i: hermano i del término C_j; Sin'_j: hermano j del término C'; $Sim(Sin_i,Sin'_j)$: medida de similitud entre los hermanos de los términos C_j y C'; Des_i: hijo i del término C_j; Des'_j: hijo j del término C'; $Sim(Des_i,Des'_j)$: medida de similitud entre los hijos de los términos C y C'; n: máximo nivel de similitud léxica entre C_j y C'.

Enunciado 11. Se considera que m_d representa al término raíz del subárbol de DISCO, en donde los términos C_j y C' alcanzan un valor de similitud máxima.

Enunciado 12. Sea $TC=\{(c,id_1,C_1,Ms_1,m_d),\cdots,(c,id_n,C_n,Ms_n,m_d)\}$ un dataset de términos de conocimiento válido, donde c indica el tipo de perfil (1 académico, 2 laboral), id_i el identificador del perfil, C_n el término de conocimiento C_j de los perfiles, M_s la máxima medida de similitud semántica $Simsem(C_j,C'_i)$ según ec. (10), y m_d el término raíz del subárbol del tesauro DISCO según Enunciado 11.

C. Comparación estructural con la taxonomía de Bloom

El proceso de comparación se realiza con los términos de habilidad del tesauro BLOOM (González-Eras, 2018), lo cual implica identificar el nivel cognitivo que es raíz del subárbol en donde se localizan los términos comparados, para lo cual se plantean los siguientes enunciados:

Enunciado 13. m_b representa el término raíz del subárbol de BLOOM, en donde los términos habilidad H_j y H'_i alcanzan un valor de similitud máxima.

Enunciado 14. $TH=\{(c,id_1,H_1,m_b),\cdots,(c,id_n,H_n,m_b)\}$ es un dataset de términos de habilidad válido, donde c indica el tipo de perfil (1 académico, 2 laboral), id_i el identificador del perfil, H_n el término de habilidad de los perfiles H_j , y m_b el término raíz del subárbol del tesauro BLOOM según Enunciado 13.

Con esas definiciones, se define el siguiente macro algoritmo (ver tabla 5).

2.4. Actualización

Esta fase consiste en la población de la ontología de competencias OC, que tiene la estructura de la Figura 2 (clases, propiedades y relaciones). Algunas de las clases principales son: categoría gramatical (clase para el Análisis Morfo-léxico); cobertura de conocimiento (tópicos del nivel superior del tesauro DISCO); competencia (dominio académico, laboral y patrones de competencia); dominio cognitivo (tópicos del nivel superior de BLOOM); y patrón (axiomas de la Tabla 1).

Inicio

Entrada
Dataset de tópicos de conocimiento P' según Enunciado 10
Árbol del tesauro DISCO D de tópicos de conocimiento según el Enunciado 6
Árbol del tesauro BLOOM B de tópicos de habilidad según el Enunciado 7
Procedimiento
1. Para todo el Dataset P'
1.1. Obtener el término de conocimiento (C_j)
1.2. Para todo el Arbol D
1.2.1. Calcular la similitud semántica entre C_j y C'_i *Simsem(C_j,C'_i)* según Definición 8
1.2.2. Obtener el término raíz del subárbol del tesauro DISCO m_d de máxima *Simsem(C_j,C'_i)* según Enunciado 11
1.2.3. Agregar al dataset de términos de conocimiento TC C_j máx *Simsem(C_j,C'_i)*, m_d, id y c según Enunciado 12
1.3. Fin Para
2. Para todo el Dataset P'
2.1. Obtener el término de habilidad (H_j)
2.2. Para todo el Arbol B
2.2.1. Obtener el término raíz del subárbol del tesauro BLOOM m_b donde se encuentran los términos de habilidad H_j,H'_i) según Enunciado 14
2.2.2. Agregar al dataset de términos de habilidad TH H_j, m_b, id_i y c según Enunciado 14

Tabla 5 – Cálculo de la Similitud Semántica

En esta fase se realiza la población ontológica, la cual considera dos momentos, el primero, cuando se realiza la obtención de términos desde los documentos siguiendo los patrones lingüísticos de la Tabla 1, y el segundo durante el proceso de alineamiento de términos con los tesauros DISCO II y BLOOM. En esta fase se realizan dos cosas: i) poblar la ontología y ii) actualizar los elementos de competencia. Así, el modelo ontológico contiene una parte libre de contexto correspondiente a los patrones lingüísticos, que puede ser aplicado en documentos de cualquier dominio de conocimiento; y una parte dependiente del contexto, que corresponde a la actualización de las competencias según el Dominio_Cognitivo (por ejemplo, el tesauro BLOOM) y la Cobertura_Conocimiento (por ejemplo, el tesauro DISCO II) donde se usará la ontología. Por otro lado, la Tabla 1 presenta ejemplos de los axiomas utilizados en el proceso de población ontológica, en Lógica de Primer Orden.

Para evaluar la calidad de la ontología actualizada, se usan las siguientes medidas:

Completitud. Una ontología de competencia OC es considerada completa con referencia al perfil profesional o académico id_i, si contiene todos los términos relevantes extraídos del perfil id_i. La Ecuación (14) presenta la definición de Completitud.

$$Completitud(OC, id_i) = \frac{\sum_{i=1}^{m} Trelevantes(id_i) \cap Terminos(OC)}{\sum_{i=1}^{m} Trelevantes(id_i)} \quad (14)$$

En donde: *Trelevantes(id_i)* son los términos cuyo *Score(id_i,C_j)* se encuentra en el rango definido en el U_R (entre 0,3 y 1) y *Terminos(OC)* son los términos en la OC.

Robustez. Una ontología de competencia OC se considera robusta con referencia al conjunto de perfiles profesionales y académico, si sus términos C_j son relevantes para el perfil id_i.

$$Robustez(OC, id_i) = \frac{\sum_{i=1\&Cj \in Terminos(OC)}^{m} Score(idi, Cj)}{|Terminos(OC)|} \quad (15)$$

En donde, si ese promedio es menor a 0.3, implica que se han usado términos no relevantes para poblar a OC.

Para establecer un análisis cualitativo de los resultados de calidad, se usa la escala de valoración propuesta en (Mendonca, 2015), que considera 3 criterios: Alto (valor mayor que 0.5), Medio (valor entre 0.3 y 0.5) y Bajo (valor inferior a 0.3).

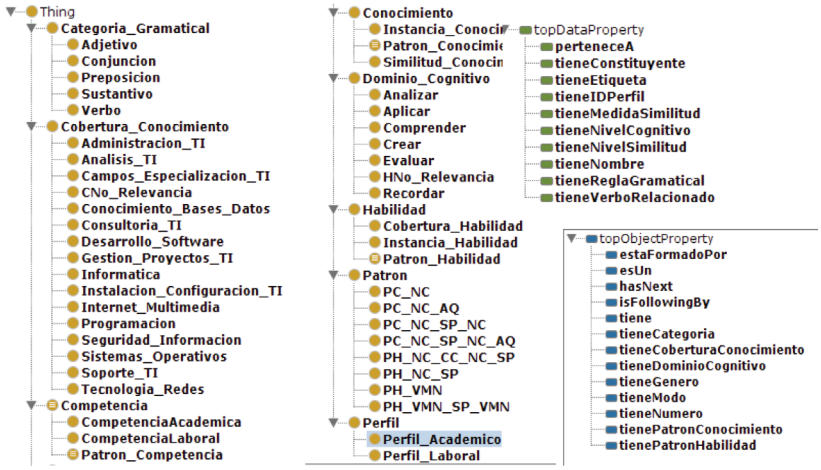


Figura 2 – Clases y propiedades de la ontología de competencias OC

3. Experimentación

Para el experimento, se tomó como entrada 35 perfiles profesionales en español: 20 ofertas de carrera obtenidas de portales universitarios (id_1,...,id_{20}), y 15 ofertas laborales obtenidas de portales de empleo en Internet (id_{21},..id_{35}), donde se usó esta nomenclatura para proteger la identidad de los autores de los perfiles. De cada perfil, fueron seleccionados extractos de texto que se encontraban bajo apartados como: descripción, objetivos, competencias, para respetar la normativa de propiedad intelectual de los sitios fuente de los documentos. Además, se utilizaron dos tesauros para la fase de alineamiento: DISCO (Tc_1,...,Tc_{15}) y BLOOM (Th_1,...,Th_6), los cuales se modelan en las clases "cobertura de conocimiento" y "dominio cognitivo" de la ontología (ver Fig. 2 y (González-Eras, 2018), para más detalles).

3.1. Fase de extracción

En esta fase se buscan los términos de conocimiento, habilidad y competencia. Además, se realiza el filtrado de términos. La Tabla 6 presenta casos de filtrado de términos,

para k_1=1,2 b=0,75 y δ=0,5. Se observa que los términos "software", "aplicación computacional" y "hardware", presentan buenas medidas de relevancia.

Término	$n(C_j)$	$IDF(C_j)$	id_1	id_2	id_3
base de datos	14	0,3087	0	0,5658	0
Hardware	9	0,5371	0,4371	0	0,6559
Software	11	0,4357	0,3545	0	0,5321
sistema operativo	13	0,3484	0,2835	0	0,4255
aplicación computacional	12	0,3906	0,3178	0,7159	0
actividad de especificación	9	0,5371	0,437	0	0
aplicación informática	18	0,1679	0	0,3078	0,2051
proyecto informático	16	0,2352	0	0	0,2872
lenguaje sql	5	0,8195	0	0	0
desarrollo de software	13	0,3485	0,2835	0	0,4255
Sistema	18	0,168	0,1367	0	0,2051
proyecto de Sistema	18	0,168	0,1367	0	0,2051

Tabla 6 – Cálculo de algunos filtrados de términos

3.2. Fase de comparación

En esta fase se realiza el alineamiento de los términos obtenidas en la fase de extracción, con los tesauros DISCO II y BLOOM. La similitud léxica es calculada en la Tabla 7, y la semántica en la Tabla 8, para los tesauros DISCO II y BLOOM.

Términos	Similitud Léxica
base de datos	0,233
Software	0,379
proyecto informático	0,640
aplicación informática	0,717
lenguaje sql	0,397
Hardware	0,425
sistema operativo	0,817
actividad de especificación	1,007
aplicación computacional	1,867

Tabla 7 – Cálculo de la Similitud Léxica

En la Tabla 8 se observa que la medida calculada entre los términos C y C' (Definición 8), permite identificar el término raíz del subárbol de DISCO, con el cual los términos de los perfiles tienen alta similitud. Por ejemplo, "aplicación computacional", "proyecto informático", "lenguaje sql", "lenguaje java" y "java" tienen una similitud con el término

del tesauro "Programación", perteneciendo al mismo contexto, porque los valores de Ms se encuentran cerca del umbral $U_s > 0,45$. En cuanto a habilidad, se aprecia, que los términos "administrador", "dirigir", "interacción" y "planificar" tienen una relación semántica con el término del subárbol del tesauro BLOOM "Síntesis" (relación de sinonimia).

id	C	M_s	m_d	H	M_s	m_b
1	sistema de comunicación	0.67	Instalación y configuración TI	administrador	0,7	Síntesis
1	actividad de especificación	0.76	Desarrollo de software	dirigir	0,8	Síntesis
2	aplicación computacional	0.59	Programación	desarrollo	0,8	Aplicación
2	proyecto informático	0.53	Programación	planificar	1	Síntesis
21	base de dato	0.70	Conocimientos de bases de datos	interacción	0.7	Síntesis
21	lenguaje sql	0.61	Programación	programar	0,9	Aplicación
21	lenguaje java	0.58	Programación	conocer	1	Evaluación
21	Java	1	Programación	programar	0,9	Aplicación

Tabla 8 – Cálculo de la similitud semántica en términos de conocimiento (C) y habilidad (H) con DISCO II y BLOOM, respectivamente

3.3. Fase de actualización

En esta fase, se realiza la población ontológica del modelo de la Figura 2. La Tabla 9 presenta los resultados generales de la calidad de la ontología en cuanto a su completitud y robustez. En general, para el 71,44% de perfiles, la ontología tiene una completitud Alta, y, para el 22,85% de los perfiles una completitud Media; la completitud de nivel Bajo corresponde a dos perfiles (5,71%). Eso implica que, en promedio, todos los términos relevantes son usados para poblar OC, un resultado excelente que indica la utilidad del proceso. En cuanto a la robustez, el 28,57% de los perfiles tienen una relevancia Alta, mientras que el 68,57% de los perfiles tienen una relevancia media. En este caso, prácticamente todos los componentes con los que se pobló el OC son relevantes para los perfiles usados.

4. Comparación con otros trabajos

La Tabla 10 presenta una comparación de la propuesta con trabajos anteriores según técnicas de NLP usadas, Representación de conocimiento y habilidad, Medidas de similitud, información extra usada (tesauros, etc.) y métricas de Validación del modelo ontológico. Se puede observar, que existen coincidencias entre los trabajos y nuestra propuesta en el uso de medidas de similitud léxicas, pero la nuestra es la única que usa similitud semántica. En cuanto al uso de información extra, muchos usan tesauros, lexicones o enlazado de datos. Así también, en las técnicas de NLP que se emplean, la mayoría usa técnicas de Reconocimiento de Entidades Nombradas (NER), y nuestro trabajo es el único basado en patrones lingüísticos. En cuanto a la validación ontológica,

nuestra propuesta es la única que usa los criterios de Completitud y Robustez para determinar la utilidad y calidad del proceso de extracción de términos, y su uso para poblar la OC, en lugar de la métrica de Precisión. Otra diferencia radica en el idioma de las fuentes de los datos, donde nuestra propuesta y (Rodríguez, 2014) son los únicos para el español. Finalmente, la mayoría usa ontologías, y nuestra propuesta es la única que combina documentos de la web con perfiles.

Perfil	Completitud	Valor	Perfil	Robustez	Valor
id_{2}	1	Alto	id24	0,94	Alto
id_{14}	0,93		id21	0,66	
...			...		
id_{20}	0,67		id1	0,53	
id_{3}	0,51		...		
...			id33	0,5	
id_{31}	0,51		id13	0,49	Medio
id_{35}	0,48	Medio	id34	0,48	
id_{28}	0,46		...		
id_{1}	0,42		Id3	0,42	
...			id2	0,39	
id_{24}	0,26		...		
id_{16}	0,19	Bajo	id6	0,24	
id_{23}	0,14		id23	0	Bajo

Tabla 9 – Cálculo de la Completitud y Robustez del modelo ontológico OC

Trabajo	**NLP**	**Representa conocimiento y habilidad**	**Medida de similitud**	**Tesauros extra**	**Fuente de datos**	**Validación**
(Malzahn, 2013)	Patrones	Competencias como entidades	Levenshtein	WordNet	Ofertas laborales	Experto
(Sateli, 2017)	NER	Competencias como entidades	Levenshtein WordNet	LinkedData ontología competencia	Perfiles	Precisión y relevancia
(Faria, 2014)	NER	Entidades y acciones	Léxicas WordNet	Ontología competencia	Corpus	Precisión Recall
(Rodríguez, 2014)	NER	Competencias como entidades	Semántica	Lexicón y Onomasticón	Texto de la Web	Precisión y relevancia
(Gatteschi, 2012)	NER	Ontología de perfiles de competencia	Léxicas	WordNet	Texto de la Web	Precisión
Nuestra propuesta	Patrones	Competencias conocimiento y habilidad	Léxicas Semántica	DISCO II BLOOM	Ofertas laborales y carrera	Completitud Robustez relevancia

Tabla 10 – Comparación de la propuesta con otros trabajos

5. Conclusiones y trabajos futuros

Este trabajo presenta un esquema para la actualización de competencias profesionales y académicos desde perfiles obtenidos de la Web. Las competencias son modeladas usando ontologías, para descubrir sus elementos de habilidad y conocimiento mediante patrones lingüísticos, que, al alinearse con tesauros mediante medidas de similitud y relevancia, definen sus importancias. Los resultados obtenidos permiten poblar el modelo ontológico con los tópicos de conocimiento y habilidad de los tesauros DISCO II y BLOOM, que se encuentran en los documentos de los perfiles obtenidos en la Web.

La diferencia con nuestro trabajo radica en las técnicas de PLN, medidas de similitud y tesauros considerados, además de las medidas de la calidad del modelo ontológico. Nuestra propuesta puede extraer y poblar una ontología de competencias de cualquier dominio, sobre textos en español en contextos académicos y laborales, contribuyendo así con sistemas que generen perfiles académicos de carreras según necesidades laborales, o determinen las nuevas competencias profesionales a considerar en un ámbito de trabajo. Por otro lado, la calidad del proceso de extracción es determinada por las métricas de calidad.

Los futuros trabajos deben definir el uso de otras medidas de similitud, tesauros y márgenes de relevancia en cada una de las fases de la arquitectura. Además, deben diseñar una métrica de precisión que complemente las de calidad del modelo ontológico usadas en este trabajo. Finalmente, se debe replicar el modelo propuesto con otras bases de conocimiento en el área de Computación, como es el caso de ACM, y en otras áreas de conocimiento (ciencias jurídicas, medicina, etc.).

Referencias

Aguilar, J. (2001). A general ant colony model to solve combinatorial optimization problems. Revista Colombiana de Computación-RCC, 2(1).

Alqadah, F., & Bhatnagar, R. (2011). Similarity measures in formal concept analysis. Annals of Mathematics and Artificial Intelligence, 61(3), 245–256.

Dijkman, R., Dumas, M., Van Dongen, B., Käärik, R., & Mendling, J. (2011). Similarity of business process models: Metrics and evaluation. Information Systems, 36(2), 498–516.

Faria, C., Serra, I., & Girardi, R. (2014). A domain-independent process for automatic ontology population from text. Science of Computer Programming, 95, 26–43.

Gatteschi, V., Lamberti, F., & Demartini, C. (2012). Lo-match: A semantic platform for matching migrants' competences with labour market's needs. In Global Engineering Education Conference (EDUCON), 2012 IEEE (pp. 1-5). IEEE.

González-Eras, A. G., & Aguilar, J. (2015). Semantic Architecture for the Analysis of the Academic and Occupational Profiles Based on Competencies, Contemporary Engineering Sciences, 8, 1551–1563.

González-Eras, A., Aguilar, J. (2018) Determination of professional competencies using an alignment algorithm of academic and professional profiles, based on competence thesauri and similarity measures. Enviado para publicación.

Gluga, R., Kay, J., & Lever, T. (2013). Foundations for modeling university curricula in terms of multiple learning goal sets. IEEE Transactions on Learning Technologies, 6(1), 25–37.

Hassan, F. M., Ghani, I., Faheem, M., & Hajji, A. A. (2012). Ontology matching approaches for erecruitment. International Journal of Computer Applications, 51(2).

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. In Soviet physics doklady (Vol. 10, No. 8, pp. 707-710).

Malzahn, N., Ziebarth, S., & Hoppe, H. U. (2013). Semi-automatic creation and exploitation of competence ontologies for trend aware profiling, matching and planning. Knowledge Management & E-Learning, 5(1), 84.

Manning, C., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S., & McClosky, D. (2014). The Stanford CoreNLP natural language processing toolkit. In Proceedings of 52nd annual meeting of the association for computational linguistics: system demonstrations (pp. 55-60).

Mendonça, M., Perozo, N., & Aguilar, J. (2015, July). An approach for Multiple Combination of Ontologies based on the Ants Colony Optimization Algorithm. In Computer Aided System Engineering (APCASE), 2015 Asia-Pacific Conference on (pp. 140-145). IEEE.

Montuschi, P., Lamberti, F., Gatteschi, V., & Demartini, C. (2015). A semantic recommender system for adaptive learning. IT Professional, (5), 50–58.

Müller-Riedlhuber, H. (2009). The European Dictionary of Skills and Competencies (DISCO): an Example of Usage Scenarios for Ontologies. In I-SEMANTICS (pp. 467-479).

Paquette, G., Rogozan, D., & Marino, O. (2012). Competency comparison relations for recommendation in technology enhanced learning scenarios. CEUR Workshop Proceedings.

Rodríguez, T., & Aguilar, J. (2014). Aprendizaje ontológico para el marco ontológico dinámico semantico. DYNA, 81(187), 56–63.

Sateli, B., Löffler, F., König-Ries, B., & Witte, R. (2017). ScholarLens: extracting competences from research publications for the automatic generation of semantic user profiles. PeerJ Computer Science, 3.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: xx/xx/2016
Aceitação/Acceptance: xx/xx/2016

Algoritmo Memético Autoadaptativo para Solución de Problemas Combinatorios

Yuri Cristian Bernal Peña[1], Alvaro E. Ortiz Dávila[2], Astrid Z. Peñaranda Capera[3]

ycbernalp@correo.udistrital.edu.co, aeortizd@udistrital.edu.co, azpenarandac@correo.udistrital.edu.co

[1] Universidad Distrital "Francisco José de Caldas", Facultad de Ingeniería, Bogotá, Colombia.

[2] Universidad Distrital "Francisco José de Caldas", Facultad de Ingeniería, Bogotá, Colombia.

[3] Universidad Distrital "Francisco José de Caldas", Facultad de Ingeniería, Bogotá, Colombia.

Pages: 448–462

Resumen: Los problemas de optimización combinatoria son problemas de alta complejidad, cuyas mejores soluciones se han dado a través del uso de metaheurísticas. Los Algoritmos Meméticos (MA) hibridan un Algoritmo Genético (GA) con un Algoritmo de Búsqueda Local (LS), dando como resultado una metaheurística con ventajas en la exploración tanto global como local. La presente investigación propone un MA que durante la exploración ajusta su método y parámetros de Búsqueda Local dependiendo de la diversidad poblacional; así mismo propone que en la exploración no se descarten todas las ramificaciones provenientes de agentes (individuos o soluciones) infactibles, permitiendo así la búsqueda a través del espacio infactible de soluciones. El MA Autoadaptativo se pone a prueba utilizando el problema de la mochila o knapsack.

Palabras-clave: Algoritmos Meméticos, Autoadaptativo, Optimización Combinatoria, Algoritmos Genéticos, Espacio Infactible.

Self-adaptive Metadata Algorithm for Solution of Combinatorial Problems

Abstract: The problems of combinatorial optimization are problems of high complexity, whose best solutions have given with the use of metaheuristics. Memetic Algorithms (MA) hybridize a Genetic Algorithm (GA) with a Local Search Algorithm (LS), resulting in a metaheuristic with advantages in both global and local exploration. The present investigation proposes an MA that during the exploration adjusts its method and parameters of Local Search depending on the population diversity; likewise, it proposes that the exploration does not rule out all the ramifications coming from unfeasibles agents (solutions), thus allowing the search through the unfeasible space of solutions. The Self Adaptive MA is tested using the knapsack or backpack problem.

Keywords*:* Memetic Algorithms, Self-adaptive, Combinatorial Optimization, Genetic algorithms, Infeasible Space.

1. Introducción

En los problemas de optimización combinatoria (*Combinatorial Optimization Problem* COP) el objetivo es encontrar el máximo o el mínimo de una función sobre un conjunto finito de soluciones discretas. Básicamente en un problema combinatorio de optimización se desea encontrar un orden específico entre muchos posibles de un conjunto de elementos (Aarts & Lenstra, 2003) (Sait & H., 1999). Los COP entran muy a menudo como mínimo en la categoría de complejidad computacional NP-Hard (Leeuwen, 1998) (Lawler, Lenstra, & Rinnooy, 1985), ya que el número de posibles combinaciones de elementos del problema (siendo cada combinación de elementos una posible solución) puede ser muy grande, lo que dificulta, e incluso en ocasiones imposibilita alcanzar el óptimo en un tiempo finito, y si lo hace, no hay manera de asegurar que sea la solución óptima.

1.1. Algoritmos Meméticos

Los Algoritmos Meméticos (*Memetic Algorithm* MA) son meta-heurísticas que complementan conceptos de diferentes técnicas de resolución, como lo son algoritmos evolutivos (EA) (Holland, 1975) (Goldberg, 1989), (Davis, 1991) y algoritmos de búsqueda local o *Local Search* (LS). El adjetivo memético viene del término inglés meme, acuñado por (Dawkins, 1976) para designar al análogo del gen en el contexto de la evolución cultural. Los MA son una buena opción para abordar los problemas combinatorios (COP). La razón de la hibridación de la búsqueda evolutiva, de los cuales hacen parte los Algoritmos Genéticos ó *Genetic Algorithm* (GA), con algoritmos LS, radica en que los primeros son buenos en exploración global, garantizando así mayor diversificación y evitando el estancamiento en óptimos locales, mientras que los segundos son buenos optimizadores locales, que aceleran la llegada a un óptimo (Moscato & Cotta, 2003);

1.2. Autoadaptación

En los MA se ha estudiado la autoadaptación proponiendo incorporar diferentes métodos LS al proceso de competición (selección) y cooperación (cruce), aumentando la probabilidad de aplicar a los agentes aquellos métodos que han tenido mejores resultados durante la exploración (Ong & Keane, Meta Lamarckian Lerning in Memetics algorithms, 2004). También se ha propuesto incorporar procedimientos de aprendizaje como parte del cromosoma, pudiendo ser parte de dicho cromosoma una secuencia de heurísticas a aplicar al respectivo individuo o agente, incluso con auto-configuración en los procesos de cooperación, competición, topología de búsqueda, mejora individual y topología de búsqueda (Neri, Toivanen, Cascella, & Ong, 2007) (Nogueras & Cotta, 2015).

La presente investigación se concentra en un concepto simple: adaptar la profundidad de exploración en la búsqueda local e incluir eventualmente la búsqueda en el espacio no factible. Se incluyen dos parámetros nuevos, los cuales se ajustarán en cada nueva generación; el primero lo llamaremos *fitness2*, el cual nos permite explorar también el espacio no factible; el segundo es *la distancia promedio al mejor*, clave para establecer a que individuos o agentes se les aplicará una búsqueda local de mayor profundidad, adaptando así la búsqueda de forma ajustada para cada nueva población. Dichos parámetros se calculan y ajustan en el operador de evaluación del GA. Se espera que dicho ajuste de operadores y recálculo de parámetros disminuya el tiempo de búsqueda,

encontrando soluciones tan buenas o mejores que los GA's y MA's básicos, a pesar del recurso computacional extra que dicha autoadaptación pueda implicar.

2. Metodología

Para la elaboración del Algoritmo Memético (MA) final, se surtieron las siguientes etapas:

- Desarrollo del Algoritmo Genético (GA); sirve tanto de comparación, como de elemento constitutivo del MA.
- Desarrollo del Algoritmo de Búsqueda Local (LS), elemento constitutivo del MA.
- Desarrollo del MA simple de comparación.
- Desarrollo del MA autoadaptativo con los parámetros adaptativos de mejor resultado.
- Pruebas de contrastación de la eficiencia del MA autoadaptativo final con base en los anteriores.

El problema escogido para realizar las pruebas es el *Knapsack Problem* o Problema de la Mochila, es uno de los problemas NP-completos catalogado así por el informático teórico Richard Karp (Karp, 1972). Ha sido estudiado desde que fue planteado formalmente a finales del siglo XIX. Su fama se debe a lo sencillo de su planteamiento, pero lo complejo que es para resolver. Como todos los problemas estructurados representativos, las técnicas para su resolución pueden extrapolarse a problemas de optimización de diversas ramas del conocimiento.

En su forma básica consiste en llenar una mochila que soporta como máximo un peso determinado, con todo o parte de un conjunto de objetos, cada uno con un peso y valor específicos. Los objetos colocados en la mochila deben maximizar el valor total sin exceder el peso máximo.

Elementos: $x_i \quad i \in \{1..n\}$ Pesos: $w_i \quad i \in \{1..n\}$

Valores: $v_i \quad i \in \{1..n\}$

Capacidad Máxima: W

$$Máximizar \sum_{i=1}^{n} x_i v_i \tag{1}$$

S.A.

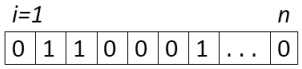
$$\sum_{i=1}^{n} x_i w_i \leq W \; \forall \; i \in \{1..n\} \tag{2}$$

$$3)\; x_i \in \{0,1\}$$

2.1. *Codificación del problema*

Cada agente representará una solución con la siguiente información:

Cromosoma: Cada posición del cromosoma (gen) será binaria, y representará la inclusión (1) o no inclusión (0) de un elemento x_i en la mochila.



2.2. Algoritmo de Búsqueda Local usado (Hill Climbing Voraz)

El algoritmo usado como operador de mejoramiento local individual o LS es el *Hill Climbing* o ascendiendo la colina, tipo *greedy* o voraz. Consiste en encontrar a la mejor solución aledaña o individuo vecino, y si este tiene mejor *fitness* o mejor adaptación que el individuo actual nos desplazamos para allá. Este proceso continuará hasta que no encontremos un vecino mejor, lo que significa que estaremos en un óptimo local, o tal vez, el óptimo global.

Para pasar a una solución vecina lo que se hace es cambiar un solo gen del cromosoma, y verificar la nueva adaptación. Se hace en forma exhaustiva para todas las combinaciones posibles, lo que denominamos combinación de un grado por nivel nodal. Esto supone que el número de combinaciones exploradas por nivel es de (l), o longitud del cromosoma.

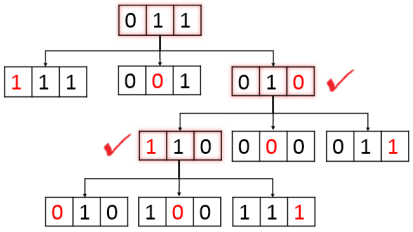


Figura 1 – Hill Climbing de combinación de un grado por nivel nodal (HC1)

El método propuesto realiza también en algunos de los agentes una exploración de mayor profundidad, en la que se cambian dos genes para pasar a una solución vecina a dos dimensiones o grados de distancia. A esto lo denominamos combinación de dos grados por nivel nodal. Sigue explorándose aquella rama que presenta la mayor mejoría, y termina cuando no se encuentre una mejor rama. Esto supone que el número de combinaciones exploradas por nivel es de (l^2), o el cuadrado de la longitud del cromosoma.

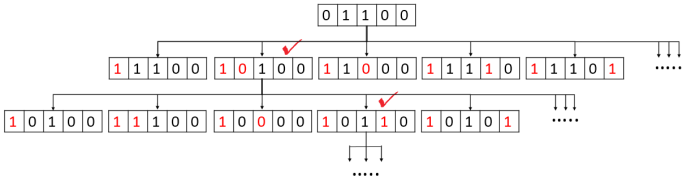


Figura 2 – Hill Climbing de combinación de dos grados por nivel nodal (HC2)

Debido al consumo de recurso computacional que requiere esta exploración de mayor profundidad, se debe ser cuidadoso con el número de individuos o agentes a los que se le aplica esta búsqueda local, pues de lo contrario el algoritmo puede llegar a perder

eficiencia en tiempo. En el método propuesto se planteó aplicar la exploración de dos grados a un porcentaje de individuos con los mejores puntajes, pero dependiendo de su separación respecto del mejor, con el objetivo de explorar en zonas dispersas, incentivando la diversidad poblacional.

2.3. Algoritmo Genético

En el GA, el operador de selección de individuos para cruce utiliza el método Montecarlo o regla de la ruleta y una población elite (que se mantiene en la siguiente generación) del 2%, donde el puntaje criterio lo calcula el operador de evaluación para cada individuo; el operador de cruce es mono punto con punto de inicio aleatorio; el operador de mutación tiene ocurrencia del 0.01, la cual es alta pero garantiza una mayor diversidad, de forma que compense la presión selectiva del algoritmo LS (Araujo, 2009).

2.4. Parámetro o Indicador Fitness principal y Fitness2 o complementario

En los GA el operador de evaluación se encarga de calcular la aptitud, puntaje o *fitness* del individuo (solución) mediante una función de adaptación; también se calcula la probabilidad de selección y probabilidad acumulada de selección, información que será utilizada por el operador de selección.

Normalmente dentro de los GA y los MA, a los individuos infactibles se les asigna cero como puntaje, lo que implica que la exploración está limitada al espacio factible. Algunos autores han propuesto no penalizar en extremo a las soluciones infactibles, e incluso se han realizado estudios comparativos de dichas técnicas (Alfonso, Bermúdez, & Stark, 2001). Hay investigaciones en las que incluso una fracción de la población puede estar conformada por soluciones infactibles (Singh, Isaacs, Ray, & Smith, 2008).

Se han propuesto métodos de penalización efectivos donde no sólo se penaliza a los individuos que violan restricciones, sino también a aquellos que "desperdician" el espacio de la mochila (Olsen, 1994) (Michalewicz & Dasgupta, 1996).

En el algoritmo propuesto el operador de evaluación calcula un puntaje de adaptación adicional o *fitness2,* en los agentes de cromosoma no factible, de modo que no se descarte su posible exploración del todo. En estos individuos infactibles su *fitness* principal sigue siendo cero. Dentro del operador de evaluación también se calcula la distancia del cromosoma del agente o individuo actual respecto al mejor agente, dato que permitirá estimar la diversidad poblacional.

Tratamiento a soluciones no factibles propuesto: Para mantener la filosofía de realizar la búsqueda por los espacios tanto factibles como infactibles, es necesario dar un puntaje diferente de cero a las soluciones infactibles, pero en un *fitness* complementario; para esto se propone el siguiente razonamiento. Establecemos cuantas unidades promedio el cromosoma del individuo infactible sobrepasa el límite de peso (Fig 2.2) estimándolo así:

$$ue = \frac{\sum_{i=1}^{m} w_i - W_{max}}{\frac{\sum_{i=1}^{n} w_i}{n}} \quad (3)$$

Siendo ***ue*** el número de unidades en exceso de la solución actual, calculado como la relación entre la diferencia entre el peso de las *m* unidades *w* de la solución y el peso máximo W_{max}, sobre el peso promedio de una unidad del total de *n* unidades.

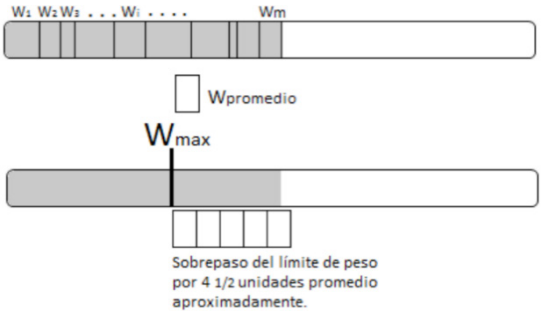


Figura 3 – Unidades promedio en exceso

El valor promedio del exceso ***ve*** se calcula multiplicando las unidades promedio de exceso por el valor promedio por unidad (Suma de los valore *v* sobre el número de unidades *n*):

$$ve = ue * \frac{\sum_{i=1}^{n} v_i}{n} \tag{4}$$

El *fitness2,* que es el que nos interesa como indicador de aptitud complementario, se calcula restándole al *fitness* que tendría la solución si fuera factible (al ser infactible su aptitud real será cero) el valor promedio de exceso:

$$fitness2 = fitness - ve \tag{5}$$

En esencia, el fitness2 representa la potencial aptitud que podría tener el individuo si eliminara las unidades que posee en exceso. Este valor no es utilizado por el algoritmo propuesto durante el proceso de selección, sino, durante el proceso de mejoramiento individual, o búsqueda local, con el objetivo de aproximar también por el espacio infactible.

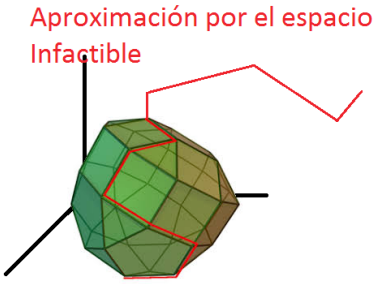


Figura 4 – La búsqueda por espacio infactible puede mejorar la eficiencia

2.5. Parámetro o Indicador Hamming para Diversidad Poblacional

Es sustancial para la eficiencia del algoritmo de búsqueda local, utilizar una heurística adecuada con conocimiento del problema (Hart & Krasnogor, 2004) (Ong, Lim, & Zhu, Classification of adaptive memetic algorithms: A comparative study, 2006), pero aún igual o más importante, escoger correctamente en que momento y a que individuos o agentes se le debe aplicar la mejora por LS en el proceso de exploración.

No usar en toda la población la búsqueda local, sino únicamente sobre algunas soluciones seleccionadas aleatoriamente, o en función de su calidad, o bien según algún otro criterio, es lo que se conoce como *lamarckismo parcial* (Cotta, 2005) (Houck, Joines, Kay, & Wilson, 1997). También cuenta la intensidad del aprendizaje, o sea durante cuánto tiempo o que complejidad debe tener la fase de aprendizaje individual (Sudholt, 2012).

La diversidad poblacional es clave para que la búsqueda local aumente su efectividad, las soluciones sobre las que se aplique deben pertenecer en lo posible en diferentes zonas del espacio de soluciones, así como tener una calidad razonable. Es por eso que se han realizado propuestas como el Algoritmo Memético Equilibrado con Diversificación Voraz (AMEDV), el cuál añade diversidad a la población en caso de que sea necesario, sustituyendo aquellas soluciones que verifiquen determinado criterio de semejanza con otras de la población por nuevos cromosomas generados por un algoritmo voraz aleatorizado (Herrera & Herrera, 2015).

Por lo tanto, en los MA la diversidad poblacional se puede utilizar como medida para seleccionar y activar los memes más apropiados (Caponio, Cascella, Neri, & Salvatore, 2007), coordinando los memes para que exista un balance entre diversidad y convergencia local.

En concordancia con lo anterior, el algoritmo desarrollado busca propender por la búsqueda en diferentes sectores del espacio de soluciones, por lo que se experimentó su aplicación con la búsqueda local (LS) de la siguiente forma:

El operador de evaluación calcula previamente para todos los agentes de la población, la distancia *Hamming* (Wang, Wang, & Yang, 2009). En esencia su aplicación es igual para cromosomas binarios, enteros o reales, pero para cromosomas binarios se aplica contabilizando el número de genes diferentes por posición entre los dos cromosomas a comparar. De este modo, entre el cromosoma **01010101** y **00001111** hay una distancia de 4 bits, genes o unidades. Este dato es guardado por el agente.

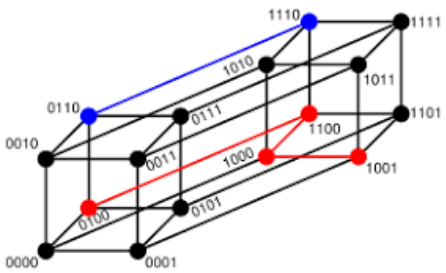


Figura 5 – Distancia Hamming entre cromosomas

2.6. MA Autoadaptativo propuesto

El Algoritmo Memético propuesto utiliza los elementos estándar de los Algoritmos Genéticos para los procesos de cooperación y competencia entre agentes. El método de Búsqueda Local que utiliza es *Hill Climbing* tipo *Greedy*, con capacidad de búsqueda por nivel nodal de uno y dos grados o dimensiones de combinación en profundidad (ver sección 2.2.). La aplicación de dicho LS sobre los agentes se realiza en forma adaptativa teniendo como variable de control el nivel de diversidad poblacional usando como indicador la distancia promedio al agente de mejor aptitud o *fitness*.

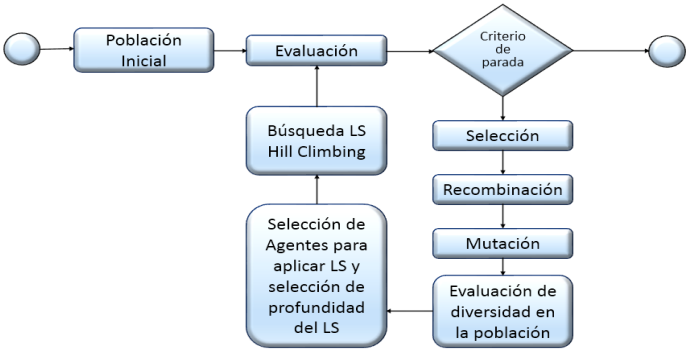


Figura 6 – Diagrama de Flujo del MA propuesto

2.6.1. Población Inicial

La población inicial es generada de forma aleatoria (Back & Hoffmeister, 1991), y es llevada a soluciones factibles rápidamente por el algoritmo LS profundo basado en el *fitness2*.

El tamaño de la población empleado es como mínimo de diez veces la longitud del cromosoma (10l), según lo estima la teoría (Alander, 1992).

2.6.2. Operadores de Selección, Cruce y Mutación

Los procesos de selección, cruce y mutación utilizados son los mismos usados en el Algoritmo Genético (GA) de prueba básico descrito en la sección 2.3.

2.6.3. Criterio de Parada

El algoritmo termina su proceso de búsqueda una vez converge en un óptimo. Se considera que el algoritmo convergió si pasadas *l* generaciones, la mejor solución obtenida no ha mejorado, siendo *l* el largo del cromosoma.

2.6.4. Operadores de Búsqueda Local

La aplicación del LS sobre los agentes se realiza en forma adaptativa teniendo como variable de control el nivel de diversidad poblacional; La idea es que cada generación

ajuste cuántos y cuáles individuos o agentes se les aplicara la búsqueda local y con qué profundidad.

Para cuantificar este nivel se usa como indicador la distancia *Hamming* promedio de la población respecto al agente de mejor aptitud o *fitness;* En los agentes infactibles dicha distancia se calcula con base en el *fitness2* o complementario. La exploración LS *Hill Climbing* de un grado de profundidad por nivel nodal se aplica a toda la población en cada nueva generación; se espera que su uso mejore en los agentes factibles el *fitness* y en los infactibles el *fitness2* independientemente del uso de la exploración adicional a dos grados de profundidad nodal en forma complementaria. El hecho de incluir a los no factibles de la población debe incentivar la diversidad.

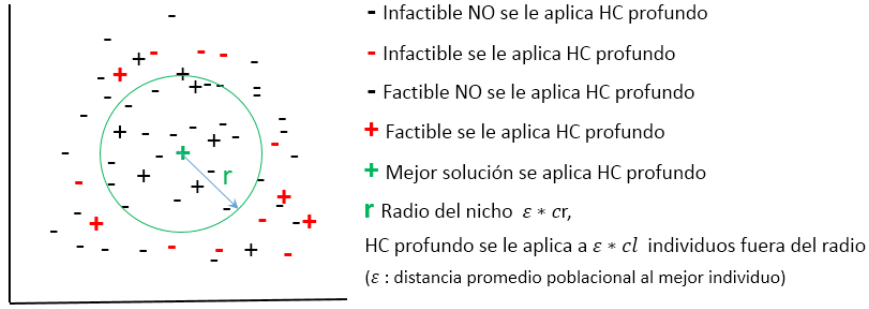


Figura 7 – Espacio de candidatos a mejora HC2 dos grados de profundidad nodal.

El método *Hill Climbing* a dos grados de profundidad por nivel nodal se aplica a los agentes o individuos clave (de importancia en la eficiencia como ya se mencionó); para esto se aplica dicho LS sobre los mejores ε*cl, siendo ε la distancia Hamming promedio al mejor agente o individuo en la respectiva generación y ***cl*** el coeficiente de longitud, que es un coeficiente para calcular el número de individuos seleccionados; se experimentó seleccionando individuos que además de cumplir con la condición anterior, también tuviesen una distancia *Hamming* (ver sección 2.5) al mejor individuo mínimo de ***ε*cr***, siendo ***cr*** un coeficiente radial; se aplica con agentes o individuos factibles e infactibles, de forma tal que se favorezca la búsqueda en zonas lejanas a la zona de la mejor solución actual, o sea espacios con baja probabilidad de exploración en la etapa evolutiva (ver figura 7).

2.6.5. Pruebas Hill Climbing Grado 2 y Grado 1

Debemos encontrar un buen coeficiente **cr** para la escogencia de un sector poblacional por fuera de radio donde se aplicará exploración profunda o LS HC2, así como un adecuado coeficiente **cl** para calcular el número de agentes o individuos por fuera de dicho radio a quienes se les aplicará LS HC2. Así mismo se debe decidir si se usa exploración local HC1 en forma complementaria, aplicándola a toda la población o a ningún individuo en cada generación.

El problema de prueba es el de la mochila con 120 grados de libertad. Para verificar cuáles son los coeficientes más apropiados, se realizaron las pruebas combinando el coeficiente del número de elementos a quienes aplicar LS HC2 (grado dos por nivel de

profundidad), con el radio mínimo de distancia al mejor para aplicar LS HC2, y si se aplica o no LS HC1 (grado uno por nivel de profundidad) a toda la población:

Se aplicó HC1 a toda la población	SI					NO					
Coeficiente Cr Para distancia mínima al mejor	0,01	0,05	0,1	0,15	0,2	0,25	0,3	0,35	0,4	0,45	0,5
Coeficiente Cl Para número de individuos seleccionados	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1	1,1	1,2

Tabla 1 – Combinaciones de valores de los tres parámetros en el experimento.

Por ejemplo, una combinación sería: Se aplica HC1 + Cr: 0,05 + Cl: 0,4.

El número de muestras tomadas (corridas) para cada una de las combinaciones es de 52. Se calcula utilizando la prueba estadística paramétrica asumiendo distribución normal con un nivel de confianza del 95% y un error máximo permitido de 5%. Se utiliza el promedio de las mismas para el análisis de los datos.

2.6.6. Algoritmo MA Autodaptativo vs Los Básicos

El experimento permite tener para el MA los parámetros ajustados que permitan autoadaptar para cada nueva generación la exploración de forma eficiente. Este Algoritmo Memético Autoadaptativo debe comparar su eficiencia con los algoritmos básicos, de forma tal que provea evidencia de tener un mejor desempeño.

3. Resultados

A continuación, se presentan los resultados gráficos del experimento:

Se concluye que el algoritmo NO debe incluir como parámetro un radio mínimo (***cr***) dinámico respecto al mejor individuo para seleccionar candidatos a búsqueda HC2, pues esto no garantiza diversidad, y si empeora la eficiencia.

Se concluye que el parámetro coeficiente a multiplicar por el promedio dinámico de diversidad poblacional (ε*cl) para el número de seleccionados a quienes se aplicará la búsqueda HC2 en cada generación, es conveniente introducirlo con valor de 0.4, lo cual mostró los mejores resultados.

Se concluye que aplicar HC1 a toda la población en cada generación mejora la eficiencia.

El seudocódigo de la exploración local en el MA Autoadaptativo quedaría así:

```
//Fitness –Fitness2 - ε EvaluarPoblación (poblacion);
for 1 to totalpoblacion do
  MejoraLocal HC1(Individuo);
End
```

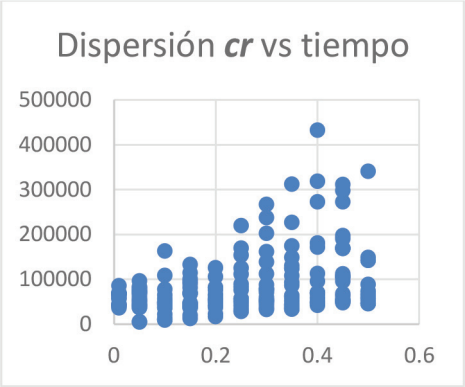


Figura 8 – Los coeficientes más bajos (radio tiende a cero) dieron menores tiempos

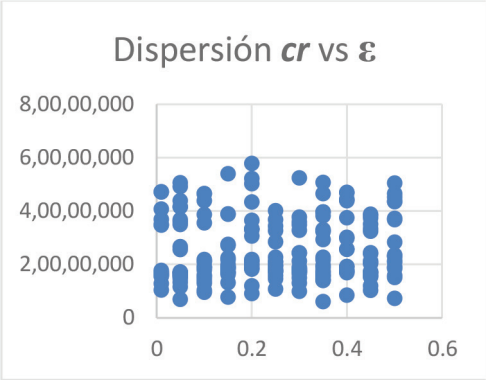


Figura 9 – El incluir un radio mínimo de aplicación no parece influir en la diversidad

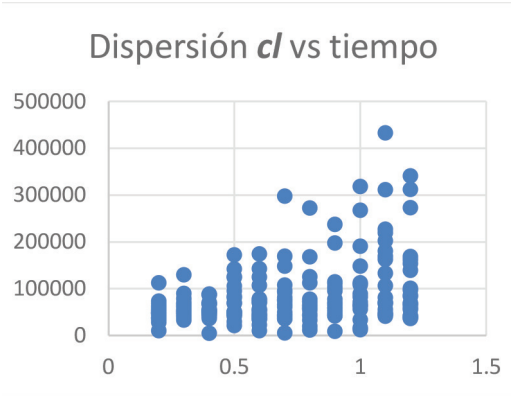


Figura 10 – El coeficiente ***cl*** =0.4 dio los menores tiempos de ejecución

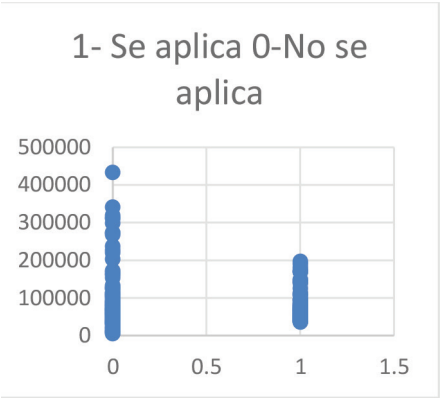


Figura 11 – La aplicación de HC1 a toda la población es más consistente

*//Selecciona los 0.4*ε mejores de la población*
selecciona_individuos*(0.4*ε);*
for all *individuos seleccionados* ***do***
  MejoraLocal *HC2(Individuo);*
end

3.2. Resultados Algoritmo MA Autodaptativo vs Los Básicos

El algoritmo MA autoadaptativo se puso a prueba frente a los siguientes algoritmos básicos:

- Algoritmo Genético simple.
- Algoritmo Memético Hill Climbing grado 1 a toda la población con aproximación exclusiva por espacio factible.

- Algoritmo Memético Hill Climbing grado 1 a toda la población con aproximación por espacio factible e infactible.

El problema de prueba es el de la mochila con 120 grados de libertad. La muestra o número de corridas utilizadas para todos los algoritmos fue de 47. Para estimar el número de corridas necesarias para cada algoritmo se empleó la prueba estadística paramétrica asumiendo distribución normal, con un nivel de confianza del 95% y un error máximo permitido de 5%. Los resultados son los siguientes:

Los algoritmos en orden de eficiencia (velocidad de respuesta y calidad de la misma) de mayor a menor rendimiento quedan así:

- Algoritmo MA autoadaptativo con aproximación por espacio factible e infactible.
- Algoritmo MA HC grado uno con aproximación por espacio factible e infactible.
- Algoritmo MA Simple HC grado uno con aproximación sólo por espacio factible.
- Algoritmo GA simple.

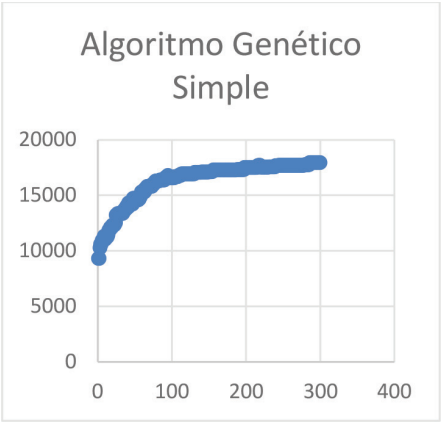


Figura 12 – Puntaje vs Generaciones

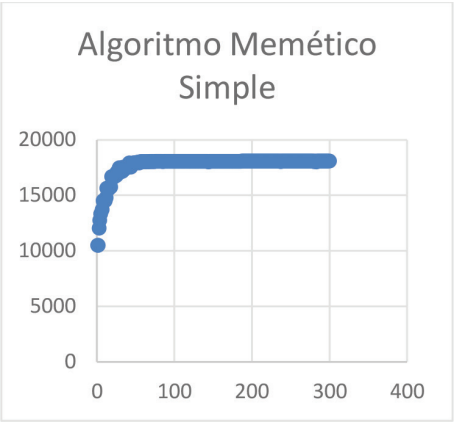


Figura 13 – Puntaje vs Generaciones

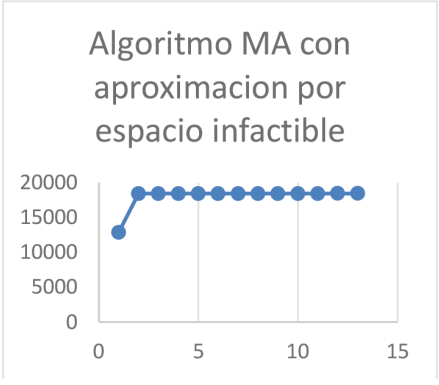


Figura 14 – Puntaje vs Generaciones

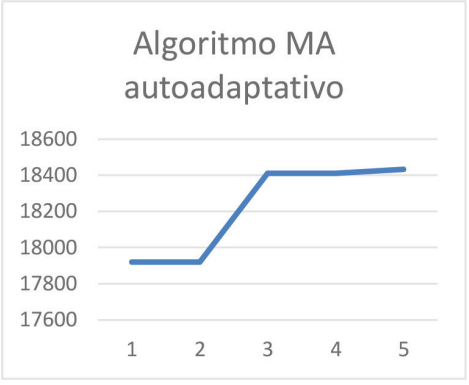


Figura 15 – Puntaje vs Generaciones

4. Conclusiones

La expectativa de la presente investigación se centraba en la capacidad de autoadaptación del algoritmo memético como técnica que lo hiciera más eficiente; sin embargo, la innovación más importante puede ser la creación de un segundo *fitness* o aptitud para los agentes; un puntaje adaptativo que indica para las soluciones no factibles cuál sería su valor si se convirtiera en una solución factible. Esta segunda aptitud permite el mejoramiento de estos agentes sin ignorarlos como suele pasar con la gran mayoría de métodos, pero sin entorpecer el mejoramiento de los factibles como ocurre cuando se unifican en un solo valor *fitness* para factibles e infactibles. Este aporte permite aproximarnos por el espacio factible e infactible al óptimo, mejorando en buena medida la eficiencia.

En cuanto a la autoadaptación, se investigó el efecto de la aplicación de una exploración más profunda con Hill Climbing de dos grados de exploración por nivel nodal (HC2) con el número de individuos y el tipo variando adaptativamente respecto al grado de convergencia de la población. Se observó que seleccionar los mejores individuos más lejanos con una distancia mínima al individuo óptimo, no contribuye a la diversidad poblacional, y además termina siendo ineficiente en la búsqueda convergente.

Se observó que la cantidad de individuos seleccionados para aplicar HC2 puede ajustarse a un factor de 0.4 veces la distancia promedio de la población respecto al mejor individuo, o 0.4ε. Si se aumenta este factor en exceso se incrementarán los cálculos necesarios para la búsqueda aumentando así el tiempo de ejecución; pero al disminuirlo, baja el nivel de exploración, lo que resiente la calidad de la búsqueda. Por lo tanto, el número de individuos a explorarse es adaptativo a la diversidad poblacional. Este valor por ahora es verificable sólo para el Problema de la Mochila.

Se puede considerar incluir en futuros trabajos las tablas tipo *hash* o estructuras de datos de matriz asociativa, las cuales permiten referenciar aquellas soluciones que ya se han explorado en profundidad para no perder tiempo computacional re-explorándolas (Joyanes & Zahonero, 2004). En dichas tablas la ubicación de soluciones ya exploradas debe ser un proceso mucho más rápido que el proceso de exploración nodal para búsqueda local, de esta manera mejorando la eficiencia global del algoritmo. Esto se puede lograr mediante el uso de la aptitud de la respectiva solución como llave para su referenciación en la tabla *hash*, con pocas colisiones esperadas. También es de sugerir investigar en trabajos futuros, las implicaciones de exploración por espacios no factibles en otras metaheurísticas poblacionales, como por ejemplo en los algoritmos genéticos simples GA, tanto en etapas de inicialización como de selección poblacional.

Referencias

Aarts, E., & Lenstra, J. K. (2003). Local search in combinatorial optimization. *Eindhoven y Atlanta: Princeton University Press.*, 1-137.

Alander, J. (1992). On optimal population size of genetic algorithms. *Proceedings CompEuro, Computer Systems and Software Engineering, 6th Annual*, (págs. 65-70).

Alfonso, H., Bermúdez, C., & Stark, N. (2001). Análisis Comparativo de Diferentes Soluciones Alternativas para el Problema de la Mochila. *Universidad Nacional de La Plata, Argentina*.

Araujo, L. &. (2009). *Algoritmos evolutivos: un enfoque práctico*. Madrid, España: RA - MA.

Back, T., & Hoffmeister, F. (1991). Adaptive search by evolutionary algorithms. *Models of Self-organization in Complex Systems. Mathematical Research. Akademie-Verlag*, 17–21.

Caponio, A., Cascella, G., Neri, F., & Salvatore, N. (2007). Afast adaptative memetic algorithm for on_line control design of pmsm drives. *IEEE Transactions on Systems Man and Cybernetics Part B, Special Issue onMemetic Algorithms, 37*(1), 28-47.

Cotta, C. (2005). Memetic algorithms with partial lamarckism for the shortest common super-sequence problem. *Artificial Intelligence and KnowledgeEngineering Applications: a Bioinspired Approach*, 84-91.

Davis, L. (1991). *Handbook of Genetic Algorithms*. New York: Van Nostrand Reinhold.

Dawkins, R. (1976). *The Selfish Gene*. Oxford: Clarendon Press.

Goldberg, D. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. MA: Addison-Wesley.

Hart, W., & Krasnogor, N. (2004). *Recent Advances in Memetic Algorithms*. Heidelberg: Springer-Verlag.

Herrera, A., & Herrera, F. (2015). Algoritmo genéticoequilibrado condiversificaciónvoraz. *Congreso de Metaheurísticas, Algoritmos Evolutivos y Bioinspirados MAEB*, *10*, págs. 9-18. España.

Holland, J. (1975). *Adaptation in Natural and Artificial Systems*. University of Michigan Press.

Houck, C., Joines, J., Kay, M., & Wilson, J. (1997). Empirical investigation of the benefitsof partial lamarckianism. *Evolutionary Computation 5*, 31-60.

Joyanes, L., & Zahonero, I. (2004). *Algoritmos y estructuras de datos*. Madrid: McGraw-Hill.

Karp, R. (1972). Reducibility Among Combinatorial Problems. *Complexity of Computer Computations. Nueva York: Plenum*, 85-103.

Lawler, E. L., Lenstra, J. K., & Rinnooy, A. H. (1985). *The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization*. John Wiley & Sons.

Leeuwen, J. v. (1998). *Handbook of Theoretical Computer Science. Vol. A, Algorithms and complexity*. Amsterdam: Elsevier.

Michalewicz, Z., & Dasgupta, D. (1996). Evolutionary algorithms for constrained engineering problems. *Computers & Industrial Engineering, 30*(4), 851-870.

Moscato, P., & Cotta, C. (2003). An Introduction to Memetic Algorithms. (19), 131-148.

Neri, F., Toivanen, J., Cascella, G., & Ong, Y. (2007). An adaptive multimeme algorithm for designing HIV multidrug therapies. *IEEE/ACM Transactions on Computational Biology and Bioinformatics, 4*(2), 264–278.

Nogueras, R., & Cotta, C. (2015). *Algoritmos Meméticos con Propiedades Self-* para la Optimización de Problemas Complejos*. Malaga: Tesis Doctoral. Universidad de Málaga.

Olsen, A. (1994). Penalty Functions and the Knapsack Problem. *IEEE First Conference on Evolutionary Computation*, (págs. 554-558). Orlando.

Ong, Y., & Keane, A. (2004). Meta Lamarckian Lerning in Memetics algorithms. *IEEE Transactions on Evolutionary Computation, 8*, 99-110.

Ong, Y., Lim, M., & Zhu, N. (2006). Classification of adaptive memetic algorithms: A comparative study. *IEEE Trans. Syst.,Man, Cybern. B, Cybern.36, no. 1*, 141-152.

Sait, S. M., & H., Y. (1999). Iterative computer algorithms with applications in engineering: Solving combinatorial optimization problems. . *IEEE Computer Society Press*, 1-248.

Singh, H. K., Isaacs, A., Ray, T., & Smith, W. (2008). Infeasibility Driven Evolutionary Algorithm (IDEA) for Engineering Design Optimization. *21st Joint Australiasian Conference on Artificial Intelligence AI-08*, (págs. 104–115). Australia.

Sudholt, D. (2012). Parametrization and balancing local and global search. En F. Neri, C. Cotta, & P. Moscato, *Handbook of Memetic Algorithms ,tomo 379 de Studies in Computational Intelligence* (págs. 55-72). Berlin: Springer-Verlag.

Wang, H., Wang, D., & Yang, S. (2009). A memetic algorithm with adaptive hill climbing strategy for dynamic optimization problems. *Soft Comput, 13*, 763–780.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 16/09/2018
Aceitação/Acceptance: 03/11/2018

Estabelecimento de Requisitos para Aplicações de Realidade Aumentada

Marcelo de Paiva Guimarães[1,2], Valéria Farinazzo Martins[3], Claudio Kirner[4], Maria Amélia Eliseo[3], André Luiz Satoshi Kawamoto[5]

marcelodepaiva@gmail.com, valeria.farinazzo@mackenzie.br, ckirner@gmail.com, mariaamelia.eliseo@mackenzie.br, kawamoto@utfpr.edu.br

[1] Open University of Brazil (UNIFESP) , Rua Sena Madureira 1500, 04021-001, São Paulo, Brazil

[2] FACCAMP's Master Program, R. Guatemala 167, 13231-230, Campo Limpo Paulista, Brazil

[3] Mackenzie Presbyterian University, Rua da Consolação 930, 0130-907, São Paulo, Brazil

[4] Federal University of Itajubá, Av. BPS 1303, 37500-903, Itajubá, Brazil

[5] Federal University of Technology – Paraná, Via Rosalina Maria dos Santos, 1233, 87301-899, Campo Mourão, Brazil

Pages: 463–475

Resumo: A Realidade Aumenta explora os sentidos humanos sobrepondo elementos virtuais (por exemplo, imagens 3D, videos e áudios) no ambiente real. Para isso, utiliza recursos que não são comumente usados nas aplicações convencionais, o que torna seu desenvolvimento uma tarefa específica. Embora possa ser mais eficaz utilizar um processo de desenvolvimento centrado nas peculiaridades de Realidade Aumentada, estas aplicações têm sido construídas levando em consideração apenas a experiência de seus desenvolvedores. Assim, torna-se objetivo deste artigo contribuir para a engenharia de requisitos das aplicações de Realidade Aumentada, que é uma fase fundamental no ciclo de vida de desenvolvimento de softwares. Para isso, discute os diversos aspectos das aplicações de Realidade Aumentada, em relação aos seus requisitos funcionais, não funcionais e específicos deste tipo de aplicação.

Palavras-chave: realidade aumentada; engenharia de requisitos; requisitos funcionais; requisitos não funcionais.

Establishment of Requirements for Augmented Reality Applications

Abstract: The Augmented Reality explores human senses by overlaying virtual elements (e.g. 3D images, video, and audio) in the real environment. To do this, it uses resources that are not commonly used in conventional applications, which makes its development a very specific task. Although it may be more effective to use a development process focused on the peculiarities of Augmented Reality, applications have been built taking into account only the experience of their developers. Thus, this paper aims to contribute to the Requirements Engineering Augmented Reality applications, which is a key phase in the software development

lifecycle. To that end, it discusses the various aspects of the Augmented Reality applications related to their functional, non-functional and specific requirements.

***Keywords**:* augmented reality; requirements engineering; functional requirements; non-functional requirements.

1. Introdução

Os avanços tecnológicos associados à diminuição dos custos nos últimos anos têm criado diversas novas aplicações computacionais e, consequentemente, novos desafios. Dentre elas, as que se baseiam na Realidade Aumentada (RA), que é uma tecnologia que ganhou notoriedade no momento atual devido ao aumento da capacidade de processamento do hardware, como dos celulares e dos computadores pessoais. Essa tecnologia possibilita que os usuários interajam com o mundo real com elementos virtuais sobrepostos (por exemplo, imagens 3D, videos e áudios) em tempo de execução, ou seja, o mundo real é enriquecido com objetos virtuais sintetizados por computadores no momento em que o usuário navega e interage com o mesmo. Dessa maneira, as aplicações de RA combinam objetos reais e virtuais em um mesmo ambiente real e executam de forma interativa (por exemplo, é possível ler e ouvir a história de uma estátua ao apontar a câmera do celular para mesma). O uso de aplicações de RA está atrelado a diversos contextos, como na Educação (Bacca et.al.,2014), Saúde (Wrzesien, 2011), indústria (Wrzesien, 2014) e disseminação de produtos e serviços (Olsson, 2013).

Embora seja uma tecnologia derivada da Realidade Virtual (RV) –vista como uma interface avançada na qual dispositivos multissensoriais são usados para navegar e interagir com um ambiente 3D sintetizado por computador -, a RA simplifica isso. Ela não imerge o usuário em um mundo totalmente sintetizado, ou seja, o usuário continua consciente do mundo real.

Cabe ressaltar que os elementos virtuais adicionados no mundo real, no caso da RA, podem ser voltados para estimular qualquer um dos sentidos humanos (visão, audição, tato, olfato e paladar). Além disso, envolvem pontos como mobilidade dos usuários, dispositivos específicos de visualização (por exemplo, celular, monitores e óculos de visualização) e de interação (marcadores ou não). Elas podem ser voltadas para dar suporte ao processo de ensino-aprendizagem de crianças com dificuldades de aprendizagem, auxiliar a realização de tarefas, como a troca de um pneu, ou até auxiliar o processo de fisioterapia de idosos, ou seja, elas podem estar presentes nos mais variados contextos. Como consequência, a forma de desenvolvimento, uso e validação das mesmas é diferente das aplicações convencionais, como, por exemplo, de um sistema de gestão de pessoas. Isso aponta uma oportunidade de investigação de técnicas, métodos e metodologias que auxilem o desenvolvimento das mesmas.

Na maioria das vezes, aplicações de RA têm sido criadas sem a utilização de uma metodologia que leve em consideração as especificidades desta classe de aplicação, em que somente a *expertise* dos desenvolvedores é utilizada, não contando com os recursos que a Engenharia de Software pode prover. Os trabalhos encontrados sobre Engenharia de Requisitos mais próximos de RA são referentes à Realidade Virtual (RV), os quais também possuem diversas lacunas de pesquisa (Santos, Delamaro & Nunes, 2013).

Este artigo tem como principal contribuição a investigação no levantamento dos requisistos das aplicações de RA, que é uma das fases fundamentais para o desenvolvmento de qualquer sistema (Robinson, Pawlowski & Volkov,2003; Nazir, Raana & Majeed, 2014; Martins, Guimarães & Kirner, 2015), a fim de que o comportamento do sistema e suas restrições devem acontecer. Esse estudo é fundamentado no conceitos oriundos da Engenharia de Software e na experiência dos autores no desenvolvimento de aplicações de RA e RV de diversos tipos.

O restante do artigo está organizado da seguinte forma: a seção 2 apresenta e discute as especificidades dos diversos tipos de aplicações de RA, levando em conta aspectos como forma de visualização e interação; a seção 3 discute os requisitos para aplicações de RA; e, por fim, a seção 4 apresenta as conclusões.

2. Tipos de Aplicações de Realidade Aumentada

As aplicações de RA podem ser classificadas de diversas maneiras (Hugues, Fuchs & Nannipieri, 2011;Kalawsky et.al., 2000; Normand, Servi`eres & Moreau, 2012; Peddie,2017). Quando considerado como o mundo real é apresentado (*display*) para os usuários, elas podem ser organizadas como:

- Visão ótica por projeção: o usuário visualiza o ambiente real com elemento virtuais projetados sem nenhum equipamento auxiliar, ou seja, os objetos virtuais são projetados diretamente na superfície do ambiente real. Como requisito específico desse tipo de projeção é a necessidade levar em conta que a projeção das imagens é feita em uma superfície e que se fizer uso do projetor, o mesmo deve ficar posicionado de tal maneira que o usuário não fique entre a imagens e a superfície projetada.
- Visão baseada em monitor: a imagem do ambiente real é capturada por uma câmera e misturada com os elementoss virtuais. Nela, o usuário visualiza as imagens resultantes em uma tela, como de um computador, celular ou *tablet*. Os requisitos desse tipo de aplicação devem considerar, por exemplo, a mobilidade do usuário ao usar um celular, ou não, ao usar um *desktop*. Além disso, os tipos de elementos virtuais, pois a capacidade de processamento de um desktop é diferente dos dispositivos móveis (celulares e *tablets*).
- Visão ótica direta: os objetos 3D são projetados nas lentes transparentes de um óculos ou de um capacete. Assim, o usuário visualiza o mundo real e os objetos 3D. Esse tipo de solução usa hardware específico, dessa maneira a elicitação dos requisitos deverá levar em conta pontos como a usabilidade dos equipamentos e fadiga.
- Visão direta por vídeo: os objetos 3D são misturados às imagens capturadas por uma câmera e são visualizadas pelo usuário, ou seja, o usuário visualiza um vídeo. Esse tipo de aplicação deve, por príncipio, utilizar um capacete de visualização, que comumente limita a movimentação do usuário e requer alto poder de processamento.

As diferenças entre os tipos de aplicações de RA podem acarretar grandes mudanças na forma de especificação dos requisitos e, consequentemente, na implementação. Por exemplo, elas podem ser criadas com baseadas no uso de marcadores, ou não.

Em aplicações baseadas em marcadores (*marker-based*), são detectados padrões (por exemplo, imagem 2D ou QRCode) nas imagens capturadas pelas câmeras, que foram previamente cadastradas. Ao detectar o padrão, o elemento virtual é adicionado na cena.

As aplicações que não usam marcadores (*markerless*), por sua vez, utilizam de recursos como GPS ou reconhecimento de imagens para determinar a posição e o momento na qual o elemento virtual é adicionado na cena.

As aplicações de RA podem suportar as formas tradicionais de interação com o usuário, como mouse e teclado, e também as não convencioais, como movimentação do marcador e do próprio dispositivo (por exemplo, o celular) ou até mesmo o usuário se movimentando no ambiente (a posição do mesmo pode determinar os elementos virtuais adicionados). Assim, nota-se que existe uma enorme gama de possibilidades de funcionalidades para as aplicações de RA, sendo que boa parte delas podem ser vistas como não convencionais, ou seja, não estão presentes nas aplicações comuns.

3. Levantamento de Requisitos para Aplicações de Realidade Aumentada

Para fazer o levantamento dos requisitos funcionais e não funcionais de uma aplicação, é necessário que os desenvolvedores de software trabalhem com os usuários finais do sistema a fim de que possam entender o domínio da aplicação, quais serviços que o sistema deve fornecer, o desempenho esperado do sistema, as restrições de hardware, entre outros (Sommerville & Sawyer, 1997). Os requisitos são levantados utilizando técnicas como: observação, cenários, ponto de vista, entrevistas, reuniões, casos de uso, além do levantamento e análise de documentação pertinente. Esta fase envolve a superação de diversas barreiras, como: pode existir entre os desenvolvedores e usuários falta de conhecimento compartilhado entre eles, ocorrência de discursos imprecisos, ideias pré-concebidas e terminologias diferentes da visão dos *stakeholders* da aplicação. Assim, esta fase deve envolver ações a fim de capturar e registrar informações que poderão servir para o entendimento o mais completo e correto possível entre os *stakeholders* (Sommerville; Sawyer, 1997; Dick; Hull; Jackson, 2017). Assim, pode-se focar esta fase em três elementos distintos: os usuários, o ambiente no qual a aplicação será executada, e as restrições impostas pelo hardware.

Uma descrição das características importantes dos usuários (por exemplo, habilidade, experiência, educação e treinamento) é imprescindível. Assim, pode-se definir questões ligadas ao desenvolvimento da aplicação, entre as quais pode-se citar: se a aplicação será mono ou multiusuário; se os usuários atuarão no local ou remotamente; e se os usuários do sistema serão pessoas experientes ou principiantes. Para aplicações de RA, identificar o tipo de usuário predominante da aplicação trará implicações no *design* da interface, na facilidade de uso, entre outros requisitos de usabilidade. Por exemplo, a Figura 1 apresenta uma aplicação de RA desenvolvida para o ensino de percepção

musical destinada ao público infantil. Percebe-se que o design da interface foi focado em imagens infantis e com cores vibrantes.

Figura 1 – Aplicação baseada em marcadores para apoiar a aprendizagem das propriedades sonoras (Fonte: Fecchio et. al. 2018)

Pode-se definir o ambiente em que a aplicação será usada de diversas maneiras, como o ambiente físico, de trabalho e/ou social. Em relação ao ambiente físico, praticamente todas as aplicações de RA utilizam câmera, portanto as condições de iluminação podem interferir diretamente no rastreamento de objetos da cena real. Assim, aspectos como esse, além de outros (ergonômicos e locomoção, por exemplo) devem ser considerados na especificação do ambiente físico de execução. A Figura 2(a) mostra um exemplo de projeto de posicionamento dos computadores criado que levou em consideração a iluminação. Observa-se que as mesas ficaram enfileiradas perpendiculares às janelas, e foram dispostas de tal maneira que os computadores ficaram divididos igualitariamente em cinco para cada fileira visando melhor distribuir os alunos. Na Figura 2(b), como o projeto foi executado no ambiente real.

Projeto (b) Ambiente Real

Figura 2 – *Layout* e ambiente real de laboratório construído para favorecer a iluminação

Restrições de Hardware. As restrições de hardware descrevem o equipamento necessário para executar uma aplicação de RA. Em geral, aplicações de RA utilizam dispositivos como câmeras e/ou sensores, mas alguns aplicativos exigem a criação de objetos com antecedência, como marcadores ou objetos virtuais tridimensionais. Devido à alta interatividade, aplicações de RA devem responder rapidamente às entradas do usuário (a uma razão mínima de 30 fps). Assim, o desempenho de todo o equipamento deve ser compatível – não apenas o processador, bem como as câmeras e demais dispositivos. Se, por exemplo, a aplicação utilizar interface gestual, será necessário um sensor de captura de movimentos. Caso a aplicação seja destinada a dispositivos móveis, poderá ainda fazer uso do giroscópio e do GPS - o uso de sinal de GPS também deve ser considerado, pois, muitas vezes, o sinal pode ser difícil ou demorado para ser detectado.

3.1. Requisitos Funcionais

Os requisitos funcionais especificam qual deve ser o comportamento de uma aplicação. Se forem descritos em alto nível, então são voltados para que sejam compreendidos pelos usuários finais. Contudo, se as funcionalidades forem detalhadas, suas entradas e saídas e outros pontos, então são voltados para os desenvolvedores. Eles podem ser representados, por exemplo, por meio de diagramas de UML, tal como o apresentado na Figura 3, por um diagrama de Caso de uso de uma aplicação de RA desenvolvida pelos autores.

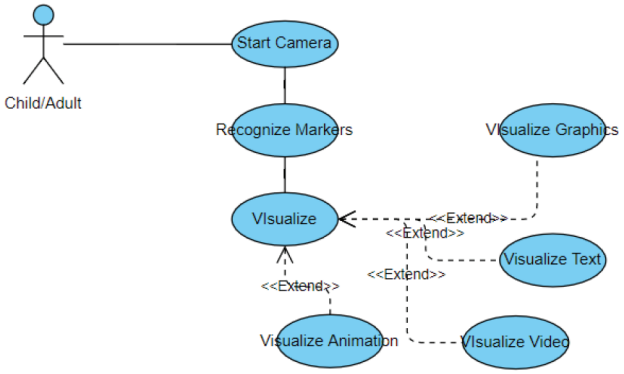


Figura 3 – Diagrama de Caso de Uso de uma aplicação de RA

Levando em consideração as peculiaridades das aplicações de RA, é necessário ponderar requisitos funcionais específicos para este tipo de aplicação. Devido à semelhança de alguns pontos de RA com RV (por exemplo, geração de objetos virtuais, interação com os objetos virtuais, entre outros), pode-se considerar três enfoques possíveis, que impactam de maneira distinta a especificação funcional da aplicação, conforme proposto por (Kirner, Salvador, 2007; Santos, Delamaro, Nunes, 2013; Stuart, 1996):

- Comportamento 1: quando o ambiente virtual reproduzirá tarefas e ações executadas por usuários em um ambiente real e se pretende retratar esta situação como ela é, ou seja, com realismo. Os sistemas de RA que utilizam interfaces gestuais podem desempenhar este enfoque com mais naturalidade, pois os movimentos podem se assemelhar aos movimentos dos usuários no seu cotidiano.

- Comportamento 2: quando os usuários desempenham determinadas tarefas no mundo real, mas se decide que o ambiente virtual reproduzirá tais tarefas de forma diferente ao que ocorre na realidade. Por exemplo, enquanto no mundo real a atividade de mudança é demorada, no ambiente virtual é realizada de forma instantânea.
- Comportamento 3: quando o ambiente virtual pretendido compreenderá a realização de tarefas e interações que não são desempenhadas no mundo real. Por exemplo, a utilização de um marcador a ser apresentado a uma câmera e a geração de um objeto 3D a partir daí é uma tarefa não desempenhada no mundo real.

Para os casos de uso descritos na Figura 3, por exemplo, os três comportamentos foram utilizados: assistir a um vídeo (*Visualize Video*) utiliza o Comportamento 1; visualizar gráficos *(Visualize Graphics)* é uma tarefa executada no mundo real, mas que aqui acontece de uma maneira diferente (Comportamento 2); e, finalmente, interagir com um objeto virtual não é uma tarefa executada no mundo real (Comportamento 3).

3.2. Requisitos Não Funcionais

Os requisitos não-funcionais são restrições sobre as funções ou os serviços ofe-recidos pelo sistema. Estão relacionados a software, hardware ou fatores externos, que determinem condições ou restrições ao comportamento do sistema pre-tendido, tal como restrições de *timing* e padrões. Esses requisitos são classificados em organizacionais, externos e de produto (Sommerville; Sawyer, 1997; Dick; Hull; Jackson, 2017):

- Requisitos **organizacionais** descrevem padrões de processos que devem ser adotados, requisitos de implementação (linguagem de programação ou método do projeto usado) e requisitos de entrega.
- Requisitos **externos** dizem respeito a interoperabilidade, aspectos legais e éticos. Por exemplo, ter que se comunicar com um banco de dados, atender a normas e legislação vigente no país, ou omitir dados de cunho pessoal.
- Requisitos de **produto** estão relacionados ao desempenho (tempo de resposta, consumo de recursos, *throughput*, etc.), confiabilidade, portabilidade, usabilidade (facilidade de uso, nível de satisfação de usuários, etc.), segurança, entre outros. Por exemplo, é preciso estabelecer um equilíbrio entre o número de polígonos que os objetos virtuais terão para garantir o realismo visual, mas que não impactem no tempo de resposta das interações e na apresentação dos objetos virtuais. Por outro lado, requisitos de portabilidade devem ser pensados e, por exemplo, atendidos com aplicações Web e/ou em dispositos móveis. Ao analisar a usabilidade, há de se conhecer o perfil do público-alvo que utilizará a aplicação de RA e adequá-la, atendendo aos três princípios da usabilidade: eficácia, eficiência e satisfação do usuário.

A usabilidade é um aspecto crítico para esta classe de aplicações, visto que envolve uma tecnologia com interface não convencional que ainda está sendo disseminada. Este requisito deve ser buscado por meio da identificação de formas que visem facilitar a interação, tais como:

- Definição de metáforas claras e compatíveis com os sistemas de RA e em concordância com o perfil dos usuários.

- Vídeos e informações textuais iniciais para orientar os usuários em como utilizar o sistema.
- Proporção de comandos e objetos tridimensionais sobrepostos; neutra-lidade das cores dos objetos virtuais frente ao mundo real.
- Persistência do posicionamento dos elementos sobrepostos ao mundo real independentemente da velocidade de movimentação da câmera.
- Boa localização dos elementos de controle a fim de facilitar a interação.
- Quantidade adequada de objetos virtuais e informações de maneira a não causar sobrecarga de informações, entre outros.
- Ergonomia cognitiva: Descrita através de critérios, tais como: condução (por exemplo, *feedback* imediato, clareza nas características lexicais), carga de trabalho, controle explícito, adaptabilidade, gestão de erros, etc.) (Scapin & Bastien, 1997).

Assim como qualquer outro tipo de sistema de software, as aplicações de RA devem atentar aos requisitos não funcionais, que devem ser especificados com clareza, de forma precisa.

3.2.1. Requisitos Não Funcionais Específicos para Aplicações de RA

Além dos aspectos de qualidade destacados na subseção 3.2, as aplicações de RA agregam peculiaridades que demandam a consecução de requisitos não funcionais adicionais (Guimarães; Martins, 2014; Martins; Kirner; Kirner, 2015). Neste sentido, espera-se que um sistema de RA algumas características, descritas a seguir.

Realismo visual. As aplicações de RA envolvem o desenvolvimento de elementos que não são comumente criados para as aplicações, como objetos 3D, texturas e áudios. A modelagem dos objetos 3D das aplicações de RA é complexa, pois além de representar o objeto desejado, o modelo deve ter baixo número de polígonos, caso contrário, a leitura e manipulação do objeto não atenderá aos requisitos de interação. Como consequência, a modelagem pode levar um tempo considerável do projeto como um todo para ser realizada. As texturas e os áudios também devem buscar ter alta qualidade, a fim de proporcionar realismo visual, mas seu tamanho não pode afetar o desempenho da aplicação.

Configuração do ambiente. Os aplicativos de RA exigem dispositivos especiais, como sensores e / ou câmeras. Além disso, marcadores podem ser necessários, tais como marcadores fiduciais. A configuração do ambiente deve ser o mais simples possível (Guimarães; Martins, 2014).

Interface e Interação. As aplicações de RA podem ser voltadas para diversos tipos de dispositivos, como celulares, *tablets* e computadores. Isso gera a necessidade de levar em conta a apresentação e a interação da aplicação conforme o dispostivo. Por exemplo, a interface e a forma de interação em um computador pessoal são diferentes quando comparado com a de um celular. A interação dos usuários com a aplicação é um aspecto de destaque nas aplicações de RA. Faz parte dessa abordagem estabelecer o uso ou não

de marcadores; se a interface permitirá a interação via mouse e teclado; se terá botões; as mensagens apresentadas para os usuários e outros aspectos. Por exemplo, a aplicação apresentada na Figura 4 é voltada para crianças, então foi necessário levar em conta o tamanho de suas mãos para definir o tamanho dos marcadores com o objetivo de facilitar a manipulação. Nela é sobreposto um dado 3D no ambiente quando o marcador é detectado.

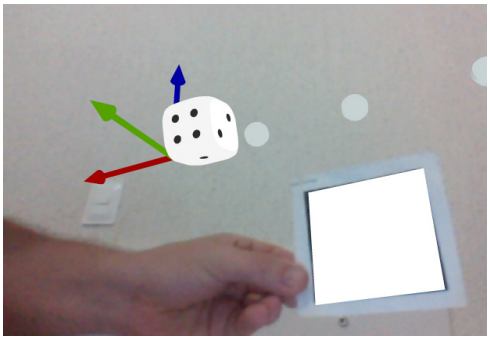


Figura 4 – Aplicação de RA para apoiar a aprendizagem de geometria (Fonte: Alves et. al., 2017)

Tempo de Resposta. O tratamento das interações deve ser em tempo real, dessa forma permitirá, por exemplo, que os objetos 3D, texturas e áudios sejam inseridos na cena no momento em que o marcador é apresentado para a câmera. Aplicações de RA que envolvam som e animação devem respeitar um baixo tempo de resposta para que a cena e o áudio estejam sincronizados.

Mapeamento da cena e Correspondência entre o sistema e o mundo real. O design do sistema deve seguir as convenções do mundo real, fazendo com que as informações apareçam em uma ordem natural e lógica e os objetos virtuais e reais se misturem adequadamente em relação a seu posicionamento. Manter a proporção dos objetos virtuais e reais é fundamental para que o usuário sinta-se mais imerso no ambiente (Martins; Guimarães; Corrêa, 2013; Guimarães; Martins, 2014).

Calibração. Sobretudo sistemas de RA que não utilizam marcadores devem ser calibrados com a câmera que fará a detecção e identificação dos objetos reais da cena. Por exemplo, quando se utiliza o corpo como marcador, existe a necessidade do sistema se adaptar às características físicas de cada pessoa, sendo assim necessário realizar a calibração

Ergonomia. O sistema deve ter um fator de forma ergonômico que permita a manutenção do dispositivo de RA por longos períodos sem fadiga excessiva e realizando operações típicas com alta conveniência (Schall et al., 2009). Aplicações de RA que utilizam gestos tem mais problemas em causar fadiga muscular do que aqueles que são feitos com marcadores (*desktops* ou celulares), então é preciso levar este requisito em consideração mais seriamente, dependendo do tipo de aplicação de RA, para não comprometer seu sucesso (Martins; Niedermeyer; Guimarães, 2016).

Cabe ressaltar que as aplicações de RA envolvem o desenvolvimento de elementos que não são comumente criados para as aplicações, como objetos 3D, texturas e áudios. A modelagem dos objetos 3D das aplicações de RA é complexa, pois além de representar o objeto desejado, o modelo deve ter baixo número de polígonos, caso contrário, a leitura e manipulação do objeto não atenderá os requisitos de interação. Como consequência, a modelagem pode levar um tempo considerável do projeto como um todo para ser realizada. As texturas e os áudios também devem buscar ter alta qualidade, mas o tamanho não pode afetar o de-sempenho da aplicação.

Em resumo, a partir do que foi discutido nas seções anteriores, é possível esquematizar requisitos, tal como apresentado nas Tabelas 1 e 2.

Requisito	**Descrição**	**Específico para RA?**
Comportamento 1	As ações no mundo virtual são iguais as do mundo real.	Sim
Comportamento 2	As ações no mundo virtual são executadas de maneira diferente das ações no mundo real.	Sim
Comportamento 3	As ações só existem no mundo virtual.	Sim
Organizacionais	Padrões de processos que devem ser usados, requisitos de implementação (linguagem de programação ou método do projeto usado) e requisitos de entrega.	Não
Externos	Interoperabilidade, legais e éticos	Não
De Produto	Desempenho (tempo de resposta, consumo de recursos, *throughput*, etc.), confiabilidade, portabilidade, usabi-lidade (facilidade de uso, nível de satisfação de usuários, etc.), segurança, entre outros.	Não

Tabela 1 – Requisitos Funcionais para Aplicações de RA

Requisito	**Descrição**	**Específico para RA?**
Realismo visual	Qualidade apropriada dos objetos virtuais para ter aparência dos objetos reais.	Sim
Configuração do ambiente	Os usuários têm que ser capaz de configurar a aplicação facilmente, por exemplo o uso da câmera.	Sim
Interface e interação	Uso de marcadores ou não, tipos de dispositivos e de elementos de interface.	Sim
Tempo de resposta	Baixo tempo de resposta para garantir a sincronização de animação, visualização, áudio.	Sim
Mapeamento da cena	Objetos virtuais e reais se misturem adequadamente em relação a seu posicionamento.	Sim
Correspondência entre o sistema e o mundo real	Design do sistema deve seguir as convenções do mundo real, fazendo com que as informações apareçam em uma ordem natural e lógica.	Sim
Calibração	Calibrar a câmera que fará a detecção e identificação dos objetos reais da cena.	Sim
Ergonomia	Evitar fadiga muscular.	Sim

Tabela 2 – Requisitos não-funcionais para Aplicações de RA

4. Conclusões

As aplicações de RA buscam explorar os sentidos humanos fundamentais sobrepondo elementos virtuais no ambiente. Por isso, oferecem e envolvem recursos diferenciados, por exemplo, de visualização, interação e navegação, quando comparadas com as aplicações convencionais. Contudo, pelo fato de envolver recursos relativamente recentes, muitas aplicações tem sido desenvolvidas levando em conta apenas com a experiência dos desenvolvedores, sem a utilização de métodos e metodologias provenientes da Engenharia de Software.

Este trabalho focou no levantamento dos requisitos, uma fase fundamental de todo o processo de desenvolvimento de uma solução de RA. Devido à complexidade dos diversos tipos de aplicações de RA, esse artigo abordou os principais pontos que podem interferir no sucesso ou não da solução criada (Nuseibeh & Easterbook, 2000), tratando desde as funcionalidades da aplicação em si até aspectos como o tamanho do marcador. Então, é esperado que a fase de levantamento dos requisitos resulte em um documento com a especificação dos requisitos funcionais e não funcionais dos mesmos, incluindo itens como os referentes a dados, ambiente físico, usuários, usabilidade e outros abordados neste trabalho. Além disso, apesar de difícil, o levantamento dos requisitos deve buscar a completude, que garante a definição de todas as funcionalidades, e a consistência, que trata de não existir definições contraditórias. Por isso, sempre que possível, é desejável associar métricas aos requisitos.

Como trabalhos futuros, pretende-se validar, e eventualmente incluir e alterar todos os requisitos aqui propostos. Esse processo será colaborativo e conta com o apoio de outros desenvolvedores de aplicações de RA. Além disso, visa-se a expansão deste estudo de forma a abranger todo o processo de engenharia de requisitos, desde a elicitação até a validação destes requisitos.

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de nível superior - Brasil (CAPES) - Programa de Excelência - Proex 0653/ 2018

Referências

Bacca, J., Baldiris, S., Fabregat, R., Graf, S. (2014). Augmented reality trends in education: a systematic review of research and applications. Journal of Educational Technology & Society. 17(4). 133.

Wrzesien, M., Burkhardt, J. M., Alcañiz, M., Botella, C. (2011). How technology influences the therapeutic process: a comparative field evaluation of augmented reality and in vivo exposure therapy for phobia of small animals. IFIP Conference on Human-Computer Interaction. Springer. Berlin, Heidelberg. 523-540.

Wang, X., Truijens, M., Hou, L., Wang, Y., Zhou, Y. (2014). Integrating Augmented Reality with Building Information Modeling: Onsite construction process controlling for liquefied natural gas industry. Automation in Construction. 40. 96-105. doi: doi10.1016/j.autcon.2013.12.003

Olsson, T., Lagerstam, E., Kärkkäinen, T., Väänänen-Vainio-Mattila, K. (2013). Expected user experience of mobile augmented reality services: a user study in the context of shopping centres. Personal and ubiquitous computing. 17(2). 287-304.

Santos,C. C. D, Delamaro, M. E. & Nunes, F. L. S. (2013). The Relationship between Requirements Engineering and Virtual Reality Systems: A Systematic Literature Review, 2013 XV Symposium on Virtual and Augmented Reality, Cuiaba, 53-62.

Robinson, W. N., Pawlowski, S. D. & Volkov, V. (2003). Requirements interaction management. Journal ACM Computing Surveys (CSUR). 35:132–190.

Nazir, A., Raana, A. & Majeed, N. (2014). Highlighting the role of Requirement Engineering and User Experience Design in Product Development Life Cycle. International Journal of Modern Education and Computer Science. vol. 6. no. 1. 34-40.

Martins, V. F., Guimarães, M.P. & Kirner, C. (2015). Development and Evaluation of Augmented Reality Educational Applications. In: Raymond Hughes. (Org.). Augmented Reality: Developments, Technologies and Applications. 1ed. Hauppauge: Nova Science Publisher. v. 1. 1-22.

Hugues, O., Fuchs & P., Nannipieri, O (2011). New Augmented Reality Taxonomy: Technologies and Features of Augmented Environment. Springer New York. New York. NY. 47–63. doi:10.1007/978-1-4614-0064-6 2.

Kalawsky, R. S., Stedmon, A.W., Hill, K. & Cook, C.A. (2000). A taxonomy of technology: Defining augmented reality. Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 44 (5). 507–510. doi:10.1177/154193120004400506.

Normand, J.M., Servi`eres, M., Moreau, G. (2012). A new typology of augmented reality applications. Proceedings of the 3rd Augmented Human International Conference. AH '12, ACM. New York. NY. USA. 1–18. doi:10.1145/2160125.2160143

Peddie, J. (2017). Augmented Reality: Where We Will All Live. Types of Augmented Reality. Springer International Publishing. Cham. 9–46 doi:10.1007/978-3-319-54502-8 2.

Sommerville, I., Sawyer, P. (1997). Requirements engineering: a good practice guide. John Wiley & Sons. Inc.

Stuart, R. (1996). The Design of Virtual Environmets. McGraw-Hill. Fairfield. PA. USA.

Dick, J., Hull, E., Jackson, K. (2017). Requirements engineering. Springer.

Pressman, R. S. (2011), Engenharia de Software: Uma Abordagem Profissional. 7a. Edição. McGraw-Hill.

Nuseibeh, B. & Easterbook, S. (2000). Requirements engineering: A roadmap. International Conference on Software Engineering (ICSE'00). Limerick, Ireland. June 04. 1 35–46.

Carvalho, B. A (2016). Modelo de desenvolvimento de aplicações de realidade aumentada para empacotamento e distribuição no formato de objetos de aprendizagem. Dissertação de Mestrado em Ciência da Computação. Dezembro. Centro Universitário Campo Limpo Paulista.

Fecchio, D. Y.; Fenley, F. J. G.; Nunes, S. S.; Correa, A. G. D.; Martins, V.F. Instrumenteca RA: Um Objeto de Aprendizagem para Auxiliar a Percepção Musical através da Diferenciação dos Sons de Instrumentos Musicais. RENOTE. REVISTA NOVAS TECNOLOGIAS NA EDUCAÇÃO, v. 16, p. 1-10, 2018.

Kirner, T. G. & Salvador (2007), V. F. M. Salvador. Contribution to the Requirements Engineering of Virtual Environments. In: 9th International Conference on Enterprise Information Systems. Volume HCI. Funchal, Madeira. Portugal. June 12-16. 142-157.

Martins, V. F., Niedermeyer, A., & de Paiva Guimarães, M. (2016, October). Educational game for aid in literacy process: Comparative study between kinect and mouse use as interaction and motivation way. In Learning Objects and Technology (LACLO), Latin American Conference on (pp. 1-9). IEEE.

Alves, B., Dias, D. R. C., Borges, S. de S., Durelli, V. H. S., Bressan, P. A., Martins, V. F., & de Paiva Guimarães, M. (2017). On Capitalizing on Augmented Reality to Impart Solid Geometry Concepts: An Experimental Study. In Universal Access in Human–Computer Interaction. Designing Novel Interactions (p. 105–117). Springer International Publishing. https://doi.org/10.1007/978-3-319-58703-5_8

Souza, R. C., Moreira, H. D. F., & Kirner, C. (2012). FLARAS 1.0–Flash Augmented Reality Authoring System.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 22/10/2018
Aceitação/Acceptance: 06/12/2018

Plataforma y aplicación móvil para proporcionar información del transporte público utilizando un dispositivo embebido de bajo costo

Jorge Buele[1], Franklin Salazar L.[1], Santiago Altamirano[1], Abigail Aldás R.[1], Pilar Urrutia-Urrutia [1]

jl.buele@uta.edu.ec, fw.salazar@uta.edu.ec, santiagomaltamirano@uta.edu.ec, naldas3666@uta.edu.ec, elsapurrutia@uta.edu.ec

[1] Facultad de Ingeniería en Sistemas, Electrónica e Industrial, Universidad Técnica de Ambato, 180103, Ambato, Ecuador

Pages: 476–489

Resumen: En este proyecto se presenta el desarrollo de un sistema informativo de transporte público urbano, para mejorar la movilidad de los usuarios. Este sistema consta de una pantalla informativa ubicada en la parada, por donde un bus realiza su recorrido. Adiciona la reproducción de mensajes auditivos, que promueven la inclusión de personas con discapacidad visual y la población analfabeta. Para una mayor interacción con los pasajeros se ha implementado una plataforma donde se puede apreciar la ruta a seguir, horarios, clima y el tiempo pronosticado hasta que otra unidad de transporte llegue a dicho lugar. Además de una aplicación intuitiva compatible con teléfonos inteligentes que utilicen sistema operativo Android. Para el procesamiento de los datos se ha utilizado la tarjeta embebida Raspberry Pi 3 como unidad central del sistema. Paralelamente se ha incorporado un servidor web, que permite el procesamiento de información y su almacenamiento en una base de datos.

Palabras-clave: Aplicaciones móviles; Transporte público; Servidores; Bases de datos.

Platform and mobile application to provide information on public transport using a low-cost embedded device

Abstract: This work presents the development of an urban public transport information system, to improve user mobility. The system contains an informative screen located in the stops where a bus makes its journey. It adds auditory messages, which promote the people inclusion with visual disabilities and the illiterate population. For greater interaction with passengers, a platform has been implemented where you can see the route to follow, timetables, weather and the predicted time until another bus unit reaches that stop. In addition, an intuitive application compatible with Android smartphones has been presented. For data processing, the Raspberry Pi 3 embedded board has been used as the central unit of the system. At the same time, a web server has been incorporated, which allows the processing of information and its storage in a database.

Keywords: Mobile applications; Public transportation; Servers; databases.

1. Introducción

La acelerada evolución de los sistemas de comunicación, ha permitido que el conocimiento se difunda de una forma más rápida y eficiente (Johnson, Lukaszewski & Stone, 2016). Esta realidad ha contribuido a que la información se desplace a nivel global, en un menor periodo de tiempo y sin limitaciones (Shakir, Hammood & Muttar, 2018). Las principales herramientas utilizadas para este fin son el internet y los servicios web, que permiten realizar consultas sobre el clima, ubicación en tiempo real, sitios turísticos, alimentación, centros médicos, entre otros (Arciniegas Aguirre, y otros, 2018). Todo ello desde la comodidad de un ordenador personal o desde un teléfono inteligente que presente las prestaciones necesarias de conectividad. La inmersión de estos dispositivos ha llegado a representar una verdadera revolución a nivel mundial en las últimas décadas, donde su producción en masa ha crecido exponencialmente (Silva, Toasa, Guevara, Martinez, & Vargas, 2018). De forma paralela el desarrollo de interfaces y aplicaciones se ha convertido en una nueva tendencia comercial sin discriminación.

De esta forma el ser humano se encuentra interconectado con el medio que lo rodea. Lo cual le ha permitido mejorar su calidad de vida y para ello se han implementado sistemas que involucran a las Tecnologías de la Información y la Comunicación (TIC) e Internet de las Cosas (IoT por sus siglas en inglés) (Harris, 2017). Un claro ejemplo es su aplicación en los entornos urbanos, permitiéndoles convertirse en las denominadas "ciudades inteligentes". Dentro de ello, se debe recalcar que la gestión del transporte público se ha convertido en un aspecto de primer orden, ya que es utilizado por la mayoría de los ciudadanos cotidianamente (Abdelsalam, y otros, 2018). De su correcto manejo y organización depende el crecimiento político, demográfico, económico y social de la región local. Por ejemplo, con la implementación de los contadores de pasajeros en los buses de transporte urbano, se puede obtener estadísticas del uso diario de este servicio. Así, los organismos responsables y las agencias de tránsito locales pueden tomar mejores decisiones, que optimicen el servicio que se le presta al usuario y permite la inclusión de personas con diversas discapacidades (Guo, Derian & Zhao, 2015).

Este trabajo basado en las investigaciones descritas en el capítulo posterior, propone el desarrollo de interfaces informativas para mejorar el servicio del transporte público urbano. Este sistema se implementa utilizando tecnología de bajo costo, que permite mantener a los usuarios informados sobre las prestaciones que se les brinda y la inclusión de personas con capacidades diferentes. Como unidad de procesamiento central se utiliza Raspberry Pi 3 y para el procesamiento y almacenamiento de información un servidor web cargado en la misma tarjeta. Para validar este prototipo, se realizan las respectivas pruebas experimentales utilizando 2 buses y 3 parada. Al realizarlas con línea de vista directa o sin ella, se observa una respuesta eficiente como se muestra en la sección de resultados.

Por tanto, este artículo se organiza de la siguiente manera: la introducción en la sección 1, en la sección 2 se muestra el estado del arte de investigaciones relacionadas. En la sección 3 se presenta la descripción del sistema y los elementos que forman parte del mismo. Los resultados de las pruebas experimentales realizadas y las conclusiones se describen en las secciones 4 y 5 respectivamente.

2. Estado del arte

Innovar los sistemas de transporte público, es una labor que requiere una fuerte inversión económica y un estudio que pueda demostrar la factibilidad de su implementación. Por lo cual se han desarrollado algunas investigaciones que han tratado esta temática como se presenta a continuación. Para poder ofrecer un mejor servicio es necesario entender el comportamiento humano e identificar patrones sobre su movilización, como se analiza en el trabajo de (Siła-Nowicka, y otros, 2016). El resultado de este estudio es un conjunto de datos enriquecido contextualmente, que admite nuevas posibilidades para modelar el comportamiento del movimiento humano. En (Mazloumi, Currie & Rose, 2010) por su parte se analiza la variabilidad del tiempo de viaje del transporte público, mediante el uso de datos obtenidos con un dispositivo GPS (Global Positioning System en inglés). Esta propuesta se implementa en una ruta de autobús en Melbourne, en distintos momentos del día, variaciones del tráfico y otros factores que contribuyen a la investigación. Con la intención de identificar las causas por las cuales existe congestión vehicular en las urbes en (Stoll, 2016) se realiza el monitoreo de un autobús. Se ha creado un conjunto de herramientas para ayudar a visualizar las tendencias en el comportamiento y el movimiento mientras realiza su recorrido. Este tipo de información es importante para las autoridades de tránsito que buscan mejorar los tiempos de funcionamiento del transporte público y para la investigación que se está realizando. De manera similar en (Bie, Gong & Liu, 2015) se presenta un algoritmo basado en datos tomados con un GPS, para dividir las horas de operación de un bus en intervalos de tiempo durante un día. Este método consta de tres pasos y se prueba en una ruta en Suzhou, China para validar lo propuesto.

A través del desarrollo tecnológico se han implementado nuevas herramientas para realizar el monitoreo de las unidades de transporte. En (Lee, Tewolde & Know, 2014) se desarrolla un sistema económico para el seguimiento y rastreo de vehículos desde cualquier ubicación y en cualquier momento. Para ello se combina el uso de un dispositivo GPS con la tecnología GSM/GPRS, que es una de las formas más comunes para el seguimiento vehicular. Para la visualización de la información se presenta una aplicación para teléfonos inteligentes, que recibe la información desde un microcontrolador. De manera similar en (Megalingam, y otros, 2014), utilizando módulos GPS y GSM se establece comunicación entre la unidad móvil de trasporte y la parada. De esta manera se muestra un tiempo estimado del arribo del bus a la próxima estación y se presenta la información mediante una pantalla y una aplicación en su smartphone. Un sistema de información para el trasporte público utilizando GPS, el protocolo TCP/IP y otros servicios web se presenta en (Sungur, Babaoglu & Sungur, 2015). Mediante monitores en el interior de la unidad móvil y en las paradas, se puede conocer su ubicación en tiempo real. En (Handte, Foell, Wagner, Kortuem & Marrón, 2016) utilizando el sistema de posicionamiento global, los pasajeros pueden ver la ruta trazada desde el interior del bus. Como complemento se presenta una serie de recomendaciones sobre rutas alternas, dependiendo de los niveles de multitud que esperan un próximo abordaje.

3. Descripción de la propuesta

En este contexto, los estudios y algoritmos realizados están orientados a la formación de un sistema informativo, que mejore la movilidad de los usuarios al utilizar el transporte

público. En la Figura 1 se muestra un esquema general del sistema y los elementos que forman parte del mismo.

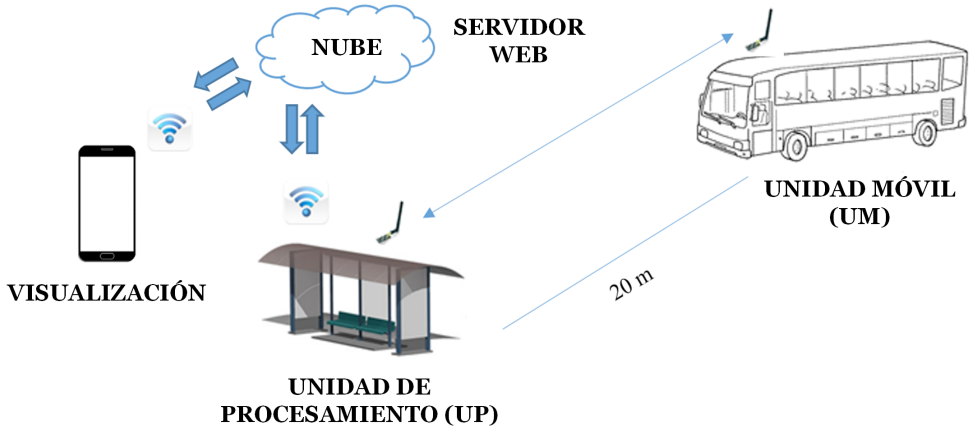


Figura 1 – Esquema general del sistema propuesto

3.1. Servidor web

Para iniciar el desarrollo de los algoritmos de programación, se instala el sistema operativo Raspbian en la tarjeta embebida Raspberry Pi 3. A continuación, se describe el proceso de instalación del servidor; que inicia con la actualización de librerías y repositorio, para evitar inconvenientes al momento de configurar y conectar periféricos. De esta manera, la instalación del servidor Apache puede empezar al introducir el comando de programación correspondiente. El directorio donde se guardan los archivos o páginas que se muestran en el servidor web de Raspbian es: /var/www/html. Aquí index.html es la página que aparece por defecto. Se instala el lenguaje de código abierto PHP por sus prestaciones en el desarrollo de aplicaciones web y se crea el fichero info.php en el directorio de la página web. Al abrir el navegador en la dirección donde se creó el fichero se verifica la correcta instalación de PHP en el servidor. Para el almacenamiento de la información se instala una base de datos MySQL. Se configura una nueva contraseña para el usuario root, que será utilizada para la autenticación de identidad. Se inicia el servicio y se comprueba su funcionamiento con los comandos respectivos.

Para la administración de la base de datos, se selecciona el servidor web que será utilizado y se genera una contraseña mediante phpMyAdmin. Se modifica el fichero de configuración de Apache “apache2.conf”, se guardan los cambios y se reinicia el servicio. Utilizando FTP se puede compartir archivos desde cualquier computadora que tenga los permisos y el acceso para el uso de este servicio, hacia el servidor web que se ha creado. Para permitir la escritura en el servidor es necesario modificar el fichero de configuración de FTP, así como el reinicio del servicio para actualizar los cambios efectuados. Finalmente, para la creación y administración del sitio web se utiliza WordPress. Esta herramienta necesita una base de datos para el almacenamiento de su información, por lo cual se procede a crearla, se configura el sitio y contraseña.

Funcionamiento

Los datos adquiridos son enviados a la base de datos (BD), procesados por el servidor y presentados a los usuarios finales. De esta manera se pueden efectuar consultas desde una computadora o un smartphone, como se muestra en la Figura 2. Para probar esta propuesta se han definido dos rutas de transporte público en un entorno urbano denominadas “La Joya” y “Ruta 2”, la Figura 3 muestra el recorrido de dichas rutas por las paradas llamadas “P24”, “P07” y “P29”.

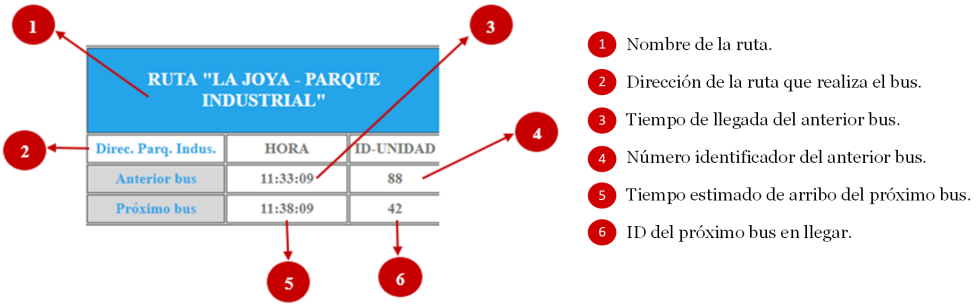


Figura 2 – Tabla informativa presentada al usuario final

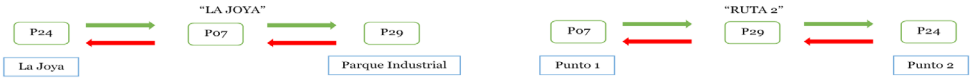


Figura 3 – Paradas que forman las rutas “La Joya” (izquierda) y “Ruta 2” (derecha)

El algoritmo que se ha desarrollado en el servidor para determinar el próximo arribo de una UM, se presenta en la Figura 4. Para explicar el proceso realizado, se toma como ejemplo la parada de bus P07. Al ingresar los datos de un bus perteneciente a una ruta que transita por esta parada. Éste es registrado y en la pantalla de información se muestra la hora de arribo y el ID correspondiente a la unidad móvil. Según la frecuencia de buses establecida por la cooperativa de transporte, el tiempo estimado de arribo del siguiente bus es visualizado en la parada. El tiempo dependerá del horario, tráfico y si es un día laborable o no. Para visualizar en el informativo el ID del próximo bus a llegar, se realiza una búsqueda en la tabla de datos correspondiente a la anterior parada, sea esta P24 o P29 (según la dirección de la ruta). Se busca el N_ID (número identificados de la UM) del bus recién llegado, para luego buscar el vehículo con la misma ruta y dirección registrado después. Obtenido el ID de la siguiente unidad registrada en la anterior parada, este se visualiza en el informativo de la parada o en el medio digital donde se realiza la consulta. Si no se encuentra un bus registrado después del actual en la anterior parada, en el informativo se visualiza el mensaje “No definido”. De esta manera el mismo algoritmo se puede utilizar en todas las paradas.

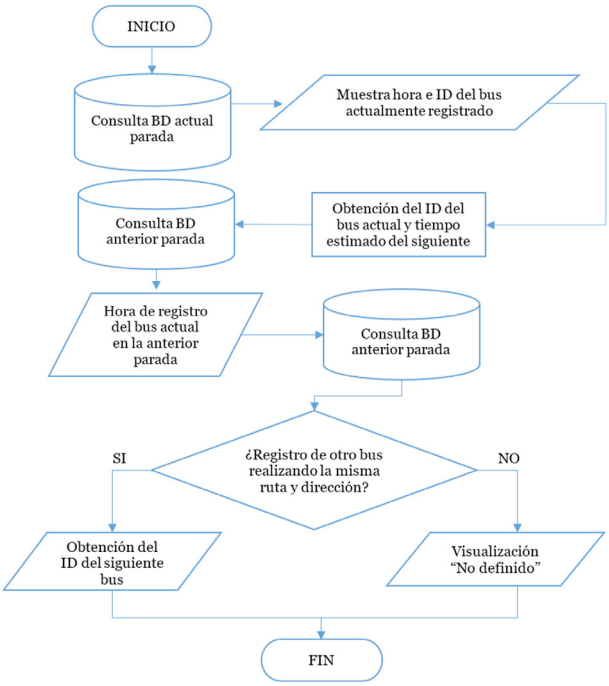


Figura 4 – Diagrama de flujo del algoritmo desarrollado en el servidor web

3.2. Unidad móvil

El operador de la unidad móvil, introduce los datos de la ruta a realizar y el sentido (norte - sur o viceversa) en el dispositivo implementado. Se establece comunicación entre dicho dispositivo y la UP automáticamente, mientras se encuentren dentro del rango establecido por el alcance de los módulos de transmisión inalámbrica. Es decir, no es necesaria la activación manual por parte del chofer. Cuando los parámetros han sido ingresados, el módulo de comunicación en la UM se dispone en modo espera y al verificar que hay comunicación envía los datos a la UP. Ésta envía un mensaje de confirmación que controla la reproducción de audio alojado en la memoria interna de la UM y activa un indicador que señala que el proceso de transmisión ha finalizado. Si la parada no pertenece a la ruta elegida, el proceso la cierra y la UM vuelve a estar a la espera de un enlace disponible para la comunicación. De este modo si el automotor ha transitado por algún lugar cercano a otra parada no correspondiente, que se encuentra dentro del rango de comunicación no le afectará en ningún sentido. La trama de información enviada está conformada por datos del tipo entero y sus parámetros son: la ruta, dirección, cooperativa y la Id de la unidad. La Figura 5 ilustra el proceso realizado por la UM.

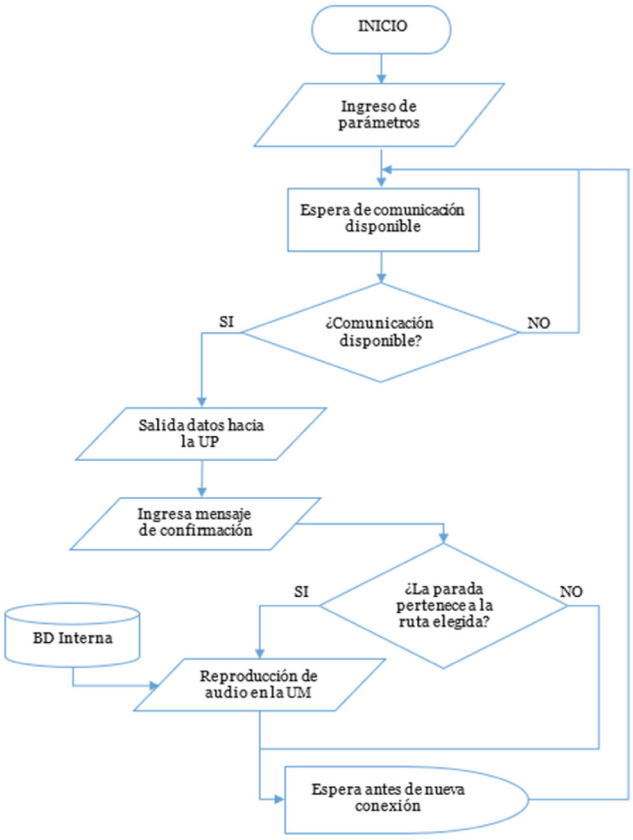


Figura 5 –Diagrama de flujo del algoritmo desarrollado en la unidad móvil

3.3. Unidad de procesamiento

Está ubicada en la parada de bus y se encuentra programada para recibir los parámetros enviados desde la UP inalámbricamente. Si la información es válida, la UP envía la información a la base de datos presente en el servidor web. De forma paralela envía un mensaje a la UM confirmando su validación, lo cual controla la reproducción de un audio informativo en su parada. El audio alojado en la memoria interna indica la parada actual en la que se encuentra el transporte y la próxima estación. En la pantalla informativa se indica el arribo del transporte y la ruta que realiza. En la Figura 6 se muestra el procedimiento descrito.

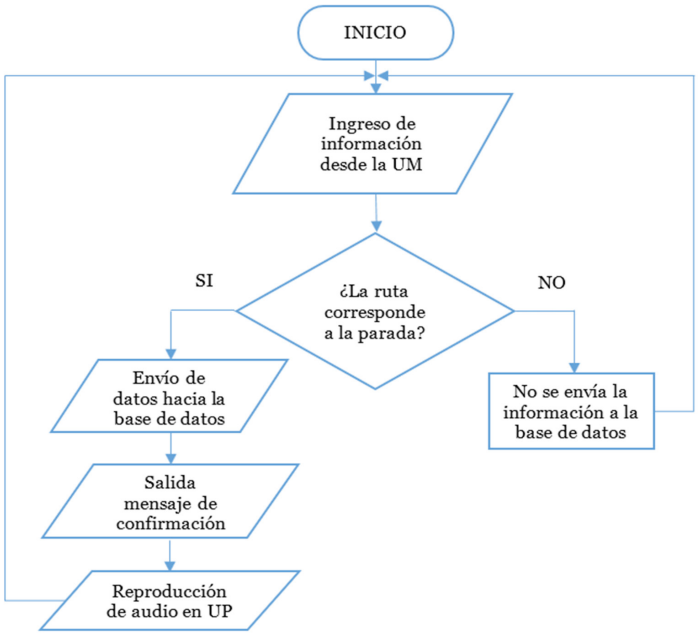


Figura 6 – Diagrama de flujo del algoritmo desarrollado en la unidad de procesamiento

3.4. Sitio web

El diseño se lo realizó mediante la herramienta WordPress, con una pantalla de inicio intuitiva y amigable con el usuario. Utilizando la herramienta Google Maps API, se implementa un mapa del entorno urbano y la ubicación de las paradas en la ruta establecida. Para elaborar este mapa se han tomado más de trescientas coordenadas con un dispositivo GPS. Al navegar por esta interface el usuario encuentra una descripción detallada de las paradas y rutas disponibles. Además, los horarios, el clima y otros parámetros que facilitan una mejor distribución del tiempo y un plan de viaje más eficiente. Dentro del menú principal del sitio web, se encuentran las opciones “Paradas Innovadoras”, “Consultas” y “Monitoreo y Control”. Al seleccionar la primera opción se muestra un mapa de la ruta y la distribución de estas paradas inteligentes, como se presenta en la Figura 7. En sentido sur – norte con color azul y en dirección opuesta con color violeta.

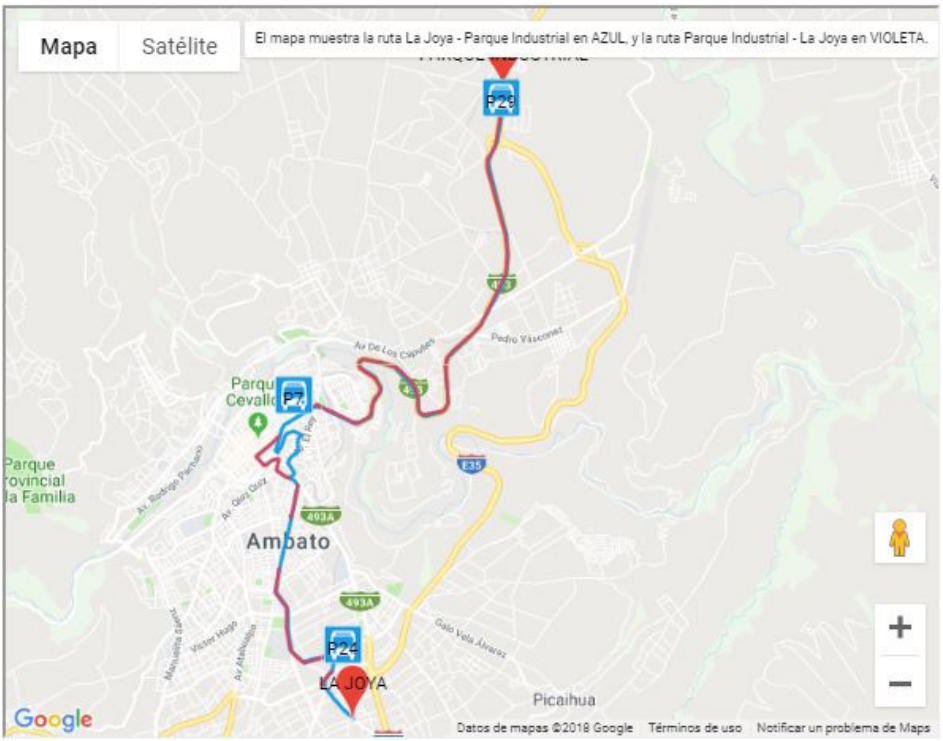


Figura 7 – Pantalla del sitio web donde se muestra el mapa de la ruta “La Joya”

3.5. Aplicación móvil

Fue desarrollada utilizando el software AppInventor, para celulares inteligentes que posean el sistema operativo Android. Según la página oficial de esta herramienta es compatible con todas las versiones de Android, aunque solo se probó con 5.1, 4.0 y 2.6. La aplicación permite al usuario realizar la consulta de los horarios del transporte público que circula por una cierta parada, mapas y rutas disponibles. La Figura 8 muestra la pantalla de inicio de la aplicación, la pantalla principal y el menú desplegable donde se selecciona la acción a realizar respectivamente. Para crear los mapas personalizados se utiliza la herramienta Google Maps Api, mediante programación en JavaScript. Para la generación del mapa “Ruta La Joya” fue necesario insertar en el código cada uno de los puntos (latitud, longitud) por el que transita la UM. La precisión del mapa depende de la cantidad de puntos tomados. Al combinar programación en JavaScript y AppInventor es posible el uso del GPS en el teléfono celular, donde se instala la aplicación para el trazo del mapa “Ubicación Paradas Innovadoras”.

Pantalla de inicio de la aplicación | Pantalla principal de la aplicación | Menú desplegable

Figura 8 – Pantallas que forman la aplicación desarrollada

4. Resultados

Los prototipos fueron implementados en dos unidades móviles (buses) y tres paradas (en ellas se encontraban los prototipos de las unidades de procesamiento), para la ejecución de las pruebas experimentales correspondientes. De esta manera se puede corroborar el correcto envío de datos en forma bidireccional, usando el módulo nRF24L01. Al no poseer acceso libre a las paradas de la localidad, estas unidades de procesamiento fueron instaladas en el interior de domicilios que se encuentran cerca de las paradas respectivas y dentro de la ruta que realizaban las unidades de transporte público.

4.1. Con línea de vista

Al enviar la información desde la UM hacia la UP, se lo realiza de manera efectiva y sin pérdida de paquetes a una distancia de hasta 20 metros. Cuando se transmiten los datos desde la UP hasta la UM la distancia máxima sin la presencia de interferencias es de 19 metros. Para el monitoreo de la recepción de paquetes se utilizó un smartphone con la aplicación “Mobile SSH”, que permite la comunicación con la consola del terminal del prototipo de la unidad de procesamiento. La Figura 9 muestra el terminal de Raspberry abierto en un celular y debajo el prototipo y en la Figura 10 la pantalla informativa presentada al usuario. El tiempo que tarda la comunicación entre los dispositivos es de 0,77 segundos.

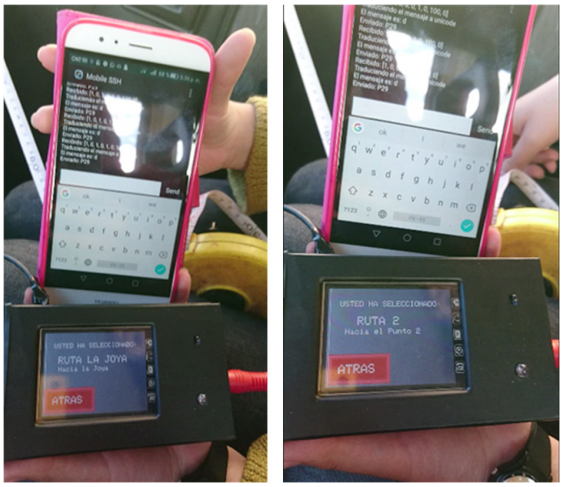


Figura 9 – Prueba del prototipo siguiendo las dos rutas establecidas

Figura 10 – Pantalla informativa presentada al usuario

4.2. Sin línea de vista

Para realizar estas pruebas se tiene como obstáculo un edificio (construcción de por medio). No existió pérdida de los paquetes enviados entre la UM hacia la UP hasta una distancia máxima de 15 metros. Si la distancia es mayor a dicho valor existe comunicación de forma intermitente.

4.3. Tiempo de respuesta del sistema

El tiempo que tarda el sitio web en actualizarse luego de recibir los parámetros respectivos desde la UP, varía desde de 1 hasta 3 y 4 segundos. Esta operación es ejecutada en la parte del cliente (servidor) e influyen factores como la calidad de conexión a internet que posee el usuario y su velocidad de transmisión. El tamaño de los ficheros que se utilizan

para el informativo presente en las paradas es de 5.8 MB en el momento de la primera carga. Para la actualización del informativo se utiliza 25.71 kilobytes y las pruebas de medición de su velocidad de carga fueron realizadas mediante la herramienta Pingdom. Se realizaron 256 peticiones y los resultados obtenidos se los detallan en la Tabla 1.

Tipo de contenido	Porcentaje	Tamaño	Porcentaje del número de peticiones requeridas	Número de Peticiones
Multimedia	89.4%	5.15MB	94.9%	243
HTML	8.7%	515.69KB	2%	5
Script	1.6%	93.18 KB	1.2%	3
Imágenes	0.2%	10.33 KB	1.2%	3
CCS (Estilos)	0.1%	5.84 KB	0.8%	2

Tabla 1 – Parámetros obtenidos en las pruebas de transmisión de la información

5. Conclusiones

El sistema informativo descrito en este trabajo, permite a los usuarios del transporte público obtener información sobre el servicio que están recibiendo. Esto contribuye a una mejor gestión y organización del tránsito urbano, ya que se ha desarrollado una propuesta tecnológica que ofrece buenas prestaciones a un costo reducido. La comunicación entre las unidades móviles (buses) y las unidades de procesamiento (paradas) fue satisfactoria utilizando módulos inalámbricos económicos, que permiten su escalabilidad. Los prototipos fueron implementados en 2 buses y 3 paradas, donde la velocidad de transmisión de datos obtenida es de hasta 2 Mbps, con un alcance de 20 metros con conexión automática y rápida con línea de vista directa. Mientras que cuando la línea de vista presente obstáculos el alcance máximo obtenido es de 15 metros. Se debe enfatizar que los módulos de comunicación empleados tienen un rango operativo superior (más de 500 metros), pero se han configurado de esta forma para evitar confusiones en la comunicación entre la UM y la UP respectiva. El tiempo que tarda la comunicación entre los dispositivos es de 0,77 segundos. La presentación de la información a los pasajeros se ha realizado a través de una interfaz intuitiva que muestra datos sobre la ruta que se está siguiendo, tiempo de espera para el próximo arribo, clima, entre otros. Además, para mayor facilidad del uso de este servicio se ha dotado al sistema de un sitio web, que tarda desde 1 hasta 3 y 4 segundos en actualizarse, dependiendo de la velocidad de conexión a internet existente. La aplicación es compatible con todo smartphone que posea sistema operativo Android y su efectividad también fue probada.

El trabajo que se ha presentado constituye una base para posteriores investigaciones que contribuyan a mejorar el monitoreo y control de los servicios públicos de países en desarrollo, que carecen de recursos económicos. Es por ello que inicialmente como trabajo futuro se plantea la implementación de esta propuesta en más unidades de transporte, lo cual permitirá obtener mayores datos y por ende mejores resultados. Además, se emplearán protocolos que permitan elevar el nivel de seguridad del sitio web y de la aplicación móvil para proteger los datos generados. De forma paralela, se propone desarrollar soluciones similares en otros campos de la ciencia como la agricultura y la industria, ya que los sistemas de monitoreo van tomando relevancia en la sociedad actual.

Referencias

Abdelsalam, H. M., Nazri Borhan, M., Milad, A., Imhimmed, H., Alfakhri, A.Y. Y., Alshetwia, A. B.Z. & Atiq Ok Rahmat, R. (2018). Web-Based Knowledge Evolution for Thinking Green Transportation Using Expert Systems. *International Journal of Engineering & Technology, 7. 2*, 837-839. doi: https://doi.org/10.14419/ijet.v7i2.29.14267

Arciniegas Aguirre, S. M., Mafla Rivadeneira, D. R., Guerra Torrealba, L. R., Narváez Erazo, L. D., Rivas-Echeverría, F. I. & Rivero Albarran, D. M. (2018). Metodología para el almacenamiento y visualización de datos masivos en invernadero basado en el Internet de las Cosas. *RISTI - Revista Iberica de Sistemas e Tecnologias de Informacao, E15*, 1-12.

Bie, Y., Gong, X. & Liu, Z. (2015). Time of day intervals partition for bus schedule using GPS data. *Transportation Research Part C, 60*, 443-456. doi: https://doi.org/10.1016/j.trc.2015.09.016

Guo, Z., Derian, A. & Zhao, J. (2015). Smart Devices and Travel Time Use by Bus Passengers in Vancouver, Canada. *International Journal of Sustainable Transportation, 9.5*, 335-347. doi: https://doi.org/10.1080/15568318.2013.784933

Handte, M. Foell, S., Wagner, S., Kortuem, G. & Marrón, P. J. (2016). An Internet-of-Things Enabled Connected Navigation System for Urban Bus Riders. *IEEE Internet of Things Journal, 3. 5*, 735-744. doi: https://doi.org/10.1109/JIOT.2016.2554146

Harris, N. (2017). The Design and Development of Assistive Technology. *IEEE Potentials, 36*.1, 24-28. doi: https://doi.org/10.1109/MPOT.2016.2615107

Johnson, R. D., Lukaszewski, K. M. & Stone, D. L. (2016). The Evolution of the Field of Human Resource Information Systems: Co-Evolution of Technology and HR Processes. *Communications of the Association for Information Systems, 38*, 533-553. doi: https://doi.org/10.17705/1CAIS.03828

Lee, S., Tewolde, G. & Kwon, J. (2014). Design and Implementation of Vehicle Tracking System Using GPS/GSM/GPRS Technology and Smartphone Application. *2014 IEEE World Forum on Internet of Things (WF-IoT), Seoul, South Korea*, 353-358. IEEE. doi: https://doi.org/10.1109/WF-IoT.2014.6803187

Mazloumi, E., Currie, G. & Rose, G. (2010). Using GPS Data to Gain Insight into Public Transport Travel Time Variability. *Journal of Transportation Engineering, 136.7*, 623-631. doi: https://doi.org/10.1061/(ASCE)TE.1943-5436.0000126

Megalingam, R. K., Raj, N., Soman, A., Prakash, L., Satheesh, N. & Vijay, D. (2014). Smart, public buses information system. *2014 International Conference on Communication and Signal Processing, Melmaruvathur, India*, 1343-1347. IEEE. doi: https://doi.org/10.1109/ICCSP.2014.6950068

Shakir, M., Hammood, M. & Muttar, A. Kh. (2018). Literature review of security issues in saas for public cloud computing: a meta-analysis. *International Journal of Engineering & Technology, 7.3*, 1161-1171. doi: https://doi.org/10.14419/ijet.v7i3.13075

Siła-Nowicka, K., Vandrol, J., Oshan, T., Long, J. A. Demšar, U. & Fotheringham, A. S. (2016). Analysis of human mobility patterns from GPS trajectories and contextual information. *International Journal of Geographical Information Science, 30*, 881-906. doi: https://doi.org/10.1080/13658816.2015.1100731

Silva, C. A., Toasa, R., Guevara, J., Martinez, H. D. & Vargas J. (2018). Mobile Application to Encourage Local Tourism with Context-Aware Computing. In: Rocha Á., Guarda T. (eds) Proceedings of the International Conference on Information Technology & Systems (ICITS 2018). ICITS 2018. *Advances in Intelligent Systems and Computing, 721*. Springer, Cham, 796-803. doi: https://doi.org/10.1007/978-3-319-73450-7_75

Stoll, N. B.; Glick, T. & Figliozzi, M. A. (2016) Using High-Resolution Bus GPS Data to Visualize and Identify Congestion Hot Spots in Urban Arterials. *Transportation Research Record, 2539*, 20-29. doi: https://doi.org/10.3141/2539-03

Sungur, C., Babaoglu, I. & Sungur, A. (2015). Smart Bus Station-Passenger Information System. *2015 2nd International Conference on Information Science and Control Engineering (ICISCE), Shanghai, China*, 921-925. IEEE. doi: http://doi.ieeecomputersociety.org/10.1109/ICISCE.2015.209

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 21/10/2016
Aceitação/Acceptance: 26/11/2016

Recuperación de Datos en Dispositivos de Almacenamiento SSD Utilizando File Carving

Geovanni Ninahualpa Quiña[1,2,3], Sang Guun Yoo[1,2], Teresa Guarda[4,1,5]

{gninahualpa, yyang}@espe.edu.ec, tguarda@gmail.com

[1] Centro de Posgrados, Universidad de las Fuerzas Armadas ESPE, Sangolquí, Ecuador.

[2] Departamento de Ciencias de la Computación, Universidad de las Fuerzas Armadas ESPE, Sangolquí, Ecuador.

[3] Facultad de Informática, Universidad Nacional de la Plata, La Plata, Argentina.

[4] Universidad Estatal Península de Santa Elena– UPSE, Av. La Libertad, Santa Elena, Ecuador.

[5] Algoritmi Centre, Minho University, Guimarães, Portugal.

Pages: 490–498

Resumen: Este documento tiene como objetivo presentar el estado del ar-te con respecto al tema de recuperación de datos en dispositivos de almacenamiento en SSD cuya pérdida de información ha causado daños físicos, son estos factores humanos o ambientales y propuesto los escenarios más comunes de las causas de esta pérdida golpes o humedad, también las técnicas de file carving se analizan de acuerdo con el escenario de pérdida de información, presentando como solución a este problema dos técnicas para la recuperación de datos que se utilizarán en diferentes casos, ya sea que el usuario del dispositivo sepa o no es el contenido de eso. Finalmente, se da a conocer el conocimiento actual en investigación sobre el tema.

Palabras-clave: File carving; semantic carving; fragment recovery carving; dispositivos SSD con riesgo de daño físico.

Data Recovery on SSD Storage Devices Using File Carving

Abstract: This document aims to present the state of the art regarding the subject of data recovery in SSD storage devices whose information loss has been caused physical damage are these human or environmental factors and proposed the most common scenarios of the causes of this loss being these blows or humidity, al-so the techniques of file carving are analyzed according to the scenario of loss of information, posing as solution of this problem two techniques for the recovery of data that will be used in different cases, whether the user of the device knows or not the content of it. Finally, the current knowledge on research on the subject is made known.

Keywords*:* File carving; semantic carving; fragment recovery carving; SSD devices with risk of physical damage.

1. Introducción

Durante las últimas décadas, el valor de la información ha cobrado mucha importancia, el término "sociedad de la información" se ha introducido para referirse a la sociedad actual (Poisel, Tjoa, & Tavolato, 2011). La creación, manipulación, distribución y uso de la información se han convertido en actividades fundamentales tanto en un ámbito empresarial como personal. Esta información se encuentra en distintos dispositivos de almacenamiento como memorias flash (USB), tarjetas de memoria (micro SD), discos duros de estado sólido (SSD), discos duros electromecánicos (HDD), entre otros.

Existe un índice de los años 2008-2009 que revela el comienzo del uso comercial masivo de discos duros basados en la tecnología de estado sólido (SSD), que ahora están reemplazando rápidamente los discos duros electro-mecánicos (HDD) (A & Tijan, 2016).

Los dispositivos de almacenamiento en estado sólido SSD utilizan la tecnología de almacenamiento de memoria flash, presentando una solución de almacenamiento para aplicaciones de uso intensivo de datos, debido a su baja latencia, alto ancho de banda y bajo consumo de energía en comparación con los discos tradicionales (Xie, Chen, & Roth, 2016).

Los fallos en el disco duro son la principal causa de pérdida de datos a nivel mundial, según encuestas un 72% de personas afirman que tuvieron un caso en su PC o portátil, seguidos por un 15% que sufrieron la pérdida en discos SSD y por un 13% en servidores RAID y virtuales, lo que evidencia que la inaccesibilidad en la información puede producirse en un entorno profesional y personal (ComputerWorld, 2014).

En la tabla 1 se observa que gran cantidad de personas a nivel mundial, han sufrido daños en los distintos tipos de almacenamiento siendo los más comunes los dispositivos de almacenamiento fijo o discos duros internos.

Item	Daños en dispositvos de almacenamiento Tipo	Pocentaje
1	*PC o portátil (Discos Duros Fijos)*	*72%*
2	*Discos duros de estado sólido*	*15%*
3	*Servidores RAID y virtuales*	*13%*

Tabla 1 – Daños en dispositivos de almacenamiento

En el año 2014 los ingenieros de Kroll Ontrack identificaron como causa de la pérdida de información de los dispositivos de almacenamiento SSD a los errores humanos con un 14% (ComputerWorld, 2014). En el caso de factores ambientales como la humedad, el uso de un dispositivo de almacenamiento SSD o cualquier medio de almacenamiento fuera de los intervalos ambienta-les recomendados puede afectar al ciclo de vida del dispositivo e incluso provocar una pérdida de datos (Kingston, 2018).

Actualmente los dispositivos de almacenamiento SSD, tanto internos como externos, pueden perder la información por distintos factores ya sean huma-nos o ambientales. Por este motivo, una correcta recuperación de datos se vuelve indispensable por el valor de la información que contienen.

En la tabla 2 se proponen las causas más comunes por los cuales un dispositivo de almacenamiento SSD tanto interno como externo sufre un daño físico.

Item	Causas de daños físicos Factores	Causa 1	Causa 2
1	Humanos	Golpes	Humedad
2	Ambientales	Humedad	Exposición al calor

Tabla 2 – Daños en dispositivos de almacenamiento

En respuesta a la ocurrencia de estos eventos, File Carving se ha convertido en una técnica esencial tanto para la recuperación de datos generales como para las investigaciones forenses digitales. Esta técnica consta de 3 etapas que se detallan a continuación (Poisel & Tjoa, 2013):

- Recuperación de datos basados en metadatos: se basan en el sistema de archivos que se está utilizando ya que los datos borrados y el orden de bloques de datos pueden ser preservados durante el proce-so de eliminación de archivos;
- Los metadatos del sistema de archivos contienen información sobre el tamaño de bloque utilizado. Los tamaños de bloques más grandes aumentan la precisión durante la fase de clasificación de fragmentos del proceso de file carving. Debido a que los diferentes sistemas de archivos implementan diferentes métodos para organizar los datos, en esta etapa, es de gran importancia tener conocimiento sobre el sistema de archivos desplegado ya que revela información sobre el comportamiento de fragmentación;
- Clasificación de fragmentos de archivos: Después del paso de pre-procesamiento, los bloques restantes que no se pueden asociar con un archivo se reenvían al proceso de clasificación de fragmentos de archivo. Dependiendo del método de reensamblado elegido para reordenar dichos fragmentos, se determina el tipo de fragmento para reducir aún más la cantidad de datos. De esta manera se da paso al reensamblado donde se identifican los bloques que pertenecen a los archivos específicos destinados a su recuperación;
- Reensamblaje de fragmentos: Los bloques que han sido clasificados de acuerdo a su tipo se reensamblan en su orden correcto. El resultado de esta etapa es el archivo original o, en el caso de piezas faltantes, un archivo parcialmente ensamblado.

En tal sentido el presente trabajo tiene como objetivo y debido a la poca investigación realizada en torno a las causas de pérdida de información en dispositivos SSD por daños físicos, determinar las técnicas apropiadas que se deben utilizar de acuerdo al caso.

Para cumplir con este cometido se propone dos escenarios (pérdida de datos por impacto ó golpe y pérdida de datos por humedad) en los cuales se analizará la causa de pérdida de información, el tipo de archivos que contiene y las técnicas que se utilizarán.

Así también, las técnicas de file carving propuestas para la recuperación de datos en los escenarios descritos dependen del nivel de conocimiento de la estructura y características de la información contenida en el dispositivo. Para el efecto a continuación se describe los tipos de dispositivos de almacenamiento.

1.1. Dispositivo de almacenamiento móvil

Los dispositivos de almacenamiento móvil o externos son fácilmente manipulables y pueden ser transportados de una computadora a otra de forma sencilla y al encontrarse de forma externa a la estructura del computador, el dispositivo es propenso de sufrir algún tipo daño físico por golpes, o por humedad debido a que sobre él se podrían derramar distintas sustancias líquidas como agua, café, gaseosa, etc.

Estos escenarios evidencian mayores daños físicos por factores humanos antes que factores ambientales.

1.2. Dispositivo `de almacenamiento fijo

Los dispositivos de almacenamiento fijo o interno al encontrarse dentro de la estructura del computador almacenan una gran cantidad de polvo y hume-dad que ocasionan desperfectos en los dispositivos mencionados (Poisel & Tjoa, 2013).

Este escenario evidencia mayores daños físicos por factores ambientales antes que factores humanos.

2. Técnicas de Recuperación File Carving

File Carving es una técnica que permite la recuperación de archivos sin cono-cimiento de información contextual, como los metadatos del sistema de archivos.

Debido a los avances en la investigación sobre estas técnicas, el File Carving se ha convertido en una técnica esencial para la recuperación de datos generales e investigación forense digital (Poisel & Tjoa, 2013).

La clasificación de las técnicas file carving según (Constanzo & Waimann, 2012) y (Ninahualpa, Díaz, & Guun, 2017) se muestran a continuación.

2.1. Técnicas File Carving basadas en características del archivo sin información del filesystem

Técnicas Header Footer Files: Estas técnicas basan su funcionalidad en dos identificadores “Encabezado” y “Pie/Fin”, pues cuando este se encuentra presente se conoce el inicio y el final, de modo que los bloques útiles ocupa-dos son considerados como el archivo a restaurar (Hadi, 2016; Povar, 2011).

Técnicas File Structure and Block Content: Estas técnicas basan su funcionalidad en dos partes, la estructura del archivo a restaurar y los bloques útiles contenidos en él (Alherbawi, Shukur, & Sulaiman, 2016; Garfinkel, 2007).

File Structure Based Carving: Esta técnica analiza las estructuras internas del tipo de archivo, además de las que se encuentran en la cabecera (Constanzo & Waimann, 2012),(Garfinkel & Metz, 2018).

Semantic Carving: La técnica semantic carving identifica el idioma utilizado en un bloque, y lo relaciona con bloques en el mismo idioma de forma coherente para reconstruir un texto (Constanzo & Waimann, 2012),(Garfinkel & Metz, 2018).

2.2. Técnicas File Carving Basadas en la fragmentación del archivo sin información del filesystem

Técnicas Fragmentation Issue: Este tipo de técnicas se emplean con el fin de restaurar un archivo que se encuentra fragmentado en el dispositivo de al-macenamiento. En este sentido estas técnicas no cuentan con: identificado-res o cadenas de secuencia para reconocer al fragmentov (Hadi, 2016), (Garfinkel & Metz, 2018).

Técnicas Predictive File: Este tipo de técnicas fundamenta su acción en la creación de nueva metadata que referencie a los bloques de almacenamiento válidos, utilizando un sistema de archivos (Filesystem) virtual como forma de predicción y acceso a los mismos (Alherbawi, Shukur, & Sulaiman, 2016).

Técnicas Fragment Recovery Carving: La técnica fragment recovery carving es utilizada al describir cualquier método de carving en el que dos o más fragmentos se reensamblan para formar el archivo u objeto original [9], don-de se descartan los bloques intermedios a los fragmentos que no pertenecen al mismo (Garfinkel & Metz, 2018).

Statistical Carving: Esta técnica se calcula en función matemática de los bloques buscando similitudes entre ellos para reconstruir los archivos. Algunas variantes usan funciones estadísticas, hashes, y/o signatures, entre otras (Constanzo & Waimann, 2012), (Garfinkel & Metz, 2018).

El esquema de la figura 1, presenta de mejor forma la clasificación mencionada para la recuperación de datos. El proceso de restauración de datos (Ninahualpa, Díaz, & Guun, 2017) indica la existencia de metodologías y técnicas que dependen del tipo de daño, tipo de dato a recuperar y los facto-res que afectan en el proceso, considerando las técnicas basadas en carving y file carving como las técnicas que aportan de forma eficiente en el proceso de recuperación de datos en dispositivos de almacenamiento.

Como se mencionó anteriormente los factores humanos y ambientales son la principal causa de daños físicos que provocan los daños lógicos que impedirán el acceso a la información, en tal sentido se proponen las técnicas semantic carving en caso de conocer el contenido del dispositivo siendo este texto y fragment recovery carving al no conocer el contenido del dispositivo.

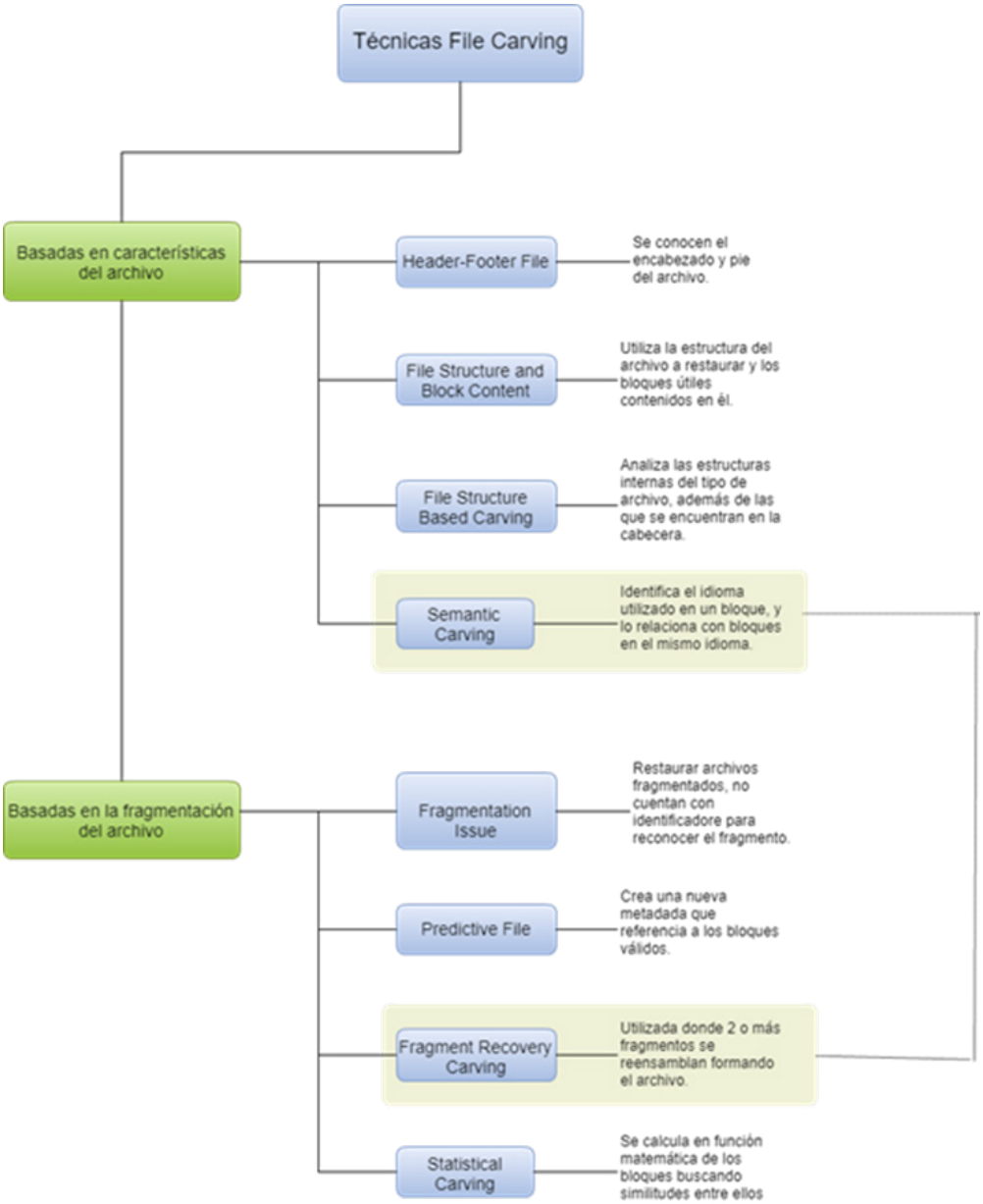


Figura 1 – Recuperación de datos según el contenido y tipo de daño

3. Proceso de Selección de Técnicas para Recuperación de Datos

La propuesta para seleccionar la técnica de recuperación de información, file carving en dispositivos de almacenamiento SSD con riesgo de daño físico según escenarios descritos y conocimiento o no del contenido del dispositivo tiene como fin el brindar la orientación sobre las herramientas software forenses que más se ajusten a requerimientos y alcancen mayor eficiencia en este proceso.

El proceso parte con la identificación del tipo de dispositivo SSD (móvil o fijo) con riesgo de daño físico. Se propone que, en caso de ser fijo, existe mayor probabilidad de que el daño haya sido causa de factores ambientales ante los factores humanos. Por otro lado, en caso de ser un dispositivo móvil, hay mayor probabilidad de que el daño haya sido causado por factores humanos ante los factores ambientales.

En caso de los dispositivos SSD que han sido afectados por factores humanos los escenarios más comunes son golpes o impacto. A diferencia de, los daños por factores ambientales cuyos escenarios más comunes son humedad y exposición al calor.

Si la causa de daños en los dispositivos no corresponde a un factor humano ni ambiental, entonces existe un deterioro normal de la vida útil del dispositivo.

Estos escenarios descritos generan daños lógicos en los dispositivos por lo que se plantea la recuperación de datos utilizando técnicas de file carving.

A continuación se realizará una verificación del tipo de contenido y estructura de la información. Así se discriminará el uso de semantic carving (usuario conoce estructura y características de la información – formato de texto) o fragment recovery carving (usuario desconoce estructura y características de la información – distinto formato).

Como resultado de la aplicación de estas técnicas se obtendrá la metadata y la data del objeto recuperado.

4. Conclusiones

Con el estado del arte realizado se ha evidenciado que no existe información suficiente ni detallada acerca de la pérdida de información en dispositivos de almacenamiento SSD por daños físicos en los ambientes propuestos.

Las principales causas de pérdida de información en dispositivos de almacenamiento SSD tanto fijos como móviles son los factores humanos y ambientales.

Queda evidenciado que las investigaciones realizadas sobre la recuperación de datos se enfocan en dar soluciones y técnicas sobre la recuperación sin analizar las causas que han llevado a la pérdida de la misma, debido a este desconocimiento de las causas no se utilizan las técnicas adecuadas.

Las técnicas de file carving no dependen del tipo de sistema de archivos para la recuperación de datos, lo que lo hace independiente del sistema operativo, además la recuperación de los metadatos es una ventaja frente a otras técnicas que no lo ofrecen.

Los fallos en los discos duros independiente si son HDD o SSD, son la principal causa de pérdida de datos a nivel mundial en un entorno empresarial o personal.

Esta investigación es un punto de partida hacia un campo que ha sido olvidado: el conocimiento las cusas de daños físicos en dispositivos de almacenamiento SSD. De esta manera se espera que aumenten investigaciones sobre este tema.

Agradecimientos

Agradezco al Departamento de Ciencias de la Computación de la Universidad de las Fuerzas Armadas- ESPE de la República del Ecuador, así como al Centro de Posgrados de la Universidad Nacional de la Plata de la República Argentina por su gran aporte y apoyo en la realización de este trabajo.

Referencias

A, S. K., & Tijan, E. (2016). Effectiveness analysis of using Solid State Disk tech-nology. *IEEE Xplore*, 1587–1591.

Alherbawi, N., Shukur, Z., & Sulaiman, R. (2016). A Survey on Data Carving. *Digital Forensic. Asian Journal of Information Technology*, 137–5144.

ComputerWorld. (2014). Los fallos en el disco duro, causa principal de la pérdida de datos en España. *ComputerWorld from IDG*. Retrieved from http://www.computerworld.es/tendencias/los-fallos-en-el-disco-duro-causa-principal-de-la-perdida-de-datos-en-espana

Constanzo, B., & Waimann, J. (2012). El estado actual de las Técnicas de File Carving y la necesidad de Nuevas Tecnologías que implementen Carving Inteligente. *Journal CADI*.

Garfinkel, S. (2007). Carving Contiguous and Fragmented Files with Object Validation. *The Digital Forensic Research Conference.*, (pp. 2–12). doi:doi:10.1016/j.diin.2007.06.017

Garfinkel, S., & Metz, J. (2018). *File Carving*. Retrieved from http://www.forensicswiki.org/wiki/File_Carving.

Hadi, A. (2016). Reviewing and Evaluating Existing File Carving. *2016 Cybersecurity and Cyberforensics Conference* (pp. 55–59). Jordan: IEEEE.

Kingston. (2018). *Uso de las SSD en sistemas integrados y destinados a un uso específico*. Kingston Technology. Retrieved from https://www.kingston.com/es/ssd/resources/embedded_purpose_built

Ninahualpa, G., Díaz, J., & Guun, S. (2017). Data Restoration and File Carving. *EEE Xplore*, 1–5.

Poisel, R., & Tjoa, S. (2013). A Comprehensive Literature Review of File Carving. *IEEE Xplore*, 475–484.

Poisel, R., Tjoa, S., & Tavolato, P. (2011). Advanced file carving approaches for multimedia files. *JoWUA, 2(4)*, 42–58.

Povar, D. (2011). Forensic Data Carving. *Digital Forensic and Cyber Crime*, 137-148.

Xie, W., Chen, Y., & Roth, P. C. (2016). Parallel-DFTL: A flash translation layer that exploits internal parallelism in solid state drives. *2016 IEEE International Conference on Networking, Architecture and Storage (NAS)* (pp. 1–10). IEEE.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 14/10/2018
Aceitação/Acceptance: 28/11/2018

Monitorização cognitiva longitudinal numa comunidade: estudo de implementação em contexto real

Ana Isabel Martins[1], Isabel Araujo[2], Andreia Sousa[2], Catarina Ruano[2], Joana Pais[2,3], Sara Félix[2], Eduardo Conde[2], Luís Ruano[3], Alexandra Queirós[1], Vítor Tedim Cruz[3,4], Nelson Pacheco Rocha[1]

anaisabelmartins@ua.pt, isabelaraujo@ua.pt, andreialeitesousa@gmail.com, c.rebelo.ruano@gmail.com, joanarpais@gmail.com, paiva.psicologia@gmail.com, eduardoconde.y@gmail.com, lmruano@gmail.com, alexandra@ua.pt, vtedimcruz@gmail.com, npr@ua.pt

[1] Universidade de Aveiro, Instituto de Engenharia Eletrónica e Informática de Aveiro (IEETA), Campus Universitário de Santiago, 3810-193, Aveiro, Portugal

[2] Neuroinova, Lda., Av. Sacadura Cabral 1694, 4410-463 Vila Nova de Gaia, Portugal

[3] ISPUP - Instituto de Saúde Pública da Universidade do Porto, Rua das Taipas 135, 4050-091 Porto, Portugal

[4] Hospital Pedro Hispano, ULS de Matosinhos, Rua Dr. Eduardo Torres, 4464-513 Senhora da Hora, Portugal

Pages: 499–512

Resumo: A monitorização cognitiva assumiu maior relevância considerando as doenças do sistema nervoso central que, devido ao aumento da esperança de vida e ao envelhecimento populacional, apresentam uma prevalência crescente. O estudo experimental reportado no presente artigo teve como objetivo a avaliação da usabilidade e da viabilidade do *Brain on Track* como um instrumento auto-aplicado para a monitorização cognitiva à distância, para o que foram envolvidos 1020 cidadãos do município de Águeda, Portugal. Do estudo pode-se concluir que o *Brain on Track* apresenta boa usabilidade e é adequado a estratégias de monitorização cognitiva de grandes populações.

Palavras-chave: Monitorização Cognitiva; *Brain on Track*.

Long term cognitive monitoring in a community setting: a real-world implementation study

Abstract: Cognitive tracking gains relevance considering central nervous system diseases, which, due to increased life expectancy and population ageing, have a growing prevalence. The present article reports an experimental study aiming to evaluate the usability and the viability *of Brain on Track* as a self-administrated instrument for remote cognitive tracking, for which were involved 1020 citizens of the municipality of Águeda, Portugal. From the study, it can be concluded that the *Brain on Track* presents good usability and is viable for implementing cognitive tracking of large populations.

Keywords*:* Cognitive tracking; *Brain on Track*.

1. Introdução

A deterioração cognitiva é uma consequência de muitas doenças do sistema nervoso central e tem-se tornado mais frequente com o aumento da esperança de vida e o envelhecimento da população a nível global (Wimo, Jönsson, Bond, & Prince, 2013). Doenças como a doença de Alzheimer, a doença de Parkinson ou a demência cerebrovascular afetam a vida de milhões de doentes e famílias, tendo um grande impacto social e económico nas sociedades ocidentais.

Um dos maiores desafios na prática clínica associada a estas doenças reside em distinguir o normal declínio cognitivo associado à idade do início da deterioração cognitiva associada a uma situação patológica, porquanto a identificação precoce dos défices cognitivos pode ser crucial para guiar a intervenção terapêutica, o treino cognitivo, a reabilitação funcional, ou a preparação de cuidadores informais, bem como para antecipar necessidades sociais (Graff, Vernooij-Dassen, & Thijssen, 2006; Gitlin, Kales, & Lyketsos, 2012; Woods & Aguirre, 2012; Bucci, Piegari, Mucci, & Merlotti, 2013).

O presente artigo reporta um estudo experimental que implementou uma solução de monitorização cognitiva à distância na população do município de Águeda, Portugal, utilizando uma plataforma web especialmente desenvolvida, o *Brain on Track* (Ruano et al., 2016)however the few available tools still depend on an external evaluator and periodic visits. We developed a self-administered computerized test intended for longitudinal cognitive testing (Brain on Track. O *Brain on Track* permite a monitorização longitudinal de declínios cognitivos ligeiros, em contraponto com o método padrão que se baseia na aplicação de uma bateria extensa de testes neuropsicológicos por um profissional clínico, o que, dado o tempo e os recursos exigidos, dificilmente é aplicável para seguimento de um grande número de indivíduos.

2. Trabalho Relacionado

A individualidade de cada ser humano resulta, em larga medida, da expressão pública do funcionamento cerebral, principalmente através de manifestações cognitivas e comportamentais. As capacidades para aprender, evoluir e gerir todos os aspetos da vida quotidiana estão dependentes do normal funcionamento do cérebro. À medida que envelhecemos, a importância do nosso estado de saúde cerebral ou cognitiva assume-se como o fator determinante para a qualidade de vida (Cicerone et al., 2011).

As doenças do sistema nervoso central estão associadas a défices cognitivos, sendo responsáveis pelas maiores taxas de incapacidade e impacto global das doenças na sociedade (Behrman, Bowden & Rose, 2013).

Independentemente das suas causas originais (e.g. acidentes vasculares cerebrais, lesões cerebrais ou doenças neurodegenerativas), os défices cognitivos raramente recuperam de modo espontâneo ou completo (Vincent et al., 2009; Rebok et al., 2012) e, na verdade, não existem intervenções farmacológicas com um efeito significativo positivo confirmado (World Health Organization, 2007; Eaton et al., 2011; Kakuma et al., 2011). Adicionalmente, em geral, os cuidados de saúde cognitiva dependem de serviços

clássicos, que exigem encontros físicos com profissionais sediados em instituições distantes do doente e respetiva família (Cicerone et al., 2011).

Tendo em consciência o enorme impacto associado às doenças neurológicas, é hoje amplamente reconhecido que os serviços e recursos disponíveis ao nível dos sistemas de saúde são desproporcionalmente escassos (Cicerone et al., 2011; Behrman, Bowden & Rose, 2013) e incapazes de lidar com a pressão imposta pelo aumento da prevalência destas. Ao nível da organização, os processos de prestação de cuidados são identificados como pouco estruturados, com perdas de eficácia, e tardios na incorporação de novas tecnologias e na adaptação a alterações sociais e a aumentos na demanda gerados por pressões demográficas (Cicerone et al., 2011; Behrman, Bowden & Rose, 2013).

Neste contexto, existem, à escala internacional, problemas de sustentabilidade críticos para a gestão da saúde cognitiva das populações, não só em termos de garantia de intervenções adequadas para se obter efeitos funcionais relevantes, mas também em termos de existência de mecanismos eficazes para identificar precocemente todos os indivíduos em risco de declínio cognitivo.

Têm sido propostas ferramentas de monitorização cognitiva de baixo custo e com tempos de aplicação curtos, mas a maioria delas ainda dependem de um avaliador com um mínimo de treino (Milne & Culverwell, 2008; Lonie, 2009). Além disso, têm pouca capacidade discriminativa para o défice cognitivo ligeiro (Lonie, 2009) e para prever a sua progressão (Lischka, 2012), principalmente por causa das dificuldades na definição de valores de referência individuais clinicamente significativos.

A monitorização individual do desempenho cognitivo ao longo do tempo tem sido proposta como um método adequado para identificar declínio cognitivo (Petersen, Stevens, & Ganguli, 2001). No entanto, medições cognitivas repetidas só são viáveis em larga escala se realizadas com o mínimo de custo e inconveniência para os avaliados. Poucas ferramentas foram clinicamente validadas para esta finalidade, e as que existem continuam a exigir visitas clínicas periódicas e um avaliador treinado, o que é inconveniente para os indivíduos e implica custos adicionais. Além do mais, a repetição do mesmo teste ao longo tempo pode levar a efeitos de aprendizagem que podem prejudicar a sua capacidade discriminativa.

3. Sistema *Brain on Track*

O Sistema Integrado de Monitorização Cognitiva *Brain on Track* (Figura 1) possibilita o acompanhamento longitudinal do desempenho cognitivo através da realização de testes cognitivos seriados utilizando uma plataforma web. Consequentemente, apresenta uma série de vantagens, nomeadamente: i) as monitorizações são realizadas à distância, sem necessidade de deslocações; ii) os custos associados são baixos; iii) a avaliação seriada do desempenho cognitivo é uma estratégia robusta de aferição do real desempenho de cada pessoa porque são valorizadas todas as alterações em relação ao melhor nível de desempenho individual previamente registado; e iv) ao contrário dos modelos clássicos que assentam em avaliações transversais, é menos vulnerável a influências internas e externas que frequentemente alteram o desempenho cognitivo (Neuroinova, 2015).

Figura 1 – Interface Brain on Track.

A partir do seu computador, o participante acede à plataforma web que disponibiliza sessões de avaliação *Brain on Track* programadas para serem realizadas autonomamente. A cada sessão de avaliação é atribuída uma pontuação, que é registada numa base de dados centralizada.

A componente nuclear da plataforma web é constituída por um conjunto de exercícios desenvolvidos para a análise do desempenho ao nível de diversas funções cognitivas como, por exemplo, atenção, funções executivas, memória ou linguagem, e que foram inicialmente desenvolvidos para um programa de treino cognitivo à distância (Cruz et al., 2013; Cruz et al. 2014). Este conjunto de exercícios foi já clinicamente validado, tendo também sido demonstrada a sua fiabilidade teste-reteste num subgrupo de uma coorte de base populacional (Ruano et al., 2016)however the few available tools still depend on an external evaluator and periodic visits. We developed a self-administered computerized test intended for longitudinal cognitive testing (Brain on Track.

Como o *Brain on Track* foi planeado para ser utilizado de uma forma repetida, recorreu-se a elementos e sequências aleatórias para minimizar os efeitos da aprendizagem. Todos os subtestes incluem pelo menos um elemento aleatório em cada tarefa ou são compostos por várias tarefas semelhantes pré-definidas que são selecionadas aleatoriamente para cada tentativa. Por exemplo, no subteste puzzles, existem 40 quebra-cabeças alternativos semelhantes em termos de apresentação e grau de dificuldade, sendo a seleção e a ordenação de um item específico geradas aleatoriamente no início de cada tentativa. Além disso, os subtestes foram desenhados com vários graus de dificuldade, para serem adaptativos.

Cada subteste começa com um conjunto de instruções que são apresentadas no ecrã e reproduzidas em formato áudio por uma mensagem pré-gravada e tem uma duração

limite de dois minutos. Durante esse período, o avaliado deve realizar as instruções que foram apresentadas (e.g. pressionar a tecla de espaço sempre que o número inicial for exibido numa série de números aleatórios que vão aparecendo no ecrã a cada três segundos, ou, após a apresentação de um enunciado, selecionar a resposta correta de entre quatro escolhas possíveis a três perguntas específicas).

4. Métodos

O estudo experimental reportado neste artigo teve como objetivos avaliar a usabilidade do *Brain on Track* e a sua viabilidade como um instrumento de monitorização cognitiva à distância de populações saudáveis.

O estudo foi composto por duas fases. A primeira consistiu numa avaliação presencial e a segunda numa monitorização longitudinal à distância.

A avaliação presencial teve como principal objetivo perceber o estado de saúde cognitiva de cada participante, designadamente o grau de exposição a fatores de risco e problemas de saúde que possam vir a condicionar o seu desempenho cognitivo futuro. Deste modo, para cada participante, a equipa de investigação aplicou um Questionário Geral de Saúde, a Escala de Queixas Subjetivas de Memória (QSM), e um instrumento de avaliação cognitiva, o *Montreal Cognitive Assessment* (MoCA). Ainda na sessão presencial, cada participante teve a oportunidade de usar uma versão abreviada do teste *Brain on Track*, e preencher um instrumento de avaliação de usabilidade, o *Post-Study System Usability Questionnaire* (PSSUQ).

O Questionário Geral de Saúde é constituído por uma parte inicial com dados sociodemográficos do participante e de contacto e uma segunda parte referente à saúde do participante, onde são abordados assuntos como saúde cognitiva, toma de medicação, histórico clínico ou fatores de risco (e.g. hipertensão arterial, sedentarismo ou tabagismo).

A QSM (Schmand, Jonker, Hooijer, & Lindeboom, 1996) contém dez itens com uma pontuação que varia entre zero, na ausência de queixas, e três pontos, conforme a gravidade de eventuais queixas.

Por sua vez, o MoCA é um dos instrumentos de rastreio cognitivo mais usados nos protocolos de avaliação de diversos grupos clínicos (e.g. declínio cognitivo ligeiro, doença de Alzheimer, doença de Parkinson, demência cerebrovascular, ou tumores cerebrais) (Freitas, Simões, & Martins, 2010). Avalia diferentes domínios cognitivos, nomeadamente a função executiva, a capacidade visuo-espacial, a memória, a atenção, a concentração ou a memória de trabalho, e o seu tempo de administração é de aproximadamente dez a 15 minutos.

Finalmente, o PSSUQ é composto por 19 itens que visam abordar cinco características da usabilidade de um sistema: rápida realização da tarefa, facilidade de aprendizagem, qualidade da documentação, disponibilização de informação *online* e adequação funcional (Lewis, 2002; Rosa et al., 2015).

A utilização do *Brain on Track* na sessão presencial teve como objetivo permitir aos participantes um primeiro contacto com esta aplicação e reportarem as suas opiniões

relativamente à usabilidade da mesma, bem como o esclarecimento de eventuais dúvidas a fim de facilitar a sua futura utilização autónoma em ambiente domiciliar.

Após a sessão presencial, pretendia-se que o desempenho cognitivo dos participantes fosse monitorizado através do *Brain on Track*. Para isso, na semana seguinte à realização da sessão presencial, os participantes foram convocados por *Short Message Service* (SMS) e email para realizarem uma sessão do *Brain on Track* no seu computador pessoal, em casa, com duração de cerca de 24 minutos. Posteriormente, a cada três meses receberam um SMS e email para realizarem novas sessões *Brain on Track*. Além disso, os participantes sob monitorização tiveram a possibilidade de solicitar o agendamento de uma avaliação extraordinária sempre que acharam necessário.

Para a realização do estudo foi definida uma amostra de base comunitária, não aleatorizada, constituída por munícipes de Águeda que se voluntariaram a participar.

Por forma a angariar potenciais participantes a integrar a amostra, o estudo contou com o apoio da Câmara Municipal de Águeda, e foi divulgado nos meios de comunicação social local, em ações locais de disseminação e através de folhetos colocados em pontos chave da comunidade. Adicionalmente, foram contactadas diversas instituições ligadas à administração pública, à indústria, à educação e à área social do município por forma a promover a dinamização de sessões de apresentação e esclarecimento junto dos seus colaboradores e clientes. Finalmente, para as sessões presenciais, elementos da equipa de investigação deslocaram-se a diferentes locais considerados estratégicos no município de Águeda.

Todas as pessoas que demonstraram interesse em participar no estudo deveriam, para serem selecionadas, cumprir os seguintes critérios de inclusão: i) compreender o estudo; ii) conseguir utilizar a interface rato/computador de forma autónoma; iii) ter acesso a um computador com Internet; iv) ter pelo menos um ano completo de escolaridade; e v) assinar o consentimento informado. Em termos de exclusão, foram definidos os seguintes critérios: i) défice motor, visual ou de linguagem grave que impedisse a avaliação cognitiva; e ii) não saber ler.

Todas as pessoas que se voluntariaram receberam uma folha de informações e foram questionadas sobre a sua vontade em participar no estudo. Elementos da equipa de investigação forneceram esclarecimentos sobre os objetivos, recolha de dados, duração e fases do estudo. Os interessados em participar tiveram oportunidade de pedir esclarecimentos adicionais e foram informados sobre a possibilidade de desistirem em qualquer momento, sem nenhum prejuízo pessoal e sem a necessidade de justificação. Após a informação sobre o estudo, foi pedido a cada interessado que assinasse um consentimento informado, de acordo com a Declaração de Helsínquia. Posteriormente, cada participante recebeu um folheto informativo do *Brain on Track* contendo os dados que deveriam utilizar para aceder ao *Brain on Track* (i.e. nome de utilizador e palavra pass).

Durante o estudo, todos os participantes que apresentaram declínios cognitivos foram referenciados para acompanhamento especializado.

O estudo experimental reportado no presente artigo contou com a colaboração de elementos da Câmara Municipal de Águeda, Neuroinova, Universidade de Aveiro e do Instituto de Saúde Pública da Universidade do Porto, tendo sido submetido e aprovado

pela Comissão de Ética do Centro Hospitalar de Entre o Douro e Vouga e pela Comissão Nacional de Proteção de Dados.

5. Resultados

Mais de 20 instituições sediadas no município de Águeda envolveram-se, disponibilizando as suas instalações e recebendo elementos da equipa de investigação para a realização de eventos presenciais. Cada uma destas instituições, designadas embaixadoras da saúde cognitiva, acolheu eventos de divulgação e inclusão de cidadãos, em média, pelo menos por duas vezes, num total de mais de 50 sessões presenciais. Em cada sessão foram avaliados e incluídos, em média, 20 cidadãos (mínimo dois; máximo 40). Realizaram-se também sessões públicas em eventos que ocorreram no município de Águeda e cujos objetivos estavam alinhados com a promoção da saúde ou a divulgação do potencial das soluções tecnológicas.

5.1. Participantes

Em consequência dos diversos eventos de divulgação, no decorrer do estudo experimental foram incluídos 1020 cidadãos com ligação residencial, familiar ou profissional ao município de Águeda, dos quais 62.4% eram do género feminino e 37.6% do género masculino. Do total de 1020 inicialmente inscritos no estudo, 669 compareceram e completaram efetivamente toda a avaliação presencial. Destes 669 participantes, 64.0% eram do género feminino e 36.0% do género masculino.

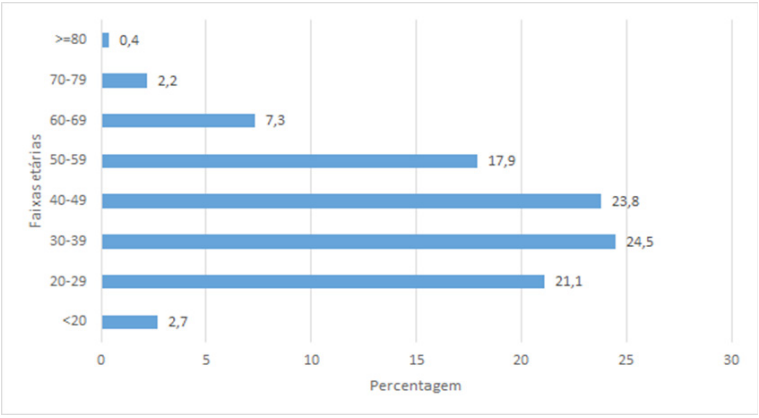


Figura 2 – Gráfico da distribuição dos participantes por faixa etária (em percentagem).

A média de idades foi de 41.15 anos (DP=13.75), sendo que o participante mais jovem tinha 18 anos e o mais velho 84 anos. O histograma com a distribuição por faixa etária dos participantes que realizaram a avaliação presencial encontra-se no gráfico da Figura 2.

No que diz respeito à escolaridade, a média de anos completos foi de 12.99 (DP=4.39), sendo a escolaridade mínima um ano e a máxima 27 anos. Salienta-se que a amostra foi na sua maioria constituída por pessoas com mais do que dez anos de escolaridade.

5.2. Fatores de Risco

De acordo com as respostas dos participantes ao Questionário Geral de Saúde, de entre os principais fatores de risco com impacto na saúde cognitiva da população destacou-se a hipertensão arterial que foi referida por 14.5% dos participantes e por mais de metade dos participantes acima de 60 anos. O sedentarismo foi reportado por cerca de 20.0% dos participantes e o tabagismo foi reportado por cerca de 15.0% dos participantes e a sua expressão foi maior em indivíduos abaixo dos 60 anos.

O sedentarismo, o tabagismo e a obesidade foram referidos desde escalões etários jovens. Por outro lado, nos escalões etários acima dos 60 anos verificou-se uma frequência relativamente elevada de diabetes, dislipidemia e arritmia. Cerca de 22.0% da população reportou ter um familiar direto com o diagnóstico prévio de alguma das formas possíveis de demência.

5.3. Cognição

Dos 669 participantes que realizaram a avaliação presencial inicial, 30.9% reportaram ter queixas acerca da sua cognição.

Relativamente aos resultados do teste MoCA, os resultados de 165 participantes não foram incluídos para análise (por não terem sido completados ou porque foram realizados por participantes com menos de 25 anos de idade, uma vez que existem apenas dados normativos para a população portuguesa a partir dos 25 anos). Assim, do subtotal de 504 avaliações válidas, a média da pontuação obtida foi de 25.24 (DP=3.67; mín.=8.00, máx.=30.00).

Da análise efetuada concluiu-se que o desempenho de 83.7% participantes se situava dentro dos parâmetros normais, considerando a idade e escolaridade e que 16.3% poderiam ter defeitos cognitivos pelo que, de acordo com o protocolo de investigação, foram referenciados para acompanhamento especializado.

5.4. Utilização do Brain on Track

Dos 669 participantes que realizaram a avaliação presencial inicial, 349 utilizaram o *Brain on Track* pelo menos uma vez nos seus domicílios. A média de idades destes 349 participantes foi de 40.63 anos (DP=12.80; amplitude: 18-75 anos) e a média de escolaridade foi de 14.23 anos (DP=3.71; amplitude: 4-27 anos).

No que respeita ao MoCA, os resultados obtidos pelos 349 (52,17%) participantes que utilizaram o *Brain on Track* pelo menos uma vez nos seus domicílios indicam uma média de 26.4 (DP=2.63). Para os restantes, a média dos resultados da aplicação do MoCA foi de 24.42 (DP=4.24).

As Tabelas 1, 2 e 3 apresentam o desempenho global do *Brain on Track* em função do género, da faixa etária e da escolaridade. Comparando os dados dos participantes que realizaram o *Brain on Track* nos seus domicílios com os que não realizaram verificou-se que os primeiros tinham, em média, idades inferiores, eram mais escolarizados, tiveram pontuações do MoCA mais elevadas e apresentaram mais queixas de memória. Contudo, as diferenças estatisticamente significativas entre os dois grupos estão apenas associadas à escolaridade e ao resultado do MoCA: o grupo de participantes que não realizou a

avaliação cognitiva remota tinha menor escolaridade (Média=11.64; DP=4.66) e os valores do MoCA foram inferiores (Média=24.42; DP=4.24).

Em termos de desempenho nos exercícios do *Brain on Track*, os resultados apresentaram uma média de respostas corretas de 218.41 (DP=57.75; amplitude: 73-361). O total de erros teve uma média de 10.34 (DP=11.91; amplitude: 0-145). Por outro lado, a análise destes resultados em função do género indicou que o género masculino apresentou mais respostas corretas (Média=246.35, DP=56.78) e menos erros (Média=9.81; DP=14.02) em comparação com o género feminino (Tabela 1). Esta diferença não é atribuível ao género, mas à diferença entre o grau de escolaridade ainda existente entre os dois géneros na amostra disponível para o estudo experimental.

Género	**Respostas corretas**		**Erros**	
	M	**DP**	**M**	**DP**
Feminino	205.69	53.79	10.51	10.81
Masculino	246.35	56.78	9.81	14.02

Tabela 1 – Desempenho geral no Brain on Track em função do género.

Faixas etárias	**Respostas corretas**		**Erros**	
	M	**DP**	**M**	**DP**
< 20 anos	279	41.88	10.22	5.19
20 - 29 anos	264.79	44.05	9.29	17.10
30 - 39 anos	239.45	53.04	8.56	10.31
40 - 49 anos	208.17	35.62	8.59	4.80
50 - 59 anos	172.31	45.24	14.56	14.76
60 - 69 anos	151.50	30.78	14.38	7.56
70 - 79 anos	145.20	34.72	12.80	4.76
Total	218.43	57.83	10.29	11.89

Tabela 2 – Desempenho geral no Brain on Track em função da faixa etária.

Faixas etárias	**Respostas corretas**		**Erros**	
	M	**DP**	**M**	**DP**
1 – 4 anos	174	56.02	14.11	7.06
5 - 9 anos	142.61	37.78	17.52	16.04
≥ 10 anos	227.35	53.30	9.45	11.27
Total	218.43	57.83	10.29	11.89

Tabela 3 – Desempenho geral no Brain on Track em função da escolaridade.

De acordo com os resultados no *Brain on Track*, considerando a idade e a escolaridade, verificou-se que as faixas etárias mais jovens são as que apresentaram melhores resultados (Tabela 2), assim como os indivíduos com níveis de escolarização mais elevados (Média=227.35; DP=53.30) (Tabela 3). A Tabela 4 apresenta as estatísticas descritivas do desempenho dos participantes no *Brain on Track* por domínio cognitivo.

5.5. Usabilidade do *Brain on Track*

A média da pontuação do PSSUQ para todos os participantes foi de 1.41 em 7.00 (DP=0.66), o que indica um nível de usabilidade e satisfação dos participantes muito elevado, pois, nesta escala, quanto menor o *score* maior o grau de usabilidade. O PSSUQ foi preenchido pelos participantes tendo em conta as suas opiniões sobre a sessão realizada presencialmente.

		Respostas corretas				**Erros**			
		M	**DP**	**Mín.**	**Máx.**	**M**	**DP**	**Mín.**	**Máx.**
Atenção	*Igual ou diferente*	11.92	4.16	2	37	0.98	1.24	0	9
Memória	*Cubos irrequietos*	11.21	3.71	0	25	0.64	1.17	0	7
	Longa Memória	17.69	3.44	4	20	0.31	0.77	0	5
	Rápida memória	10.51	2.68	1	17	0.17	0.45	0	3
Cálculo	*Mente* rápida	15.17	4.77	0	25	0.63	0.91	0	5
Funções executivas	*Combine a cor*	23.81	7.73	5	45	2.38	1.99	0	9
	Contrário	67.52	23.06	0	116	2.26	9.37	0	124
	Sequências	16.52	6.93	0	33	1.22	2.33	0	19
Linguagem	*Correspon-dências*	12.93	5.14	1	45	0.81	1.14	0	8
	Arrume as palavras	24.32	7.80	0	41	0.84	1.07	0	9
Capacidade construtiva	*Puzzles*	6.81	3.05	0	16	0.11	0.34	0	2

Tabela 4 – Estatísticas descritivas do Brain on Track.

Numa análise mais detalhada, as perguntas do PSSQ que obtiveram melhores resultados foram: i) *Foi fácil apreender a utilizar este sistema* (1.24 em 7.00); ii) *Consegui completar as tarefas e os cenários utilizando este sistema* (1.28 em 7.00); e iii) *Este sistema foi simples de utilizar* (1.31 em 7.00). Inversamente, os itens com *scores* mais baixos foram: i) *Consegui completar rapidamente as tarefas e os cenários utilizando este sistema* (3.40 em 7.00); ii) *Acredito que me tornaria rapidamente produtivo se utilizasse este sistema* (1.80 em 7.00); iii) *A informação fornecida pelo sistema (como ajuda online, mensagens no ecrã ou outra documentação) foi clara* (1.62 em 7.00); e iv) *Este sistema tem todas as funcionalidades e capacidades que eu esperava* (1.58 em 7.00).

5. Discussão e Conclusão

Neste artigo, foi reportado um estudo experimental que utilizou o *Brain on Track*, um Sistema Integrado de Monitorização Cognitiva baseado numa plataforma web, para a realização monitorização cognitiva à distância de grandes populações.

O *Brain on Track* é o resultado de um processo iterativo de desenvolvimento e foi alvo de várias avaliações e melhorias, nomeadamente em ambiente hospitalar, que culminaram num produto robusto, tal como comprova a avaliação da sua usabilidade. De acordo com o estudo experimental efetuado, os resultados da aplicação do PSSUQ revelaram um elevado nível de usabilidade.

O estudo exposto no presente artigo, considerando que se trata de uma implementação de um processo de monitorização cognitiva à escala de um município (o município de Águeda), corresponde à primeira iniciativa destinada a monitorizar o desempenho cognitivo de cidadãos, no seio da comunidade onde vivem. Tem, pois, a particularidade de ter sido realizado em contexto real e não em ambiente de investigação controlado ou instituições de saúde. Corresponde assim à primeira demonstração de que é possível acompanhar o desempenho cognitivo de cidadãos interessados, com um mínimo de literacia informática, que lhes permita utilizar um computador desde casa ou desde uma instituição comunitária que disponibilize esse acesso.

Os resultados do estudo são muito encorajadores e validam o *Brain on Track* para a realização de testes cognitivos longitudinais à distância em grandes populações.

O estudo terminou com um total de 1020 participantes incluídos, dos quais 349 realizaram pelo menos uma sessão do *Brain on Track* no domicílio. Relativamente à diferenciação entre os participantes que realizaram pelo menos uma sessão do *Brain on Track* no domicílio e os participantes que o não fizeram, verificou-se que as únicas diferenças estatisticamente significativas entre os dois grupos são a escolaridade e ao resultado do MoCA, sendo que o grupo de participantes que não realizaram qualquer sessão remota com o *Brain on Track* tinham uma menor escolaridade e obtiveram valores do MoCA inferiores.

Em termos de caracterização do perfil de fatores de risco com impacto no estado da saúde cognitiva da população a longo prazo, os resultados demonstram que: i) os participantes apresentaram uma frequência elevada de fatores de risco vasculares com impacto na saúde cognitiva e que podem vir a contribuir para o risco individual de doença de Alzheimer ou de demência cerebrovascular em idades mais avançadas; ii) os participantes apresentaram uma frequência elevada de hipertensão arterial, tabagismo, obesidade e sedentarismo; iii) diabetes, dislipidemia e arritmia estavam presentes com mais frequência nos escalões etários acima dos 60 anos; e iv) os participantes reportaram, com uma elevada frequência, a existência de familiares diretos com algum tipo de demência.

Particularmente relevante é o facto de o tabagismo, a obesidade e o sedentarismo estarem presentes desde idades relativamente jovens, dado que estes fatores exercem efeito cumulativo e são passíveis de ser modificados.

Um subgrupo reduzido de participantes (N=82) apresentou desempenho cognitivo abaixo do esperado para a sua idade e escolaridade. Na sua maioria estavam bem inseridos na sociedade e tinham profissão ativa.

As principais dificuldades enfrentadas para a implementação do estudo relacionaram-se com a inclusão de participantes. Foi necessário e essencial a implementação de estratégias de divulgação do estudo experimental junto dos munícipes e instituições do município de Águeda, para o dinamizarem, o que possibilitou o cumprimento dos objetivos estipulados. Em particular, foi muito relevante o envolvimento de instituições como embaixadoras da saúde cognitiva e a liderança da Câmara Municipal de Águeda.

Os autores mantêm a plataforma web *Brain on Track* ativa pelo que todos os cidadãos que se inscreveram no estudo experimental continuam a poder monitorizar a sua saúde cognitiva. A manutenção de um registo individual de desempenho cognitivo ao longo do tempo poderá contribuir para a motivação dos indivíduos a aderirem a comportamentos e hábitos de vida que protejam a saúde do cérebro e o desempenho cognitivo a longo prazo.

Agradecimentos

Os autores desejam manifestar o seu agradecimento à Câmara Municipal de Águeda, em particular ao Drº Gil Nadais, presidente da Câmara Municipal de Águeda na altura em que decorreu o estudo experimental reportado neste artigo, à permanente disponibilidade do Engº Carlos Alves da HFA, e às instituições que colaboraram como embaixadoras da saúde cognitiva, pela sua proatividade. Por último, mas não menos importante, os autores agradecem aos 1020 cidadãos que se voluntariaram para participar no estudo experimental. A monitorização cognitiva da população do município de Águeda foi financiado pela Câmara Municipal de Águeda e a presente publicação foi parcialmente suportada pelo Sistema de Incentivos à Investigação e Desenvolvimento Tecnológico (SI I&DT) do Programa Portugal 2020, através do Programa Operacional Competitividade e Internacionalização e Programa Operacional do Centro do FEDER - Fundo Europeu de Desenvolvimento Regional, no âmbito do projeto 017861 - Social Cooperation for Integrated Assisted Living (SOCIAL).

Referências

Behrman, A. L., Bowden, M. G., & Rose, D. K. (2013). Clinical trials in neurorehabilitation. In *Handbook of clinical neurology*, Vol. 110, pp. 61–66. Elsevier.

Bucci, P., Piegari, G., Mucci, A., Merlotti, E., Chieffi, M., De Riso, F., ... & Galderisi, S. (2013). Neurocognitive individualized training versus social skills individualized training: a randomized trial in patients with schizophrenia. *Schizophrenia research*, 150(1), 69–75.

Cicerone, K. D., Dahlberg, C., Malec, J. F., Langenbahn, D. M., Felicetti, T., Kneipp, S., ... & Laatsch, L. (2005). Evidence-based cognitive rehabilitation: updated review of the literature from 1998 through 2002. *Archives of physical medicine and rehabilitation*, 86(8), 1681–1692.

Cruz, V. T., Pais, J., Bento, V., Mateus, C., Colunas, M., Alves, I., ... & Rocha, N. P. (2013). A rehabilitation tool designed for intensive web-based cognitive training: Description and usability study. *JMIR research protocols*, 2(2).

Cruz, V. T., Pais, J., Alves, I., Ruano, L., Mateus, C., Barreto, R., ... & Coutinho, P. (2014). Web-based cognitive training: patient adherence and intensity of treatment in an outpatient memory clinic. *Journal of medical Internet research*, 16(5).

Eaton, J., McCay, L., Semrau, M., Chatterjee, S., Baingana, F., Araya, R., ... & Saxena, S. (2011). Scale up of services for mental health in low-income and middle-income countries. *The Lancet*, 378(9802), 1592–1603.

Freitas, S., Simões, M. R., Martins, C., Vilar, M., & Santana, I. (2010). Estudos de adaptação do Montreal Cognitive Assessment (MoCA) para a população portuguesa. *Avaliação Psicológica*, 9(3).

Gitlin, L. N., Kales, H. C., & Lyketsos, C. G. (2012). Nonpharmacologic management of behavioral symptoms in dementia. *Jama*, 308(19), 2020–2029.

Graff, M. J., Vernooij-Dassen, M. J., Thijssen, M., Dekker, J., Hoefnagels, W. H., & Rikkert, M. G. O. (2006). Community based occupational therapy for patients with dementia and their care givers: randomised controlled trial. *Bmj*, 333(7580), 1196.

Kakuma, R., Minas, H., van Ginneken, N., Dal Poz, M. R., Desiraju, K., Morris, J. E., ... & Scheffler, R. M. (2011). Human resources for mental health care: current situation and strategies for action. *The Lancet*, 378(9803), 1654–1663.

Lewis, J. R. (2002). Psychometric evaluation of the PSSUQ using data from five years of usability studies. *International Journal of Human-Computer Interaction*, 14(3-4), 463-488.

Lischka, A. (2012). A systematic review of screening tools for predicting the development of dementia. *CCan J Aging*, *31*(3), 295–311.

Lonie, J. A., Tierney, K. M., & Ebmeier, K. P. (2009). Screening for mild cognitive impairment: a systematic review. *International Journal of Geriatric Psychiatry: A journal of the psychiatry of late life and allied sciences*, 24(9), 902–915.

Milne, A., Culverwell, A., Guss, R., Tuppen, J., & Whelton, R. (2008). Screening for dementia in primary care: a review of the use, efficacy and quality of measures. *International Psychogeriatrics*, 20(5), 911–926.

Neuroinova (2015). Brain on Track: Sistema de Monitorização Cognitiva Contínua. Retrieved February 9, 2015, from http://www.brainontrack.pt/

Petersen, R., Stevens, J., & Ganguli, M. (2001). Practice parameter: Early detection of dementia: Mild cognitive impairment (an evidence-based review) Report of the Quality Standards Subcommittee of the American. *Neurology*, *56*(9), 1133–1142.

Rebok, G. W., Langbaum, J. B., Jones, R. N., Gross, A. L., Parisi, J. M., Spira, A. P., ... & Brandt, J. (2013). Memory training in the ACTIVE study: How much is needed and who benefits?. *Journal of aging and health*, 25(8_suppl), 21S-42S.

Rosa, A. F., Martins, A. I., Costa, V., Queirós, A., Silva, A., & Rocha, N. P. (2015, June). European Portuguese validation of the post-study system usability questionnaire (PSSUQ). In *Information Systems and Technologies (CISTI), 2015 10th Iberian Conference on*, 1–5.

Ruano, L., Sousa, A., Severo, M., Alves, I., Colunas, M., Barreto, R., ... & Lunet, N. (2016). Development of a self-administered web-based test for longitudinal cognitive assessment. *Scientific reports*, 6, 19114.

Schmand, B., Jonker, C., Hooijer, C., & Lindeboom, J. (1996). Subjective memory complaints may announce dementia. *Neurology*, 46(1), 121–125.

Vincent, C., Desrosiers, J., Landreville, P., & Demers, L. (2009). Burden of caregivers of people with stroke: evolution and predictors. *Cerebrovascular Diseases*, 27(5), 456–464.

Wimo, A., Jönsson, L., Bond, J., Prince, M., Winblad, B., & International, A. D. (2013). The worldwide economic impact of dementia 2010. *Alzheimer's & Dementia*, 9(1), 1–11.

Woods, B., Aguirre, E., Spector, A. E., & Orrell, M. (2012). Cognitive stimulation to improve cognitive functioning in people with dementia. *Cochrane Database of Systematic Reviews*, (2).

World Health Organization. (2007). *Atlas: Global resources for persons with intellectual disabilities: 2007*. World Health Organization.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 24/10/2018
Aceitação/Acceptance: 11/12/2018

Herramienta virtual usando dispositivo háptico para terapias de concentración en niños con déficit atencional

Ivón Escobar[1], Paulina Guanoluisa[1], Javier Montaluisa[1], Edwin Pruna[1] y Cristian Duque.

ipescobar@espe.edu.ec, apguanoluisa@espe.edu.ec, fjmontaluisa@espe.edu.ec, eppruna@espe.edu.ec, cristianduque@hotmail.es

[1] Universidad de las Fuerzas Armadas ESPE, Sangolqui, Ecuador.

Pages: 513–525

Resumen: Se presenta una herramienta virtual para terapias de concentración en niños con déficit Atencional. Se crearon 13 escenas consecutivas en Unity 3D, con modelos usando cubos, basados en el "Diseño con Cubos" del test WISC-IV; como parte de la estimulación se adoptó un dispositivo háptico para aumentar la capacidad atencional de los niños. La herramienta fue utilizada por 5 niños con edades comprendidas entre 9 y 12 años y bajo la supervisión de un especialista. Durante las 4 sesiones realizadas se pudo observar que la herramienta virtual aumenta la inmersión y sobre todo la capacidad de concentración de los niños, además incrementa el compromiso y motivación del paciente hacia la terapia y podría ser utilizado en terapias para niños que sufren de TDAH.

Palabras-clave: dispositivo háptico, Unity 3D, WISC-IV, terapias, TDAH.

Virtual tool with the use of a haptic device for concentration therapy in children with Attention Deficit

Abstract: A virtual tool for concentration therapies in children with attention deficit is presented. 13 consecutive scenes were created in Unity 3D, with cube models which are based on the "Design with Cubes" of the WISC-IV test; As part of the stimulation, a haptic device was adopted to increase the attention capacity of the children. The tool was used by 5 children aged between 9 and 12 years and under the supervision of a specialist. During the four carried out sessions, it was observed that the virtual tool increases the immersion and rises the concentration capacity of the children; also, it increases the commitment and motivation of the patient towards the therapy, so it could be used in therapies for children suffering from ADHD.

Keywords*:* haptic device, Unity 3D, WISC-IV, therapies, ADHD.

1. Introducción

En la actualidad es mayor la cantidad de casos de niños y jóvenes que requieren controlar sus problemas de atención. La indisciplina y el bajo rendimiento son las principales causas que aquejan a los padres de familia y a los centros educativos por la falta de atención (American Psychiatric Association, 2013). La falta de atención puede ser debida a diversas razones, por ansiedad, por depresión, debido a cambios repentinos en su vida (como el divorcio de sus padres), problemas de aprendizaje o incluso debido a que el niño sufre de trastorno de hiperactividad y déficit de atención (TDAH)(Faraone et al., 2015).

El TDAH ha sido definido como un trastorno de inicio en la infancia, caracterizado por la presencia de un patrón persistente de desatención o hiperactividad e impulsividad (Visser et al., 2014). Los niños que presentan TDAH pueden llegar a presentar dos tipos de alteraciones: de atención sostenida y discriminación atencional, de modo que son susceptibles a perder el foco de atención por interferencia ambiental irrelevante (Estévez et al., 1997) (Sánchez & Narbona, 2001).

En el interés por buscar tratamientos eficaces para niños y adolescentes con TDAH han surgido algunos estudios sobre la aplicación de test clínicos para evaluar a niños y adolescentes que presenten síntomas de TDAH (Grant & Berg, 1948) (Wechsler, 1951). La Escala de inteligencia de Wechsler para niños-IV "WISC-IV" (Psicodiagnosis.es., 2018), es un instrumento clínico que sirve para medir la inteligencia global, aplicable a niños y adolescentes de edades comprendidas entre 6 años 0 meses y 16 años 11 meses. El test es aplicable preferentemente a grupos de sujetos de los que convenga determinar su nivel de habilidades cognitivas o su funcionamiento neuropsicológico como diagnóstico de altas capacidades; diagnóstico de retraso mental leve o moderado; entre otros, así como también aplica a poblaciones con alteraciones atencionales, como el TDAH. Los ámbitos que evalúa el test WISC-IV son: Comprensión verbal, Razonamiento Perceptivo, Memoria de Trabajo, Velocidad de Procesamiento y Coeficiente Intelectual Total.

Dentro del test se encuentra el "diseño con cubos" que mide habilidades de análisis, síntesis y organización viso-espacial, a tiempo controlado. El niño dispone de 9 cubos para realizar diversos diseños. Estos cubos tienen colores en sus 6 caras, unas son rojas, otras blancas y otras mitad rojas y mitad blancas. Al iniciar la prueba el especialista presenta un modelo real construido delante del niño para que lo pueda reproducir. Posteriormente, el niño tendrá que componer diseños que aparecen dibujados en unas láminas, siempre con un tiempo controlado. Esta prueba se observa en la figura 1.

Dentro de los tratamientos también están las tareas de atención sostenida o TAS que son herramientas muy utilizadas en la evaluación clínica del déficit de atención con hiperactividad (TDAH) (Harper, Aylward & Brager, 2002). Las TAS fueron originalmente desarrolladas para detectar lesiones cerebrales en niños y adultos (Rosvoldet al., 1956)., pero ha demostrado una alta sensibilidad para distinguir niños que manifiestan conductas de inatención, distracción, agitación e impulsividad (Satterfield et al., 1972). Las tareas de atención sostenida consisten en la presentación de una secuencia de estímulos, ante la que el sujeto debe emitir una respuesta cada vez que aparece un estímulo determinado. Pueden cuantificarse las frecuencias y tiempos de reacción de los aciertos (respuesta ante la presentación del estímulo), errores por omisión (no respuesta

ante la presentación del estímulo). Existen diferentes tipos de TAS; en función de la modalidad de presentación de los estímulos, por ejemplo, se distingue entre TAS visual o auditiva.

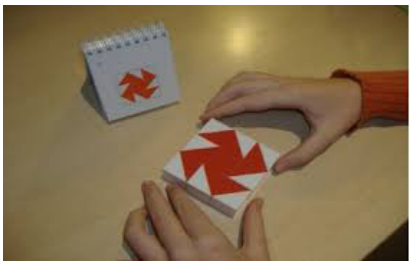


Figura 1 – Diseño con Cubos del Test WISC-IV

Para ayudar a superar los problemas de atención de los niños, los profesionales recomiendan realizar ejercicios de atención sostenida, que son actividades que busca que los niños logren focalizar su atención por periodos de tiempo específicos y que logren concluir con la tarea propuesta. Las actividades que pueden contribuir a captar la atención de los niños son: los rompecabezas, laberintos, sopas de letras, buscar diferencias, colorear. Es importante destacar que cada una de estas actividades deben ser motivadoras (Kouneiher, 2009) para conseguir la total atención de los niños.

Las nuevas tecnologías se han convertido en el eje central de multitud de áreas profesionales y actualmente ya es común su utilización en el campo de la psicología y la salud mental. (Parsey & Schmitter, 2013). La Realidad Virtual (RV), aplicada al tratamiento de casos clínicos, permite al paciente interactuar en entornos situados entre la exposición real y la exposición imaginada, para facilitar así cualitativamente el flujo terapéutico durante una sesión. La estimulación de los diferentes canales sensoriales gracias a la realimentación visual, táctil y auditiva provoca una psicología realista y respuestas de comportamiento (Parsons et al., 2017) (Bohil et al., 2011) (Iriarte et al., 2016).

Los entornos de realidad virtual se utilizan a menudo en terapias de meditación (Moseley, 2017) (Andersen et al., 2107) y en el tratamiento de trastornos mentales como la ansiedad y trastorno postraumáticos de estrés. En el estudio realizado en (Bohil et al., 2011) se demostró que estas terapias son igualmente eficaces en comparación con las terapias que emplean exposiciones a situaciones de la vida real.

La realidad virtual (VR) se ha mostrado como una herramienta efectiva en muchos dominios de terapia y rehabilitación sobre todo por la motivación que genera en los pacientes (Rizzo, 2002) (Weisset al., 2003) y con el desarrollo de tecnologías hápticas se ha vuelto posible tocar y sentir el entorno virtual y los objetos (Baud & Balzarotti, 2017), estos factores favorecen mucho a que los pacientes mejoren sus capacidades cognitivas (Wilson, 2002).

En este contexto el presente trabajo propone desarrollar una herramienta virtual basada en el “Diseño con Cubos” del test WISC-IV. Se presentan 13 pantallas consecutivas con diferentes diseños con cubos, los entornos están diseñados de tal manera que motiven y llamen la atención de los niños. Como parte de la estimulación también se adoptó un

dispositivo háptico, que provoca que las terapias efectuadas activen la parte sensorial de los pacientes, fortaleciendo a la vez la parte cognitiva. La herramienta se convierte en una tarea de atención sostenida ya que activa la parte visual táctil y auditiva., la edad de los pacientes que pueden usar la herramienta está comprendida entre 6 y 16 años tal como lo indica el test de WISC-IV. La herramienta diseñada también genera una base de datos que permite que el especialista lleve un historial de los ejercicios realizados por el paciente. El tiempo de ejecución de cada ejercicio y los errores cometidos por el paciente, se podrá observar en cada archivo generado. En la figura 2 se muestra el diagrama del sistema.

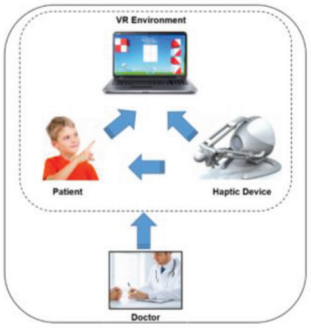


Figura 2 – Herramienta virtual utilizando dispositivo háptico

2. Metodología y Estructura de la Herramienta

El objetivo principal de esta herramienta es fortalecer la concentración de los niños que presentan atención dispersa o a su vez sufren de déficit de atención con hiperactividad (TDAH), estas dificultades en los niños suelen ser muy significativas debido a la relación que presentan con el rendimiento académico; en ambos casos los especialistas sugieren realizar tareas que logren aumentar el tiempo de concentración y tengan una progresión paulatina del tiempo de exposición y complejidad según la propia evolución del niño (Ríos et al., 2017)

La prueba de Cubos del test WISC-IV mide esencialmente la percepción visual, la capacidad de establecer relaciones espaciales y la coordinación visomotriz. Para su valoración el niño dispone de 9 cubos para realizar diversos modelos. Estos cubos tienen colores en sus 6 caras, unas son rojas, otras blancas y otras mitad rojas y mitad blancas.

El especialista al inicio de la prueba, presenta al niño un modelo que construye delante del niño y que él debe reproducir. Posteriormente, el niño tendrá que componer modelos que aparecen dibujados en unas láminas, siempre dentro de un tiempo límite, existiendo en los items más avanzados bonificación por acabar antes del tiempo señalado.

Considerando el modelo de la prueba de cubos WISC-IV y bajo el criterio de un psicólogo clínico se diseñó la herramienta virtual para ejercicios de concentración. La herramienta provee a los niños realimentación auditiva, visual y táctil, proporcionando una interacción de realidad virtual a través de una computadora y un dispositivo de

retroalimentación forzada como lo es el Novint Falcon. La herramienta propuesta es motivadora, fácil de usar y divertida. La Fig. 3 muestra el diagrama de bloques de la herramienta propuesta.

La interacción entre el dispositivo y el niño se establece mediante una comunicación bilateral donde el niño mediante la retroalimentación propioceptiva puede posicionar los objetos virtuales en el ambiente, y si realiza un movimiento incorrecto sentirá una fuerza de reacción, además una señal auditiva indica que se replicó el modelo correctamente esto provoca que el ambiente virtual sea más inmersivo e intuitivo para el niño.

En la interfaz gráfica se presenta un ambiente de visualización en 3D, donde interactúan los objetos virtuales y las fuerzas de posicionamiento que ejerce el niño, los objetos virtuales presentan colores y formas de acuerdo a los cubos del test WISC-IV, además está el modelo a ser replicado y la plantilla que servirá de base para colocar los cubos de acuerdo al modelo.

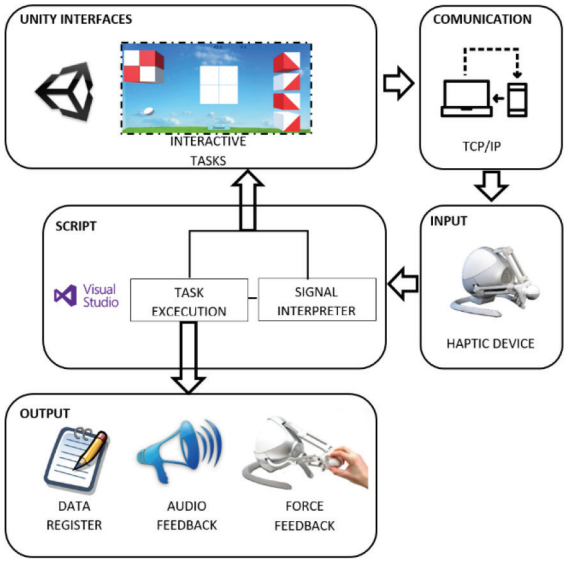


Figura 3 – Diagrama de bloques de la herramienta propuesta

3. Desarrollo de la Herramienta Virtual

El desarrollo de la herramienta virtual se centra en la adquisición de señal del periférico de entrada, el desarrollo de scripts, el diseño de entornos gráficos y finalmente la base de datos generada.

3.1. Adquisición de señal del periférico de entrada

El dispositivo háptico Novint Falcon, es el periférico de entrada en este sistema. Este dispositivo genera como salida un vector de 3 dimensiones en las posiciones cartesianas X,

Y y Z como se muestra la Figura 4a. Los estados de los botones son dados como un vector de valores lógicos, que indican cuando el botón es presionado (verdadero) o no (falso).

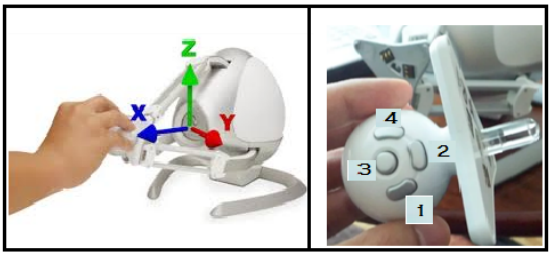


Figura 4 – a) Posiciones cartesianas del Novint Falcon. b) Disposición de botones

Los cuatros botones del controlador del dispositivo háptico han sido utilizados, y actúan mediante código de programación C Sharp (C#) para cumplir una determinada función. El botón 1 permite seleccionar el cubo que va a ser llevado a la plantilla donde se replicará el modelo indicado, los botones 2 y 3 hacen rotar el cubo elegido en el eje de las x, permitiendo seleccionar la cara del cubo, y el botón 4 cumple con la función de posicionar al cubo en la plantilla, en la figura 4b se muestra la posición de los botones.

La interacción que el paciente realiza con este dispositivo contribuye a que exista mayor concentración durante la realización de la tarea. La colocación incorrecta de un cubo, o no aceptar al final su posicionamiento generan un error o que se tenga que repetir nuevamente la tarea.

3.2. Desarrollo de scripts

La información proveniente del dispositivo Novint Falcon es interpretada por las interfaces virtuales en Unity mediante una comunicación TCP/ IP. Después de haber establecido la comunicación se reciben las señales de los sensores de movimiento que se encuentran en el dispositivo háptico. Además las interfaces virtuales están facultadas para enviar señales de realimentación de fuerzas hacia el dispositivo háptico. Esta característica le permite al sistema contar con una componente más de inmersión, pues le permite al usuario interactuar con el entorno virtual y a la vez sentir a través del tacto los eventos que se representan durante la ejecución de las tareas, así como completarlas. La figura 5 muestra el Diagrama de flujo de la conexión TCP/IP.

Los datos se interpretan desde el dispositivo de entrada con las librerías programadas en C# para la comunicación entre el dispositivo háptico y el ordenador. La codificación del sistema es desarrollado en Visual Studio. Se utilizan scripts en los cuales se manipula librerías propias de Visual Studio para organizar y administrar la información recibida de las actividades presentadas en el proyecto. Se ingresan los datos del paciente para luego elegir la tarea a desarrollar, una vez ejecutada cada tarea se pasa a la siguiente, y finalmente al concluir con todas las tareas se registra la información de tiempo y aciertos de cada tarea cumplida. En la figura 6 se indica un diagrama de flujo de la funcionalidad del sistema.

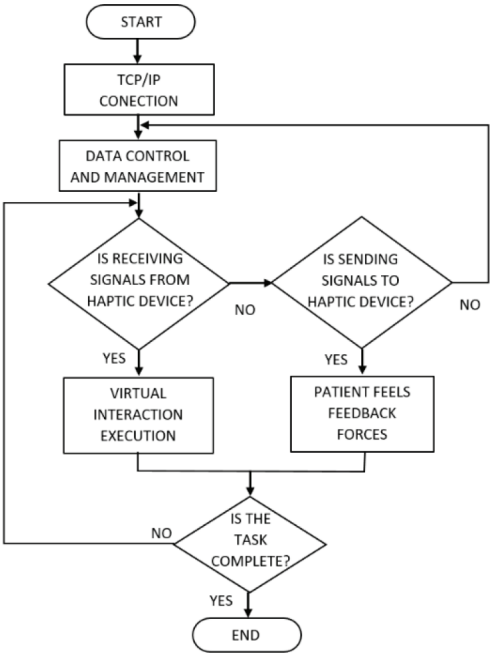


Figura 5 – Diagrama de flujo de la comunicación

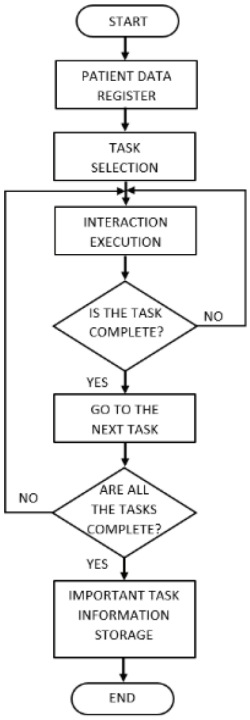


Figura 6 – Diagrama de flujo de la Funcionalidad del sistema

3.3. Diseño de entornos gráficos

Para diseñar y construir objetos virtuales, se utilizó el software Blender, en este software se bosquejan los cubos, los modelos y las plantillas para cada tarea de acuerdo a la hoja de Protocolo de Registro de Diseño de Cubos del test WISC-IV. Se asignaron a cada cara del cubo el color rojo, blanco y bicolor.

Para importar estos objetos virtuales se maneja un formato compatible al proyecto en Unity3D siendo así utilizables en la interfaz visual de la aplicación. La programación de los script se relaciona con la función que cada objeto debe ejecutar en la tarea.

La escena de cada tarea consta básicamente del modelo a ser replicado en colores blanco y rojo, la plantilla de color blanco y los cubos con sus distintas caras, y como fondo se presenta un entorno en color azul y verde sin mayores detalles para que no llame la atención del paciente.

La herramienta consta de 13 escenas, la escena del nivel de muestra y las escenas de los 12 niveles de diseños de cubos, además se incluyó en cada escena como datos de referencia para el especialista, el tiempo estimado para realizar el ejercicio y el tiempo que estará corriendo hasta que el paciente finalice la tarea. La figura 7 muestra las escenas de los 8 primeros niveles del Diseño de Cubos basadas en los modelos del test WISC-IV.

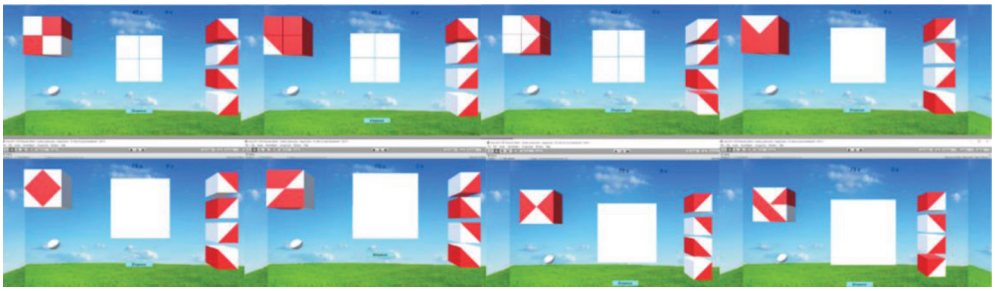


Figura 7 – Escenas de modelos basados en el test WISC-IV

3.4. Generación de la base de datos

Se ha generado una base de datos en Unity. Para esto se utilizó las librerías: using UnityEngine.UI, using System.Text, using System, using System.IO. La base de datos creada en un block de notas muestra: 1) los datos del paciente como nombre, apellido, edad, fecha de nacimiento, genero, fecha de aplicación de la sesión, y número de la sesión realizada. Estos datos son de interés para el especialista quién realizará la valoración del paciente de acuerdo al test WISC-IV. 2) El número del nivel de la tarea realizada con su tiempo de ejecución y si fue completada la tarea (acierto=1) o no (error=0). La figura 8 muestra el ingreso de datos y la base de datos generada.

a) b)

Figura 8 – a) Ingreso de datos y b) base de datos generada

4. Descripción del Uso de la Herramienta

Para dar inicio con el uso de la herramienta diseñada el paciente tiene que ingresar los datos mostrados en el menú de inicio, seguidamente debe pulsar el botón ingreso, una vez ejecutado aparecerá el mensaje “DATOS INGRESADOS”, y finalmente pulsando el botón DISEÑO DE CUBOS” aparecerá la primera tarea con el modelo de prueba.

Siguiendo el protocolo utilizado en el Test tradicional WISC-IV, el especialista indica el funcionamiento de la herramienta con la escena de prueba que se observa en la

figura 9. Se explica a los pacientes como funciona cada nivel, donde colocar los cubos y que botones debe presionar para: elegir un cubo, girar para cambiar de caras al cubo y para posicionar al cubo.

Figura 9 – Escena de prueba

Una vez indicado el nivel de prueba, el paciente continua con los siguientes niveles, para lo cual debe pulsar con el mouse el icono empezar que aparece en cada escena.

A medida que se avanza en la ejecución de las tareas, los niveles en cuanto a replica de diseños van subiendo de complejidad. Del nivel 1 al nivel 8 la complejidad es baja y las escenas presentan 4 cubos para poder replicar el modelo. Para replicar los niveles 9 y 10 se necesitan 9 cubos por lo que se consideran niveles de complejidad media, y finalmente los niveles 11 y 12 que a pesar de formarse con nueve cubos se los considera de alta complejidad porque su modelo queda en forma de rombo, generando que se ocupe uno de los botones para hacer que el cubo rote en el eje z. En la figura 10 se muestra los niveles con más complejidad.

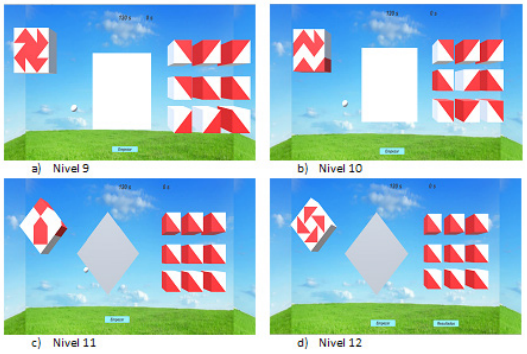


Figura 10 – Escenas de los niveles de mayor complejidad

5. Pruebas y Resultados

La herramienta como etapa inicial fue validada por un psicólogo clínico, quién en base a su experiencia realizo recomendaciones en cuanto a: modificación de colores de fondo, tamaños y realimentación en los aciertos y errores al realizar las tareas.

Una vez hechas las modificaciones en cuanto a las recomendaciones del especialista, se realizaron pruebas de usabilidad de la herramienta, 5 niños con edades comprendidas entre 9 y 12 años.

La validación de la herramienta se realizó en base a la observación (Escobar et al., 2017) y bajo la supervisión del especialista, para determinar si cumple con el objetivo de la terapia. La herramienta fue usada durante 2 semanas realizando 2 sesiones por semana.

En la tabla I se muestra un resumen de las observaciones realizadas en las sesiones, los resultados obtenidos en base a la observación, indican que todos los niños mostraron interés al conocer la herramienta, y el uso del dispositivo háptico provocó más curiosidad. Utilizando el nivel de prueba los niños entendieron como usar la herramienta y se adaptaron rápidamente. En la primera sesión y en los dos primeros niveles los niños tienen dificultad en direccionar el dispositivo háptico y en la elección de botones; una vez superado el inconveniente, se hace fácil para los niños el manejo del dispositivo en los niveles consecutivos.

Los niños se focalizan en terminar el diseño planteado, muestran mayor atención al momento de colocar cada cubo de acuerdo al color indicado, puesto que si al final no aceptan su elección, la tarea no se completa y en la base de datos se cuenta como un desacierto.

Preguntas	**Observaciones**
1. *Muestran los pacientes interés en el uso de la herramienta.*	Todos los niños muestran interés en el uso de la herramienta.
2. *El paciente entiende las indicaciones del uso de la herramienta.*	Todos los niños entienden cómo funciona la herramienta y se adaptan fácilmente.
3. *Es complicado el uso del dispositivo háptico para el paciente.*	Una vez realizado el nivel de prueba los niños usan con facilidad el dispositivo háptico.
4. *Cuál es el nivel de atención prestado por el paciente.*	Todos los niños se focalizan en terminar el diseño planteado, muestran mayor capacidad de atención
5. *Que efectos sobre la salud causa el uso de la herramienta en los pacientes.*	Un niño presento un leve y pasajero dolor de cabeza a medida que iba ejecutando las tareas en la primera sesión.
6. *En qué tiempo se ejecutaron las tareas.*	A partir de la segunda sesión las tareas se ejecutaron dentro de los tiempos establecidos por el Test tradicional WISC-IV.
7. *Existen factores de distracción en la escena presentada a los pacientes.*	No existen factores de distracción en las escenas presentadas.
8. *La realimentación visual, que efectos causa en el niño.*	Los niños se muestran satisfechos al ver el mensaje de "Tarea Completada" en cada nivel.

Tabla 1 – Resumen de observaciones

Mientras realizan las tareas los niños activan su capacidad viso-espacial, el tener que interactuar con los 4 botones del dispositivo háptico permite que aumente la capacidad atencional del niño.

En la primera sesión se pudo observar que un niño tuvo un leve y pasajero dolor de cabeza mientras realizaba las tareas; se procedió a suspender por un momento la actividad y luego se continuó sin inconvenientes.

En las primeras sesiones el tiempo utilizado para ejecutar cada nivel fue mayor al tiempo previsto en las experimentaciones con el test tradicional, esto se debe a que los niños estaban utilizando el dispositivo háptico y de inicio necesitan adaptarse y concentrarse en cómo

utilizarlo, sin embargo en la segunda sesión los niños ocuparon menor tiempo y a medida que iban subiendo los niveles iba disminuyendo el tiempo de ejecución de cada tarea.

Los colores de fondo no afectan en el foco atencional de los niños y a pesar que en cada escena se muestra el tiempo límite y existe un tiempo que está corriendo mientras se ejecutan las tareas, los niños no muestran distracción por el tiempo, puesto que al iniciar las tareas se les indica a que corresponde cada tiempo. Se ha determinado que las distracciones son de carácter externo, del ambiente donde se encuentran realizando la terapia.

Al terminar el diseño de cada nivel aparece un mensaje indicando que la tarea fue completada lo que provoca en los niños satisfacción. La figura 11 muestra el uso de la herramienta por los niños.

Figura 11 – Pruebas realizadas

4. Conclusiones y Trabajos Futuros

Se desarrolló una herramienta virtual para terapias de concentración utilizando un dispositivo háptico como periférico de entrada. El entorno virtual está diseñado en Unity 3D y está conformado por 13 tareas interactivas basadas en el Diseño de Cubos del Test WISC-IV y que cumplen con las características de una tarea de atención sostenida.

El uso del dispositivo háptico se ha vuelto un elemento importante de la herramienta virtual ya que aumenta la inmersión y sobre todo la capacidad de concentración de los niños, además permite incrementar el compromiso y motivación del paciente hacia la terapia. Los resultados obtenidos en la prueba de usabilidad indican que la herramienta tiene una aceptación para ser utilizado en terapias de concentración para niños que sufren de TDAH. Los pacientes no presentan molestias comprometedoras para su salud y tienen poca dificultad para usarlo.

Como trabajo futuro se pretende diseñar herramientas tecnológicas con el uso de gafas y ambientes virtuales inmersivos de tal forma que aumentan la motivación, el interés y sobre todo la capacidad atencional de los pacientes durante una terapia de concentración, todo esto basado en las TAS.

Referencias

American Psychiatric Association. (2013). DSM-5 attention deficit/hyperactivity disorder fact sheet. Arlington: American Psychiatric Association.

Andersen, T., Anisimovaite, G., Christiansen, A., Hussein, M., Lund, C., Nielsen, T., ... & Serafin, S. (2017, March). A preliminary study of users' experiences of meditation in virtual reality. In Virtual Reality (VR), 2017 IEEE (pp. 343–344). IEEE.

Aylward, G. P., Brager, P., & Harper, D. C. (2002). Relations between visual and auditory continuous performance tests in a clinical population: a descriptive study. Developmental Neuropsychology, 21(3), 285–303.

Baud-Bovy, G., & Balzarotti, N. (2017, November). Using force-feedback devices in educational settings: a short review. In Proceedings of the 1st ACM SIGCHI International Workshop on Multimodal Interaction for Education (pp. 14–21). ACM.

Bohil, C. J., Alicea, B., & Biocca, F. A. (2011). Virtual reality in neuroscience research and therapy. Nature reviews neuroscience, 12(12), 752.

Escobar, I., Guamán, A., Montaluisa, J., Pruna, E., & Marín, Y. (2017, October). Mobile application for vowel learning in children with Down Syndrome "LVDS-App". In Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON), 2017 CHILEAN Conference on (pp. 1–5). IEEE.

Estévez-González, A., García-Sánchez, C., & Junqué, C. (1997). La atención: una compleja función cerebral. Revista de neurología, 25(148), 1989–1997.

Faraone SV, Asherson P, Banaschewski T, Biederman J, Buitelaar JK, Ramos Quiroga JA, Rohde LA, Sonuga-Barke EJS, Tannock R, Franke B. Attentiondeficit/hyperactivity disorder. Nat Rev. 2015;1:1–23.

Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. Journal of experimental psychology, 38(4), 404.

Iriarte, Y., Diaz-Orueta, U., Cueto, E., Irazustabarrena, P., Banterla, F., & Climent, G. (2016). AULA—Advanced virtual reality tool for the assessment of attention: Normative study in Spain. Journal of Attention Disorders, 20(6), 542–568.

Kouneiher, F., Charron, S., & Koechlin, E. (2009). Motivation and cognitive control in the human prefrontal cortex. Nature neuroscience, 12(7), 939.

Moseley, R. (2017). Deep immersion with Kasina: an exploration of meditation and concentration within virtual reality environments.

Parsey, C. M., & Schmitter-Edgecombe, M. (2013). Applications of technology in neuropsychological assessment. The Clinical Neuropsychologist, 27(8), 1328–1361.

Parsons, T. D., Carlew, A. R., Magtoto, J., & Stonecipher, K. (2017). The potential of function-led virtual environments for ecologically valid measures of executive function in experimental and clinical neuropsychology. Neuropsychological rehabilitation, 27(5), 777-807.

Psicodiagnosis.es. Especialistas en Psicología Infantil y Juvenil, 2018. http://www.psicodiagnosis.es/areaespecializada/instrumentosdeevaluacion/interpretaciondelwisciv/index.php#04f9119c9410d700.

Ríos-Gallardo, Á. M., Hernández, A. G., Nieto, D. C. T., Losada, M. A. B., Bonilla-Santos, J., & Cuéllar, L. V. (2017). Efectividad de una intervención cognitiva en funciones ejecutivas para mejorar flexibilidad mental en niños con trastorno por déficit de atención e hiperactividad. RFS, 8(1), 16–22.

Rizzo, A. S. (2002). Virtual reality and disability: emergence and challenge. Disability and rehabilitation, 24(11-12), 567-569.

Rosvold, H. E., Mirsky, A. F., Sarason, I., Bransome Jr, E. D., & Beck, L. H. (1956). A continuous performance test of brain damage. Journal of consulting psychology, 20(5), 343.

Sánchez-Carpintero, R., & Narbona, J. (2001). Revisión conceptual del sistema ejecutivo y su estudio en el niño con trastorno por déficit de atención e hiperactividad. Revista de Neurología, 33(1), 47–53.

Satterfield, J. H., Cantwell, D. P., Lesser, L. I., & PODOSIN, R. L. (1972). Physiological studies of the hyperkinetic child: I. American Journal of Psychiatry, 128(11), 1418-1424.

Visser, S. N., Danielson, M. L., Bitsko, R. H., Holbrook, J. R., Kogan, M. D., Ghandour, R. M., ... & Blumberg, S. J. (2014). Trends in the parent-report of health care provider-diagnosed and medicated attention-deficit/hyperactivity disorder: United States, 2003–2011. Journal of the American Academy of Child & Adolescent Psychiatry, 53(1), 34-46.

Wechsler, D. (1951). Equivalent test and mental ages for the WISC. Journal of consulting psychology, 15(5), 381.

Weiss, P. L., Bialik, P., & Kizony, R. (2003). Virtual reality provides leisure time opportunities for young adults with physical and intellectual disabilities. CyberPsychology & Behavior, 6(3), 335-342.

Wilson, M. (2002). Six views of embodied cognition. Psychonomic bulletin & review, 9(4), 625-636.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 10/10/2018
Aceitação/Acceptance: 01/12/2018

Incorporando pruebas basadas en modelos para servicios web en un proceso de desarrollo ágil: Un caso de estudio en la industria

Brenda Aymerich[1], Andrés Solano[1], Christian Quesada-López[1], Alexandra Martínez[1]

{brenda.aymerich, andres.solanoportuguez, alexandra.martinez, cristian.quesadalopez}@ucr.ac.cr

[1] Universidad de Costa Rica, San José, 2060, Costa Rica.

Pages: 526–537

Resumen: Los equipos ágiles enfrentan dificultades para poder realizar pruebas de software a profundidad, dadas las iteraciones cortas de desarrollo. En muchos casos, las pruebas para servicios web se realizan manualmente, consumen mucho tiempo y requieren la experiencia de los miembros del equipo. Un enfoque de pruebas basadas en modelos, que permita la automatización de estas pruebas, podría mejorar la eficiencia del proceso y la calidad de los productos; sin embargo, su adopción no debería contravenir los valores, principios y prácticas de las metodologías ágiles. En este caso de estudio discutimos el proceso realizado para incorporar pruebas basadas en modelos para automatizar las pruebas de servicios web en un equipo que implementa prácticas ágiles, y analizamos su efectividad al usar la herramienta *TestOptimal* sobre servicios web *RESTful*. Asimismo, discutimos las percepciones de los miembros del equipo, los retos y oportunidades de uso de este tipo de enfoques en equipos ágiles. Los resultados indican que las pruebas basadas en modelos permiten aumentar la cantidad de casos de prueba y defectos encontrados. Por su parte, los miembros del equipo consideran que para aumentar la aceptación de estos enfoques durante el desarrollo de un proyecto ágil, son esenciales el conocimiento para el modelado y las herramientas de apoyo. A pesar de que se logra una mejora en la generación de casos de pruebas automatizados y en la detección de errores, las pruebas basadas en modelos se perciben como un enfoque complejo de aplicar.

Palabras-clave: Pruebas basadas en modelos; servicios web; *RESTful*; metodologías ágiles.

Incorporating model-based testing for web services in an agile development process: A case study in the industry

Abstract: Agile teams face difficulties to do in-depth software testing, given the short development iterations. In many cases, tests for web services are performed manually, consume a lot of time, and require the expertise of team members. A model-based testing approach that allows the automation of these tests could improve the efficiency of the process and the quality of the products; however, its adoption

should not violate the values, principles or practices of agile methodologies. In this case study we discuss the process followed to incorporate model-based testing for automating web services tests in a team that implements agile practices, and analyze its effectiveness when using the TestOptimal tool on RESTful web services. We also discuss the perception of team members, the challenges and opportunities of using this type of approaches in agile teams. The results indicate that model-based tests increase the number of test cases and defects found. Meanwhile, team members believe that to increase the acceptance of these approaches during the development of an agile project, modeling knowledge and support tools are essential. Despite achieving an improvement in the generation of automated test cases and error detection, model-based tests are perceived as a complex approach to apply.

***Keywords**:* Model-based Testing; MBT; Model-Driven Testing; web services; RESTful; agile methodologies.

1. Introducción

Dada la creciente complejidad en el desarrollo de sistemas de software, la utilización de arquitecturas orientadas a servicios ha emergido como una estrategia muy atractiva debido a su capacidad para responder con flexibilidad, adaptabilidad y rapidez a las necesidades cambiantes en las organizaciones (Krogdahl, Luef & Steindl, 2005). El desarrollo de servicios web basados en REST (*Representational State Transfer),* estándar de la arquitectura de comunicación basada en web, es deseable dada su capacidad de ser consumido por diferentes clientes tales como aplicaciones móviles Android o iOS y aplicaciones web por medio de su protocolo de transferencia de hipertexto (HTTP) para la comunicación de datos (Arcuri, 2017). Sin embargo, el proceso de verificación funcional de este tipo de servicios no es trivial, y aún es necesario desarrollar más conocimiento y estudios que ayuden a que el proceso de pruebas sea más eficiente (Fertig & Braun, 2015). En particular, las pruebas manuales para servicios web *RESTful* presentan un gran desafío debido al tiempo que consume su elaboración y ejecución, por la dificultad que presenta el contexto de los servicios web. Algunos de los principales retos que enfrentan estas pruebas son: la falta de una interfaz de usuario, que hace que las entradas y salidas involucradas no sean fácilmente manipulables; el uso de componentes débilmente acoplados, lo cual puede presentar incompatibilidad de datos; y la confiabilidad del medio de comunicación, que puede derivar fallas difíciles de distinguir (Chakrabarti & Kumar, 2009).

Por otro lado, en los entornos de trabajo ágiles se acepta el cambio como una ventaja competitiva para el negocio, y para ello se trabaja en iteraciones cortas que producen entregas incrementales del producto final. Además los desarrolladores realizan diversas tareas del ciclo de desarrollo y ninguno está específicamente dedicado a las pruebas de software (Schwaber & Sutherland, 2013), por lo que resulta importante reducir el esfuerzo humano en tareas repetitivas, acelerando así los ciclos de ejecución de pruebas, con el fin de adaptarse al cambio con una retroalimentación temprana (Collins & de Lucena, 2012).

Por esa razón, a pesar de que hoy en día aún se implementan pruebas manuales en entornos ágiles (Collins & de Lucena, 2012), es de vital importancia encontrar opciones para su automatización. El enfoque de pruebas basadas en modelos (MBT, por sus siglas en inglés: *model-based testing*) es una estrategia de automatización de pruebas de software

que deriva casos de prueba a partir de los modelos que describen el comportamiento del sistema, los cuales son diseñados por un profesional especializado en pruebas (De Cleva Farto & Endo, 2015). MBT se usa principalmente para pruebas funcionales (Utting & Legeard, 2004), y existe evidencia de que puede ayudar a automatizar el proceso de pruebas para servicios web *RESTful*. También han surgido propuestas para utilizar MBT en entornos ágiles, donde se crean modelos sencillos y se reduce la dependencia de profesionales especializados en pruebas (Thopate & Kachewar, 2012).

El objetivo de este caso de estudio es analizar un enfoque de pruebas basadas en modelos para automatizar las pruebas funcionales de los servicios web implementados mediante una *API RESTful*[1] con el propósito de evaluar su eficacia, eficiencia y aceptación en el contexto de un equipo de desarrollo ágil.

El artículo está estructurado de la siguiente manera: La Sección 2 presenta los trabajos relacionados. La Sección 3 describe el caso de estudio. La Sección 4 discute los resultados y finalmente, en la Sección 6 se presentan las conclusiones y trabajo futuro.

2. Trabajo relacionado

Las pruebas basadas en modelos han sido utilizadas para realizar pruebas en diversos dominios asociados al sistema bajo prueba tales como: aplicaciones web, aplicaciones móviles, sistemas concurrentes, sistemas integrados, sistemas distribuidos y servicios web (Villalobos-Arias, Quesada-Lopez, Martinez & Jenkins, 2018). En el caso de las pruebas para los servicios web se han usado enfoques donde se utilizan diversas herramientas tales como *GraphWalker*, *Xtext* y *Swagger*. Asimismo, se ha puesto en práctica procesos que han presentado resultados favorables para este enfoque de pruebas (Fertig & Braun, 2015). En esta sección presentamos estudios relacionados que detallan enfoques y herramientas para probar servicios web con pruebas basadas en modelos. De igual modo, discutimos estudios que han utilizado este enfoque en entornos ágiles.

Diferentes estudios han analizado la aplicación de pruebas basadas en modelos para probar distintos aspectos de los servicios web, como lo son pruebas funcionales (Pinheiro, Endo & Simao, 2010; Wu & Huang, 2013; Cleva Farto & Endo, 2015), pruebas de seguridad (Fertig & Braun, 2015), pruebas de composición (Belli, Endo, Linschulte, & Simao, 2014) y cobertura de código (Arcuri, 2017). Los enfoques utilizados para las pruebas de servicios web utilizan modelos UML, máquinas y grafos de estado finito, máquinas de estado finito extendido e incluso modelos inferidos a partir de entradas y salidas de las interfaces, los cuales son generados con herramientas como *Xtext* y *Swagger* (Fertig & Braun, 2015; Arcuri, 2017). Por otra parte, los casos de pruebas son producidos para *frameworks* como JUnit y RestAssured (Arcuri, 2017). La mayoría de los estudios anteriores implementan su propia solución para la generación automatizada de los casos de pruebas.

Para la automatización de las pruebas en entornos ágiles se han utilizado diversos enfoques de pruebas basadas en modelos. Por ejemplo, Sivanandan & Yogeesha (2018) aplican pruebas basadas en modelos mediante un marco de trabajo formado por un generador

[1] Una API RESTful es el conjunto de servicios que contiene las funciones y procedimientos necesarios para compartir información entre dos sistemas de software utilizando la arquitectura REST.

de pruebas de interfaz de usuario basado en modelos construidos a partir de máquinas de estados finitos (*GraphWalker*) y un marco de trabajo para automatización de pruebas de aceptación llamado *Robot*; el cual facilita la ejecución de pruebas independientemente de la experiencia del desarrollador. Ellos concluyen que el desarrollo de software ágil y las pruebas basadas en modelos obtienen beneficios mutuos, mejorando la flexibilidad, cobertura y mantenibilidad de las pruebas (Sivanandan & Yogeesha, 2014). Por su lado, Reider, Magnus & Krause (2018), proponen la escogencia de ciertos requerimientos que representen el comportamiento del sistema para la elaboración de los modelos y una posterior priorización en la ejecución de las pruebas, con el fin de asegurar la retroalimentación temprana y eficiente de sus resultados en un entorno ágil. Thopate & Kachewar (2012) diseñaron un modelo llamado *Chameleon* que se adapta a los cambios de los casos de prueba para responder de mejor manera al dinamismo de los desarrollos ágiles y además permite la adición de casos de prueba manuales para evitar dependencias de conocimientos especializados en pruebas automatizadas.

En términos generales, para la aplicación de pruebas basadas en modelos en proyectos ágiles los enfoques buscan adaptarse al cambio respondiendo con modelos que sean sencillos de actualizar y que a su vez no dependan de personas especializadas para realizar las pruebas. Sin embargo, tomando en consideración lo expuesto en este sección, existe escasez de información o reporte de casos con respecto a la experiencia de no solo la aplicación e incorporación de pruebas basadas en modelos para entornos ágiles, sino también para el aseguramiento de la calidad de servicios web en productos que se trabajan siguiendo estas prácticas.

Lo anterior es posible que suceda por la resistencia en general que se tiene, según Thopate & Kachewar (2012), de invertir en automatización de pruebas debido a lo riesgoso que se considera el constante cambio que pueden sufrir los casos de prueba en ambientes ágiles.

Por lo tanto, el caso de estudio que se describe en el presente artículo expone la experiencia y el análisis de la incorporación de un enfoque de pruebas basadas en modelos para la automatización de pruebas de un servicio web *RESTful* a un equipo que implementa prácticas ágiles, utilizando a su vez, una herramienta que en similitud con otros estudios, permite pruebas funcionales y modelado con diagramas de estado.

3. Descripción del estudio empírico

La estructura del diseño del estudio se realiza de acuerdo a los lineamientos para estudios experimentales propuesta por Wohlin, Runeson, Höst, Ohlsson, Regnell & Wesslén (2012). En este caso de estudio analizamos el proceso que introduce pruebas basadas en modelos para servicios web en un equipo que implementa prácticas ágiles. En este estudio se analizan las siguientes preguntas de investigación:

RQ1: ¿Cuál es la eficacia y eficiencia de aplicar pruebas basadas en modelos para automatizar las pruebas funcionales de servicios web?

RQ2: ¿Cuál es la aceptación, por parte de los profesionales, al introducir pruebas basadas en modelos en un proyecto de desarrollo ágil?

Contexto. El estudio se realiza en una organización de desarrollo de software pequeña, la cual es un emprendimiento que desarrolla aplicaciones móviles en distintas tecnologías. Los participantes del estudio son 4 desarrolladores de un equipo de desarrollo los cuales se reparten las diferentes tareas en cuanto al análisis, diseño, implementación y pruebas de los servicios web implementados utilizando el marco de trabajo Scrum. Los miembros del equipo que participan en la aplicación del proceso de pruebas basadas en modelos son graduados en ingeniería con edades entre 23 y 24 años y con aproximadamente dos años de experiencia profesional en tareas de desarrollo y con poca experiencia en procesos de pruebas de software.

Sistema bajo pruebas (SUT). El SUT utilizado en el estudio es un servicio *RESTful*, desarrollado con el *Framework Django* en *Python*. La *API* responde todas las solicitudes con *JSON*, las cuales tienen una estructura definida por el equipo de desarrollo. El servicio permite la creación de listas de compras de productos comerciales y el ingreso de estos productos utilizando su código de barras. Se selecciona una versión del servicio web liberada para las pruebas. Posteriormente, se selecciona el flujo de uno de sus componentes para realizar el modelo. El modelo representa el comportamiento del ingreso y consulta de productos utilizando el escaneo de un código de barras. Todos los desarrolladores conocen el SUT.

Herramienta de pruebas basadas en modelos. La herramienta seleccionada es *TestOptimal* la cual cumple todas las características requeridas por el equipo ágil. Los criterios para la selección de la herramienta fueron: (a) permitir realizar pruebas funcionales para el SUT, (2) definir los modelos de comportamiento mediante diagramas de estado, (3) disponibilidad en línea (4) posibilidad de utilizar la funcionalidad completa de la herramienta a lo largo del proyecto, y finalmente (5) nivel de documentación aceptable, incluyendo la descripción de todas las funciones y métodos necesarios para implementar la solución.

Una limitación a considerar es que la herramienta *TestOptimal* no es gratuita. Sin embargo, sus vendedores ofrecen una licencia académica del producto.

Métricas de evaluación. La eficacia y eficiencia será descrita en términos de los casos de prueba generados, los defectos encontrados y del tiempo requerido para modelar el flujo y generar los casos de pruebas. La aceptación del enfoque de pruebas basadas en modelos por parte del equipo se determina de acuerdo a la facilidad de uso de la herramienta, y la percepción en cuanto al valor agregado que brinda la técnica, contra el trabajo adicional requerido para llevarlo a cabo. Para esto se aplica un cuestionario siguiendo las recomendaciones del modelo de aceptación de tecnología (TAM) (Davis, 1989).

3.1. Proceso de aplicación y recolección de datos

Una vez definido el enfoque de pruebas basadas en modelos para ser utilizado por el equipo en la iteración del proyecto ágil, se aplicó el proceso empírico y se recolectó la información de las métricas a evaluar. El proceso se dividió en tres etapas:

Preparación de instrumentos y materiales. Para el proceso se diseñaron dos instrumentos de encuesta. La primera con el objetivo de conocer el grado de experiencia

de los desarrolladores en pruebas de software y pruebas basadas en modelos. La segunda con el fin de medir la percepción de los desarrolladores hacia el enfoque y la herramienta de pruebas basadas en modelos, que incluye su percepción en cuanto a su utilidad, facilidad e intención de uso futuro (Davis, 1989).

Asimismo, se confeccionó el material de capacitación para utilizarlo en las sesiones de entrenamiento del equipo de desarrollo las cuales se llevaron a cabo como *Spikes* en la iteración. Este material se compone de presentaciones y materiales que explican el enfoque de pruebas basadas en modelos y detalla el uso de la herramienta de pruebas seleccionada.

Sesión de capacitación. La sesión de capacitación para la inducción del enfoque de pruebas basadas en modelos con los miembros del equipo fue de una hora donde se explicó el enfoque y se describió los pasos a seguir para utilizar la herramienta *TestOptimal*. En esta sesión, primero se aplica la encuesta para conocer el grado de experiencia de los desarrolladores y segundo, se realiza la capacitación que incluye ejemplos del proceso y uso de la herramienta.

Ejecución del proceso. Al finalizar la capacitación, se procede con la ejecución del proceso para la aplicación de las pruebas basadas en modelos sobre las funcionalidades desarrolladas en el sistema. En este caso y siguiendo el espíritu ágil, el equipo se dividió en dos subgrupos, donde cada uno realizó su modelado, generación de casos de pruebas y ejecución de las pruebas. La división se dio de acuerdo con las habilidades y experiencia de los desarrolladores en pruebas de software para mantener el balance en los equipos. Una vez conformados los grupos, comenzaron el proceso de modelado con la herramienta. Se les especificó el componente a modelar, y se estableció un tiempo límite de tres horas para completarlo. Este tiempo incluyó el proceso de generación de casos de pruebas y la ejecución de las mismas. Al finalizar el proceso, cada desarrollador completó la segunda encuesta sobre su percepción acerca del proceso.

3.2. Evaluación y análisis del proceso de pruebas basadas en modelos

Para evaluar la eficiencia del enfoque y de la herramienta seleccionada, se comenzó por la recolección de los datos de las pruebas realizadas por el equipo para el SUT. Se extrajo la cantidad de casos de pruebas desarrollados, la cantidad de defectos encontrados y el tiempo dedicado a la generación de estos casos de prueba. Al contar con los datos de las pruebas manuales y al recolectar los datos resultantes de la aplicación del enfoque de pruebas basadas en modelos, se realizó una comparación de estos. Tomando en cuenta esta comparación y ambas encuestas, se realizó un análisis sobre la eficacia, la eficiencia y la percepción de los desarrolladores en cuanto a la aceptación del proceso y el valor agregado de utilizar pruebas basadas en modelos para el proyecto de desarrollo del equipo.

4. Resultados

En esta sección se reportan los resultados del caso de estudio. Durante la ejecución del caso de estudio dos equipos realizaron el proceso de modelado. Para el tiempo establecido solo un equipo finalizó el proceso de generación de casos de pruebas el cual es utilizado para la ejecución de las pruebas. En el caso del otro equipo, la herramienta

presentó problemas con la compilación de algunos de los *scripts* de pruebas generados. En este caso el segundo equipo aprovecha los *scripts* del primer equipo para continuar con la ejecución del proceso de pruebas sobre el SUT. A partir de la ejecución de los casos de pruebas se finaliza el proceso de recolección de los datos que incluye los casos de pruebas generadas, detección de fallos, y los tiempos requeridos para generar el modelo. En este caso los datos presentados contemplan la información del equipo que completa el proceso. Al otro equipo se le permite finalizar el proceso para que cuenten con el criterio necesario para completar la encuesta de percepción.

RQ1: Eficacia y eficiencia de aplicar pruebas basadas en modelos. La eficacia del enfoque y de la herramienta seleccionada se evalúa a partir de la cantidad de casos de prueba generados y la detección de fallos. Se comparan los datos recolectados con los datos históricos del proyecto para determinar el impacto de la aplicación de las pruebas basadas en modelos. La Tabla 1 presenta las métricas recolectadas a partir de la aplicación del proceso y la comparación con los datos de las pruebas manuales históricos del equipo. Presentamos la cantidad de casos de pruebas generados, los fallos encontrados con las pruebas, el tiempo de modelado y el tiempo de generación de los casos de prueba. Los resultados muestran que la cantidad de casos de prueba generados por la herramienta *TestOptimal* con el proceso de pruebas basados en modelos supera en 2.3 veces (13/30) el histórico de casos de prueba manuales del equipo. Del mismo modo, los errores encontrados por el proceso de pruebas basadas en modelos superan en 5.25 veces (4/21) las pruebas manuales del equipo. En este análisis se consideró como fallo cuando un servicio respondía con un código de estado 500 (Arcuri, 2017) y no se comprueba que la estructura del JSON de respuesta sea la esperada. Tomando en consideración el tiempo invertido para caracterizar la eficiencia obtenida, se tiene que ambos procesos tomaron aproximadamente la misma cantidad de tiempo para ejecutarse, ya fuera a través del modelado de pruebas basadas en modelos con *TestOptimal*, o diseñando los casos de pruebas manualmente.

Datos recolectados	**Históricos**	***TestOptimal***
Casos de prueba	13	30
Detección de fallos	4	21
Tiempo de modelaje (min)	0	190
Tiempo de generación (min)	195	0

Tabla 1 – Pruebas manuales frente a pruebas basadas en modelos.

RQ2: Aceptación al introducir pruebas basadas en modelos. La aceptación del enfoque y de la herramienta seleccionada se evalúa a partir de la aceptación por parte del equipo de desarrollo según su percepción en cuanto a la facilidad, utilidad e intención de uso futuro. Los datos recolectados con la encuesta sobre la percepción en cuanto a la facilidad, utilidad e intención de uso futuro de las pruebas basadas en modelos con la herramienta, responden las preguntas en una escala de *Likert* sobre las siguientes afirmaciones: “Fuertemente en desacuerdo”, “En desacuerdo”, “Neutral”, “De acuerdo” y “Fuertemente de acuerdo”.

Cada pregunta pertenecía a un constructo, es decir, cada una buscaba recolectar datos relacionados con la facilidad, utilidad o intención de uso. La encuesta constaba de 16 preguntas, de las cuales 7 pertenecían al constructo de facilidad de uso, 6 de utilidad y 3 de intención de uso.

Los resultados indican que en general, la herramienta se percibe como útil e impulsora de la calidad del producto, ya que la mayoría de respuestas de todos los participantes en estas preguntas se mantienen entre neutrales y de acuerdo (3 de acuerdo, 13/24 neutrales, 7/24 en desacuerdo y 1/24 fuertemente en desacuerdo). Con respecto a la facilidad de uso, la complejidad para el entendimiento, la flexibilidad y la aplicación del proceso mediante la herramienta provocaron que las respuestas tendieran a estar en desacuerdo (7/28 fuertemente en desacuerdo, 11/28 en desacuerdo, 7/28 neutrales, 2/28 de acuerdo y 1/28 fuertemente de acuerdo). Por último, las respuestas a las preguntas asociadas a intención de uso futuro tanto del enfoque de pruebas basadas en modelos como de la herramienta en el trabajo del equipo muestran neutralidad o desacuerdo (2/12 fuertemente en desacuerdo, 5/12 en desacuerdo, 2/12 neutrales y 3/12 de acuerdo).

4.1. Discusión

La experiencia de la aplicación del proceso de pruebas basadas en modelos y la herramienta *TestOptimal* muestra que es posible aumentar la cantidad de casos de prueba generados y fallos encontrados para un proyecto de desarrollo ágil que desarrolla servicios web *RESTful*. Estos resultados son comparados con los datos históricos de las pruebas manuales del equipo e indican que aproximadamente toman el mismo tiempo en ser realizadas. En primera instancia estos resultados indican que este enfoque puede ayudar a mejorar los procesos de pruebas en este contexto; sin embargo, es importante acotar que de la gran cantidad de casos de prueba fallidos que se obtuvieron con el enfoque de pruebas basadas en modelos, la mayoría fueron encontrados en un *end-point* del servicio web en particular. Debido a esto, algunos de los defectos encontrados pueden interpretarse como defectos similares. Sin embargo, esto se considera valioso dado que la concretización de los casos de pruebas con la herramienta permite generar un conjunto de entradas que el equipo de desarrollo no había comprobado. Del mismo modo, el enfoque permite encontrar fallos que no habían sido detectados por medio de las pruebas manuales.

Por otro lado, a diferencia de las pruebas realizadas con *TestOptimal* que solamente evaluaban el código de estado devuelto por el servidor, las pruebas manuales revisan adicionalmente el formato de la respuesta JSON. Como producto de esto, el enfoque de pruebas basadas en modelos no logra encontrar varios defectos que sí detectan las pruebas manuales. Esto nos indica que, aunque se encontró mayor cantidad de defectos por errores del servidor de una forma más ágil y rápida con *TestOptimal*, es igualmente necesario revisar manualmente los resultados de las respuestas exitosas, es decir, con código de estado diferentes de 500, para comprobar que la estructura del JSON recibido es el esperado. Esto implica que los enfoques de pruebas automatizadas y manuales deben ser combinados para aprovechar lo mejor de ambos enfoques. Es un trabajo en progreso intentar automatizar la comprobación de la estructura de las respuestas en formato JSON. Los casos de prueba manuales y los generados por *TestOptimal*

presentaron ciertas diferencias, factores que confirman la necesidad de la combinación de ambos enfoques en un ambiente de desarrollo. Lo anterior podría verse como una dependencia complementaria ya reportada previamente por Dudekula, Katam, Petersen & Mäntylä (2013) quien por medio de una encuesta a 115 profesionales del software hallaron que el 80% de los encuestados opinaba que no les parecía acertado pensar que las pruebas automáticas reemplazarían completamente las pruebas manuales. Finalmente, el uso de *TestOptimal* para pruebas automatizadas es un enfoque con el cual un equipo multifuncional puede generar pruebas. Sin embargo, para el mantenimiento de estas pruebas en caso de cambios producidos durante las iteraciones, se podría necesitar una persona dedicada a la actualización de los modelos.

Con respecto a la percepción de los participantes, estos en su mayoría estuvieron de acuerdo en que la herramienta resulta útil a la hora de mejorar el proceso de detección de fallos y por consiguiente puede mejorar la calidad del producto. Sin embargo, se obtiene una percepción general de que la herramienta no es sencilla de utilizar y que un profesional requiere una importante cantidad de esfuerzo para volverse habilidoso a la hora de manipularla. Esta percepción pudo verse afectada, principalmente, debido a los errores técnicos encontrados por el primer grupo, y la incapacidad de completar su modelo. Esta forma de calificar la experiencia con *TestOptimal* puede derivarse de que en este caso el equipo debió aprender tanto de un tipo nuevo de pruebas como de una herramienta ajena a su trabajo diario. Así mismo, la implementación de pruebas basada en modelos con *TestOptimal* dificultó la realización del modelado, lo que pudo haber influido negativamente en la intención de uso.

Aun así, aunque un equipo si logró completar su modelo, ambos indicaron que la herramienta en general no es fácil de utilizar. En este caso debemos considerar que todo el proceso se realizó en tres horas y una hora de capacitación, por lo cual desde nuestra perspectiva consideramos que obtuvimos resultados interesantes durante el caso de estudio y es interesante realizar otros estudios similares. Los datos indican que no hay intención de uso futuro de la herramienta *TestOptimal*, pero no descartan utilizar un enfoque de pruebas basada en modelos para agilizar sus procesos de prueba. Los profesionales percibieron la necesidad de invertir gran cantidad de esfuerzo y tiempo para poder utilizar la herramienta. Esto a su vez, entra en conflicto con el deseo de agilizar el proceso de desarrollo y pruebas.

4.2. Lecciones aprendidas

Las lecciones aprendidas obtenidas del proceso del estudio arroja dos principales experiencias: (a) La inducción de dos temas nuevos en un equipo de desarrollo, específicamente pruebas basadas en modelos y el uso de una herramienta de pruebas nueva, genera gran cantidad de dudas que se deben evacuar incluso en medio de la ejecución del ejercicio práctico a pesar de realizar una capacitación previa. (b) La aplicación de una encuesta debe conducirse con perfiles similares a los que contestarán la encuesta dado que a pesar de sus habilidades técnicas, no se puede asumir que conocen conceptos novedosos en un área particular.

Finalmente, se debe tener en cuenta que el estudio realizado presenta amenazas a la validez derivadas del trabajo con personas, de la utilización de una encuesta para recolectar datos de percepción y un sistema bajo prueba muy específico. El análisis

por parte de los investigadores podría introducir un sesgo en los resultados lo que se debe considerar como una amenaza a la validez. Finalmente, los resultados del caso de estudio no pueden ser generalizados fuera de las restricciones impuestas por el estudio y su contexto.

5. Conclusiones

En este caso de estudio se discutió el proceso realizado para introducir pruebas basadas en modelos para automatizar las pruebas de servicios web en un equipo que implementa prácticas ágiles. Los resultados indican que las pruebas basadas en modelos pueden ser utilizadas para el diseño y ejecución de pruebas funcionales de servicios web *RESTful*. La aplicación del enfoque permite aumentar la cantidad de casos de prueba y defectos encontrados.

Desde la perspectiva de los desarrolladores que implementan prácticas ágiles, a pesar de que el uso del enfoque de pruebas basadas en modelos podría mejorar la calidad del producto, es un enfoque que requiere mucho esfuerzo para el aprendizaje y la implementación. Los miembros del equipo consideran que para mejorar la aceptación de estos enfoques durante el desarrollo de un proyecto ágil, el conocimiento para el modelado y las herramientas de apoyo son esenciales. Los resultados indican que, a pesar de que se logra una mejora en la generación de casos de pruebas automatizados y la detección de errores, se percibe como una tarea compleja de aplicar.

Son necesarios más estudios que evidencien si los enfoques de pruebas basadas en modelos pueden adaptarse como una solución efectiva para la automatización de pruebas en ambientes de desarrollo ágiles. Es de interés realizar experiencias similares con otras herramientas que puedan facilitar el proceso. Sin embargo, existe la dificultad de encontrar herramientas disponibles y bien documentadas de pruebas basadas en modelos en este contexto. Adicionalmente, resultaría interesante realizar el estudio durante varias iteraciones de la metodología ágil, para determinar si la percepción y la eficacia de la herramienta varían con el tiempo y si se acopla fácilmente al constante cambio presente en estas metodologías.

Referencias

Alaimo, M. (2013). Proyectos ágiles con Scrum flexibilidad, aprendizaje, innovación y colaboración en contextos complejos. Buenos Aires: Kleer.

Arcuri, A. (2017). RESTful API automated test case generation. *Proceedings - 2017 IEEE International Conference on Software Quality, Reliability and Security, QRS 2017*, 9–20. https://doi.org/10.1109/QRS.2017.11

Belli, F., Endo, A.T., Linschulte, M., & Simão, A.D. (2014). A holistic approach to model-based testing of Web service compositions. Softw., Pract. Exper., 44, 201–234.

Chakrabarti, S. K., & Kumar, P. (2009). Test-the-REST: An approach to testing RESTfulweb-services. *Computation World: Future Computing, Service Computation, Adaptive, Content, Cognitive, Patterns, ComputationWorld* 2009, 302–308. https://doi.org/10.1109/ComputationWorld.2009.116

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS quarterly, 319–340.

De Cleva Farto, G., & Endo, A. T. (2015). Evaluating the model-based testing approach in the context of mobile applications. *Electronic Notes in Theoretical Computer Science*, *314*, 3–21. https://doi.org/10.1016/j.entcs.2015.05.002

Dudekula, M.R., Katam, K.M., Petersen, K., & Mäntylä, M. V.(2012). Benefits and limitations of automated software testing: Systematic literature review and practitioner survey. 2012 7th International Workshop on Automation of Software Test (AST),36-42. doi: 10.1109/IWAST.2012.6228988

qsEntin, V., Winder, M., Zhang, B., & Christmann, S. (2012). Introducing model-based testing in an industrial scrum project. *2012 7th International Workshop on Automation of Software Test, AST 2012 - Proceedings*, 43–49. https://doi.org/10.1109/IWAST.2012.6228989

Fertig, T., & Braun, P. (2015). Model-driven Testing of RESTful APIs. *Proceedings of the 24th International Conference on World Wide Web - WWW '15 Companion*, 1497–1502. https://doi.org/10.1145/2740908.2743045

Hossen, K., Groz, R., Oriat, C., & Richier, J. L. (2014). Automatic model inference of web applications for security testing. *Proceedings - IEEE 7th International Conference on Software Testing, Verification and Validation Workshops, ICSTW 2014*, 22–23. https://doi.org/10.1109/ICSTW.2014.47

Khan, T. A., & Heckel, R. (2009). A methodology for model-based regression testing of web services. *TAIC PART 2009 - Testing: Academic and Industrial Conference - Practice and Research Techniques*, 123–124. https://doi.org/10.1109/TAICPART.2009.23

Krogdahl, P & Luef, G & Steindl, C. (2005). Service-oriented agility: an initial analysis for the use of agile methods for SOA development. 93 - 100 vol.2. 10.1109/SCC.2005.86.

Nugroho, L. E., Azis, A., Mustika, I. W., & Selo. (2017). Development of RESTful API to support the oil palm plantation monitoring system. *Proceedings - 2017 7th International Annual Engineering Seminar, InAES 2017*. https://doi.org/10.1109/INAES.2017.8068545

Pinheiro, P., Endo, A., & Simao, A. (2010). Model-Based Testing of RESTful Web Services Using UML Protocol State Machines. *Sjc.Unifesp.Br*, (Icmc).

Reider, Martin & Magnus, Stephan & Krause, Jan. (2018). Feature-Based Testing by Using Model Synthesis, Test Generation and Parameterizable Test Prioritization. 130-137. 10.1109/ICSTW.2018.00041.

Schulze, C., Ganesan, D., Lindvall, M., McF Omas, D., & Cudmore, A. (2013). Model-based testing of NASA's OSAL API - An experience report. *2013 IEEE 24th International Symposium on Software Reliability Engineering, ISSRE 2013*, 300–309. https://doi.org/10.1109/ISSRE.2013.6698883

Schwaber, K., & Sutherland, J. (2013). The Scrum Guide. *Scrum.Org and ScrumInc*, (July), 17. https://doi.org/10.1053/j.jrn.2009.08.012

Sivanandan, S., & Yogeesha, C.B. (2014). Agile development cycle: Approach to design an effective Model Based Testing with Behaviour driven automation framework. *20th Annual International Conference on Advanced Computing and Communications (ADCOM)*, 22–25.

Thopate, H., & Kachewar, R. R. (2012). Chameleon model based automation framework design for testing in agile environments. *2012 CSI 6th International Conference on Software Engineering, CONSEG 2012*. https://doi.org/10.1109/CONSEG.2012.6349516

Utting, M., & Legeard, B. (2010). *Practical model-based testing: a tools approach*. Elsevier.

Villalobos-Arias, L., Quesada-Lopez, C., Martinez, A. & Jenkins, M. (2018) A tertiary study on model-based testing areas, tools and challenges: Preliminary results. In 21st Ibero-American Conference on Software Engineering, pp.

Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., & Wesslén, A. (2012). Experimentation in software engineering. Springer Science & Business Media.

Wu, C.-S., & Huang, C.-H. (2013). The Web Services Composition Testing Based on Extended Finite State Machine and UML Model. *2013 Fifth International Conference on Service Science and Innovation*, 215–222. https://doi.org/10.1109/ICSSI.2013.46

Yaghoubi, M., & Babanezhad, M. (2013). Software Developing with Agile Methods and Combination of Architecture, *65*(19), 33–37. Retrieved from http://research.ijcaonline.org/volume65/number19/pxc3886071.pdf

Zheng, Y., Zhou, J., & Krause, P. (2007). An automatic test case generation framework for web services. *Journal of Software*, *2*(3), 64–77. https://doi.org/10.4304/jsw.2.3.64–77

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 16/10/2018
Aceitação/Acceptance: 19/11/2018

Máquinas de soporte vectorial para clasificación supervisada de imágenes en bases de datos espaciales

Raúl Alejandro Murillo Castañeda[1], Álvaro Enrique Ortiz Davila[2]

raamurilloc@correo.udistrital.edu.co, aeortizd@udistrital.edu.co

[1] Universidad Distrital francisco José de Caldas, Vergel, 111411, Bogotá, Colombia.

[2] Universidad Distrital francisco José de Caldas, Ciudad Salitre Oriental, 111321, Bogotá, Colombia.

Pages: 538–551

Resumen: Este artículo está orientado a la implementación del método de clasificación supervisada Maquinas de Soporte Vectorial (MSV) sobre imágenes provenientes de sensores remotos almacenadas en una base de datos espacial de tipo relacional que permita contribuir a la clasificación de imágenes, donde se consiga además almacenar estos resultados dentro del mismo sistema manejador de bases de datos. Dado que el algoritmo de clasificación MSV permite tener una muy buena exactitud en el diagnóstico de las diferentes coberturas presentes en el suelo, se implementará como técnica de clasificación. Para corroborar el correcto funcionamiento de la herramienta se presenta un caso de estudio relacionado con la deforestación de la amazonia colombiana donde se demuestra la utilidad de la aplicación por medio de una clasificación supervisada la cual será comparada con el módulo de clasificación de algunos software que la implementan en la actualidad.

Palabras-clave: Maquinas de soporte vectorial; coberturas; imágenes raster; clasificación; base de datos espacial.

Implementation of the vector support machines method in spatial databases for supervised classification analysis in remote sensor images.

Abstract: This article is oriented to the implementation of the supervised classification method Vectorial Support Machines (VSM) on images from remote sensors stored in a spatial database of relational type that allows to contribute to the classification of images, where it is also possible to store these results within the same database manager system. Since the VSM classification algorithm allows to have a very good accuracy in the diagnosis of the different coverages present in the soil, it will be implemented as a classification technique. To corroborate the correct functioning of the tool, a case study is presented related to the deforestation of the Colombian Amazon where the usefulness of the application is demonstrated by means of a supervised classification which will be compared with the classification module of some software that the implemented today.

Keywords*:* Vector support machines; coverage; raster images; classification; spatial database.

1. Introducción

Los resultados obtenidos por los algoritmos de clasificación supervisada de imágenes traídas de los diferentes sensores remotos tienen muchas aplicaciones en el ámbito cartográfico, ya que pueden ser estas aplicaciones medioambientales, sociales, políticas, entre otras. Estos resultados son los principales elementos que utilizan muchos sistemas de información geográficos (SIG) para ofrecer al usuario la información requerida, este es el caso de los mapas de usos y cobertura del suelo, prevención de desastres naturales, evaluación del impacto en cuanto a los usos de los recursos naturales o el estudio del cambio climático. Con el acelerado crecimiento del volumen de estos datos obtenidos por imágenes de sensores remotos, se ha configurado como una necesidad almacenar esta información de forma estructurada y fácil de consultar(Jaime, Larín, & Garea, 2011), por lo que es de gran importancia almacenar esta información en bases de datos espaciales.

Las tendencias actuales para el tratamiento de la información geoespacial están enfocadas en el desarrollo de técnicas basadas en la representación y recuperación de estos datos desde un punto de vista semántico. En este sentido cuando hablamos de información de tipo raster aún falta mucho para que estos datos almacenados sobre una base de datos espacial se asemejen al nivel de abstracción que se logra con la información vectorial. El hecho de lograr que los análisis se hagan directamente en la base de datos le da una ventaja en ejecución a los procesos analíticos realizados por los sistemas manejadores de bases de datos(Shekhar Sashi, 2003), por lo que es muy importante lograr acercar los análisis de la información espacial en la base de datos a todos los niveles incluyendo la información raster. Por lo tanto en este artículo, se implementó una nueva herramienta que mediante la utilización del método de clasificación supervisada basado en MSV se brinde resultados precisos y ajustados, que además estén soportados sobre una plataforma no propietaria. La herramienta desarrollada se fundamentó en el algoritmo de MSV, ya que a partir de diferentes estudios se ha comprobado que genera muy buenos resultados(Argañaraz & Entraigas, 2011),(Canales, Zhang, & Liu, 2009),(Castellon, 2015)uno comercial –ENVI- y una librería de algoritmos de código abierto -Orfeo Toolbox-. Lo anterior teniendo en cuenta de que a pesar de que se usa el mismo método de clasificación los requerimientos para ejecutarlo y la implementación varían de un software a otro.","note":"DOI: 10.13140/RG.2.1.1991.1844","author":[{"family":"Castellon","given":"Juan"}],"issued":{"date-parts":[["2015",6,24]]}}}],"schema":"https://github.com/citation-style-language/schema/raw/master/csl-citation.json"} además de su facilidad en el manejo de grandes volúmenes de información, puesto que logra altos niveles de confiablidad en los resultados(Anzola, 2016).

2. Marco Teórico

2.1. Clasificación de imágenes

La clasificación de imágenes es un proceso que consiste en agrupar los pixeles de una imagen en un número finito de clases, basándose en los valores espectrales de las distintas bandas, convirtiendo de este modo la información captada por los sensores como niveles digitales a una escala categórica fácil de interpretar(Mather P., Tso B., 2009). Los pixeles que pertenezcan a la misma clase deberán tener unas características espectrales similares(Schowengerdt R.A., 1985). Los algoritmos de clasificación de imágenes son

una de las técnicas más importantes utilizadas en el ámbito de la teledetección, ya que facilitan la interpretación de una gran cantidad de información contenida en sus bandas. El objetivo de los algoritmos de clasificación de imágenes consiste en dividir los pixeles de la imagen en distintas clases, llamadas clases espectrales, teniendo en cuenta la similitud existente entre dichos pixeles. La clasificación de una imagen es una tarea que se realiza con el propósito de convertir datos cuantitativos (generalmente los niveles digitales de los píxeles en cada banda espectral) en datos cualitativos (temas o clases que son importantes en un dominio específico del conocimiento)(Richards, J. A. & Jia X., 1999). La motivación principal de una clasificación es la de representar un fenómeno que ocurre sobre la superficie terrestre a partir de la generalización y agrupación de datos obtenidos mediante sensores remotos(Jensen J.R, 2005). Una buena clasificación debe representar de manera exacta la realidad a partir de las características pictórico-morfológicas presentes en las imágenes(ASPRS, 1997).

2.3. Maquinas de Soporte Vectorial (MSV)

Las MSV son un conjunto de algoritmos de aprendizaje supervisado desarrollados por (Vapnik & Cortés, 1995) y su equipo AT&T, que han surgido como métodos relacionados con problemas de clasificación y regresión. Su buen desempeño ha llevado a su uso en una gran variedad de problemas, algunos investigadores (Fletcher, Hussain, & Shawe-Taylor, 2010),(Huang, Nakamori, & Wang, 2005),(Argañaraz & Entraigas, 2011),(Mountrakis, Im, & Ogole, 2011),(Espinosa, Sánchez, & Castilla, 2014), han utilizado MSV para solucionar problemas de clasificación y regresión relacionados a la predicción de series de tiempo y clasificación de imágenes, mostrando tener muy buenos resultados en comparación a otras metodologías tradicionales como modelos econométricos, redes neuronales y en algunos casos a modelos de aprendizaje de máquina. La construcción de las MSV se basa en la idea de transformar o proyectar un conjunto de datos pertenecientes a una dimensión n dada, hacia un espacio de dimensión superior aplicando una función kernel – Kernel Trick(Alpaydin, 2010) para que a partir del nuevo espacio creado, se operen los datos como si se tratase de un problema de tipo lineal, resolviendo el problema sin considerar la dimensionalidad de los datos(Alpaydin, 2010). La idea detrás de las MSV es que a partir de unos inputs de entrada al modelo, se etiquetan las clases y se entrena la MSV construyendo un modelo que sea capaz de predecir la clase de los nuevos datos que se introduzcan al modelo.

2.4. Funciones de desición dentro de las MSV

Se considera el problema de clasificación de un pixel cuyas características están dadas por el vector X tal que $X = (X_1, \ldots, X_p)^T$ pertenece a una de dos clases posibles. Ahora se supone que se tienen las funciones $F_1(x)$ y $F_2(x)$ que definen las clases 1 y 2 y se clasifica al pixel X dentro de la clase 1 si:

$$F_1(x) > 0, F_2(x) < 0,$$

O clasificamos al pixel X dentro de la clase 2 si:

$$F_1(x) < 0, F_2(x) > 0,$$

Estas funciones se denominan funciones de decisión. Al proceso de encontrar las funciones de decisión a partir de pares de entrada-salida es llamado entrenamiento. Los métodos convencionales de entrenamiento determinan las funciones de decisión de tal forma que cada par entrada-salida sea correctamente clasificado dentro de la clase a la que pertenece. La Figura 2 muestra un ejemplo que asume a los cuadros como la clase 1 y los círculos como la clase 2, resulta claro que los datos de entrenamiento no se intersectan en ningún momento y es posible trazar una línea separando los datos de manera perfecta(Canales et al., 2009):

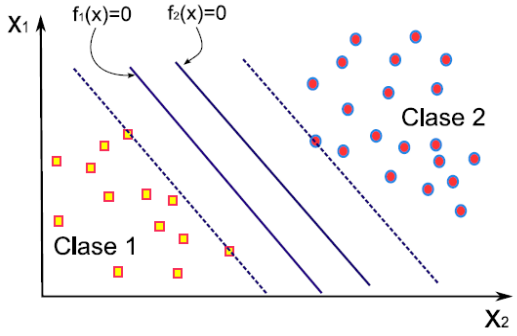


Figura 1 – Funciones de decisión(Canales et al., 2009)

2.5. Indicadores de calidad de los clasificadores

Un tema clave es la evaluación de fiabilidad de los procedimientos empleados y su calidad en los resultados finales. La calidad final de un método de clasificación o de combinación de clasificadores, se puede extraer a partir de un coeficiente de precisión, de forma que se otorga mayor confianza a aquel clasificador que haya demostrado mayor acierto. Este es el caso de la matriz de confusión de la cual se derivan distintos índices de calidad(Borràs et al., 2017). La matriz de confusión, consta de una tabla de doble entrada, que confronta los valores reales o verdad de terreno con los resultados de la clasificación. La diagonal de la matriz muestra la cantidad de píxeles reales y de la clasificación que coinciden por categoría, mientras que los restantes, vienen a mostrar aquellos que se confunden con otras categorías. Es decir, en sentido vertical se representa el porcentaje de píxeles reales que se confundieron en la verdad de terreno, y en sentido horizontal el porcentaje de píxeles después de la clasificación que se han confundido(Borràs et al., 2017). El índice Kappa propuesto por(Cohen, 1960) se usa para evaluar la concordancia de métodos cuyo resultado es categórico, con dos o más clases. Este índice representa la proporción de acuerdos observados respecto del máximo acuerdo posible más allá del azar. En la interpretación del índice Kappa hay que tener en cuenta que el índice depende del acuerdo observado, pero también de la prevalencia del carácter estudiado y de la simetría de los totales marginales(V. Abraira, 2001). En este sentido Cohen propuso el denominado índice kappa (K), que definió como(Cohen, 1960):

$$K = \frac{p_o - p_e}{1 - p_e} \quad (1)$$

Siendo la proporción de acuerdos observados y la proporción de acuerdos esperados en la hipótesis de independencia entre los observadores, es decir, de acuerdos por azar.

Finalmente Landis y Koch propusieron, la escala de valoración del índice Kappa que se describe en la tabla 1(Landis & Koch, 1977).

Kappa (K)	Grado de acuerdo
< 0,00	Sin acuerdo
0,00 - 0,20	Insignificante
0,21 - 0,40	Mediano
0,41 - 0,60	Moderado
0,61 - 0,80	Sustancial
0,81 - 1,00	Casi Perfecto

Tabla 1 – Valoración de índice Kappa

3. Metodologia

A continuación se menciona la metodología utilizada para la elaboración del proyecto la cual se describe en la figura 3.

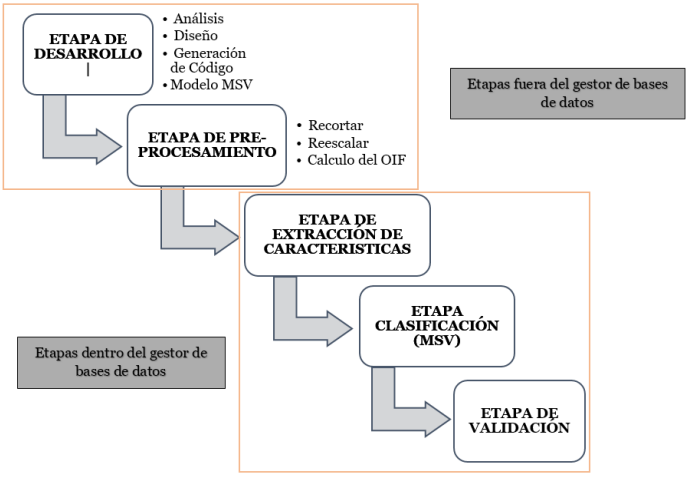


Figura 2 – Metodología propuesta

3.1. Etapa de desarrollo.

Esta sección se centra en la arquitectura de la herramienta, que está formada por una combinación de componentes relacionados con el ámbito de la teledetección y componentes relacionados con el modelo matemático de máquinas de soporte vectorial. La Figura 4 muestra la arquitectura principal del algoritmo MSV.

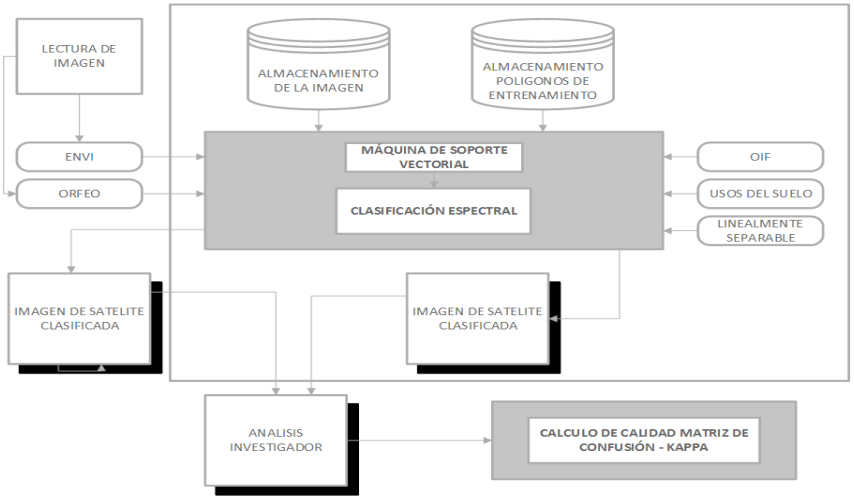


Figura 3 – Arquitectura del Algoritmo desarrollado

El algoritmo MSV tiene varios componentes de entrada y un solo componente de salida: la imagen clasificada. Para la aplicación desarrollada se tienen dos entradas: la aplicación de lectura de la imagen y la aplicación de lectura del conjunto de muestras (Polígonos de entrenamiento) los cuales fueron seleccionados, tomando de la propia imagen de satélite, varias muestras de las diferentes clases, dichas muestras están formadas por pixeles representativos que componen el denominado conjunto de entrenamiento (o firmas), sobre el que posteriormente se basó el proceso de clasificación. Este algoritmo por ser de tipo supervisado compara cada pixel de la imagen con estas firmas elegidas y, a continuación, cada pixel es etiquetado en la clase a la que más se asemeja espectralmente a partir de los pixeles que pertenecen a las clases elegidas para el proceso de clasificación.

3.2. Etapa de pre-procesamiento.

Esta sección se centra en la recolección, obtención y procesamiento digital de las imágenes a trabajar. La Figura 5 muestra la arquitectura principal de la etapa de pre – procesamiento.

La etapa de pre-procesamiento consiste en dejar la imagen con el menor porcentaje de ruido posible a partir de correcciones radiometrías y geométricas, para aumentar las posibilidades de éxito en las siguientes fases. Para lo cual se definieron las siguientes sub-etapas:

Recorte: es la parte donde se toma una porción de la escena, la cual contendrá la zona de estudio y permitirá reducir los costos de procesamiento ya que se reducirán el número de filas y columnas de la imagen y por ende los procesos serán más rápidos. El tamaño de la imagen seleccionada es de 2048 x 2048 pixeles (un total de 4.194.304 pixeles).

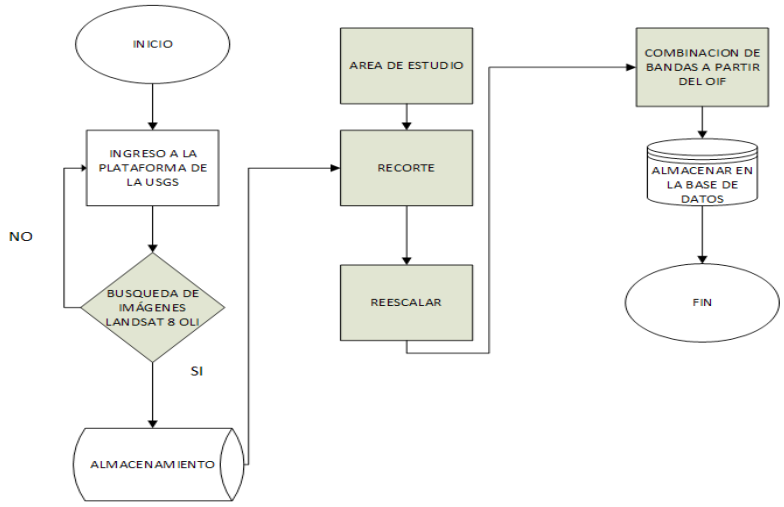


Figura 4 – Arquitectura Pre-procesamiento

Reescalar: Se procesa la imagen para dejarla en un formato estándar de 8 bits ya que se adaptara la resolución radiométrica de la imagen a la capacidad de visualización del monitor. Cada sensor codifica la reflectancia en un número de determinados niveles digitales. Ese rango digital puede no corresponder con el número de niveles de visualización que facilita la consola gráfica, por lo que resulta preciso ajustar estos parámetros, ya que al trabajar con una paleta de colores entre 0-255 niveles de color se ajustan estos niveles y además se ocupa menos espacio en la memoria del ordenador, por lo que se acelera la manipulación de la misma.

Combinación de imágenes: Permite visualizar, simultáneamente imágenes sobre diferentes partes del espectro, lo que facilita una identificación visual más precisa de algunas coberturas que ayudan a realizar una mejor clasificación. La elección de las bandas para realizar la combinación y el orden de los colores dependen del sensor elegido y la naturaleza del trabajo a realizar, para el caso de esta investigación por tratarse de un sensor Landsat 8 OLI dotado de 11 bandas de las cuales se dispuso de 7 (azul, verde, rojo, Infrarrojo cercano, dos infrarrojos lejanos y la pancromática), se tiene la posibilidad de realizar combinaciones en un amplio rango de composiciones de color. El método que permitió seleccionar la combinación de bandas que contiene la mayor cantidad de información con la menor cantidad de redundancia entre las 35 posibles combinaciones (7 tomadas de 3, sin repeticiones) se denomina Índice de Factor Optimo OIF por sus siglas en ingles. El mayor valor de OIF corresponde a la combinación de bandas con menor correlación entre ellas y con mayor desviación estándar para cada banda, indicando la mayor posibilidad de discriminación de coberturas estudiadas. Por lo tanto según el resultado obtenido la mejor combinación de bandas corresponde a la combinación RGB – 356 de OLI, correspondiente al Verde, Infrarrojo cercano, y SWIR 1.

3.3. Etapa de almacenamiento en la base de datos.

En esta etapa se definen los parámetros para realizar el almacenamiento de las imágenes en la base de datos espacial que se utilizó posteriormente para realizar la clasificación supervisada de las imágenes. La extensión para el almacenamiento de imágenes para el gestor de base de datos postgres se denomina Postgis la cual permite convertir datos raster que estén en formatos soportados por GDAL en un archivo sql que posteriormente puede ser ejecutado desde el motor de base de datos.

3.4. Etapa de extracción de características

Para la extracción de características se analizó la imagen según la combinación de bandas elegida, teniendo en cuenta que es necesario realizar un estudio de las variables ecológicas de vegetación más relevantes dentro de la zona de estudio, esto con el fin de poder establecer de manera más precisa las clases que se deben tener en cuenta dentro del algoritmo y comenzar a realizar una separación de pixeles según las características identificadas. La Tabla 2 muestra las características de las 7 clases que se pueden encontrar en la imagen de satélite.

Clase	Nombre Clase
C1	Cuerpos de agua
C2	Deforestación (Bosque fragmentado)
C3	Bosque natural
C4	Casco urbano
C5	Caño cristales
C6	Suelo desnudo
C7	Vías

Tabla 2 – Características principales presentes en la imagen satelital

Una vez se establecieron las clases, se definió una sentencia sql dentro de la base de datos que permitió ejecutar una función de análisis espacial denominada intersección entre los polígonos de entrenamiento almacenados y los pixeles de la imagen para de esta manera establecer las zonas consideradas como verdad absoluta de terreno y asociarle los niveles digitales correspondientes a cada una de las clases definidas anteriormente. Esta sentencia SQL recorre pixel a pixel la imagen y genera un recorte de la misma solo con aquellos registros de pixeles que se encuentran dentro de algún polígono de entrenamiento, cumpliendo con la particularidad de recortar los pixeles si y solo si el centroide de este se encuentra dentro del polígono, esto es muy importante ya que gracias a esta particularidad se estableció una condición de frontera que permite eliminar ruido en la muestra.

3.5. Etapa de Clasificación MSV

Esta investigación propone un algoritmo de clasificación de imágenes supervisado, basado en el algoritmo MSV. El algoritmo se implementó directamente sobre una

base de datos espacial desde el gestor Postgres aprovechando las propiedades raster definidos en la extensión Postgis sobre las imágenes, con el propósito de implementar un proceso de clasificación supervisado espectral. Por lo tanto en esta etapa aplicamos el algoritmo planteado en la etapa de desarrollo sobre una imagen satelital Landsat 8 OLI para una zona piloto en el municipio de la Macarena Meta perteneciente a la amazonia colombiana, donde se analizó el comportamiento del algoritmo comparándolo con los paquetes de procesamiento digital de imágenes ENVI para el caso de uso de software comercial y ORFEO para el caso de software libre, y se establecieron las ventajas y desventajas de cada uno de los algoritmos, además se comparó a través de cada matriz de confusión y coeficiente Kappa el rendimiento de cada clasificador para concluir según los resultados obtenidos.

3.6. Etapa de Validación de la clasificación de imágenes

Para la validación de la clasificación existen dos posibilidades, evaluar una estimación teórica del error en función de las características del algoritmo de clasificación o analizar una serie de áreas test obtenidas del mismo modo que las áreas de entrenamiento. Para la presente investigación se ha implementado el segundo modo de proceder ya que permite obtener una estimación más realista de los errores mientras la muestra de pixeles para la estimación del error sea lo suficientemente grande y representativa. Para la evaluación de los errores se utilizó una matriz de confusión de clases ya que con este tipo de análisis, se obtuvo, no sólo una caracterización del error cometido, sino también una medida sobre la adecuación de la clasificación considerada a la realidad y de los parámetros utilizados para caracterizarlas.

4. Resultados

El objetivo principal de esta investigación, fue comprobar el funcionamiento del algoritmo respecto a otros softwares que implementan la clasificación supervisada de imágenes por el método MSV para validar su funcionamiento. La comparativa de la extensión desarrollada con otros softwares de clasificación se ha llevado a cabo sobre una imagen de satélite Landsat 8 OLI, dentro de la cual se seleccionó la mejor combinación de bandas 356 a partir del OIF, con una resolución espacial de 30x30m. La región de estudio elegida fue el municipio de La Macarena, perteneciente al departamento del Meta, el cual es limítrofe con la Amazonia Colombiana por el norte. El caso de estudio planteado dentro de esta investigación corresponde al análisis de la deforestación en esta región. A continuación, se presenta una tabla donde se aprecia comparativamente los resultados obtenidos con los diferentes softwares, aplicando la clasificación supervisada MSV sobre la misma imagen.

Según la tabla anterior, las tres clasificaciones han tenido un resultado satisfactorio. En los tres casos cada una de las clases fuerón bien definidas y se logra ver claramente la disociación entre las mismas. Sin embargo algunas clases como “cuerpos de agua (azul)” y “casco urbano (rojo)” mostraron un poco de confusión, ya que algunos pixeles presentan características similares entre las dos coberturas, para el caso de ENVI y ORFEO esta confusión de pixeles es un poco más notoria que en la imagen obtenida a partir del algoritmo desarrollado.

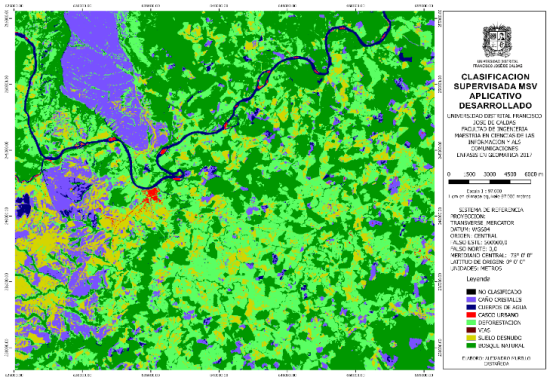
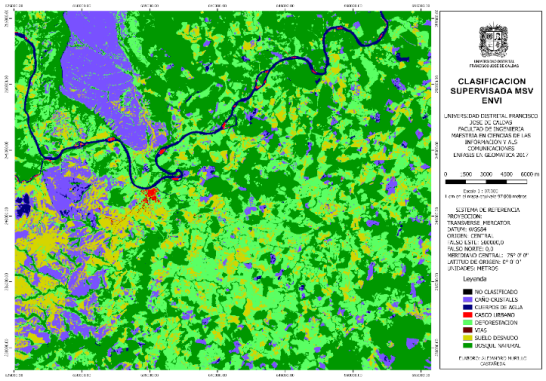
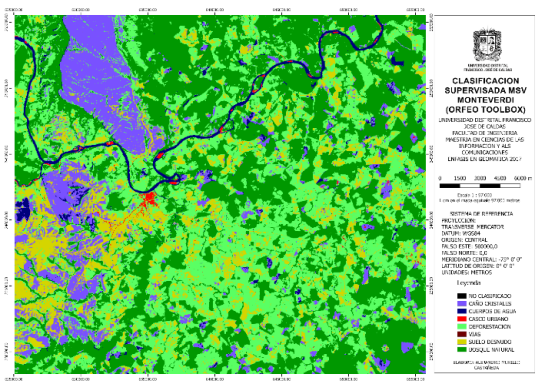
Resultado de la Clasificación de MSV	Software usado
	Aplicativo desarrollado con Python, PostgreSQL y PostGIS
	Environment fo Visializing Images (ENVI)
	Software libre MONTEVERDI y su extensión ORFEO Toolbox

Tabla 3 – Comparación de resultados de aplicar MSV sobre la misma imagen

Respecto al análisis de la deforestación las tres clasificaciones son muy concluyentes al mostrar el avance indiscriminado en la zona.

4.1. Matriz de confusión del aplicativo desarrollado

Ahora comprobaremos los resultados obtenidos a través de la matriz de confusión y su respectivo coeficiente Kappa para el software desarrollado durante este trabajo de investigación (Tabla 4), como también para el software ENVI (Tabla 5), y Monteverdi ORFEO (Tabla 6).

		Validación							Total	Comisión	EU(%)	RP(%)
		Caño Cristales	Cuerpos de Agua	Casco Urbano	Deforestación	Vías	Suelo Desnudo	Bosque Natural				
PREDICCION	Caño Cristales	3.120	-	-	24	-	58	-	3.202	82	98%	2%
	Cuerpos de Agua	-	2.001	-	-	-	-	-	2.001	-	100%	0%
	Casco Urbano	-	-	239	-	-	2	-	241	2	99%	1%
	Deforestación	140	1	88	5.585	4	2.285	985	9.088	3.503	72%	28%
	Vías	5	-	12	54	107	430	-	608	501	55%	45%
	Suelo Desnudo	8	-	-	8	-	2.727	-	2.743	16	99%	1%
	Bosque Natural	-	-	-	-	-	-	13.607	13.607	-	100%	0%
	Total	3.273	2.002	339	5.671	111	5.502	14.592	31.490			
	Omisiones	153	1	100	86	4	2.775	985				
	RP(%)	4,47%	0,05%	22,78%	1,49%	3,48%	33,53%	6,32%				
	EP/%)	96%	100%	77%	99%	97%	66%	94%				

Índices Globales: Porcentaje de Acuerdo: 86,96% Coeficiente Kappa: 0,8184

Tabla 4– Matriz de confusión clasificación supervisada MSV utilizando el aplicativo desarrollado

		Validación							Total	Comisión	EU(%)	RP(%)
		Caño Cristales	Cuerpos de Agua	Casco Urbano	Deforestación	Vías	Suelo Desnudo	Bosque Natural				
PREDICCION	Caño Cristales	3.153	-	-	133	8	8	-	3.302	149	96%	4%
	Cuerpos de Agua	-	2.027	-	-	-	-	-	2.027	-	100%	0%
	Casco Urbano	-	-	234	-	26	-	-	260	26	91%	9%
	Deforestación	4	-	2	5.585	56	5	-	5.652	67	99%	1%
	Vías	-	-	4	-	72	-	-	76	4	95%	5%
	Suelo Desnudo	83	-	60	2.375	391	2.778	-	5.687	2.909	66%	34%
	Bosque Natural	-	-	-	982	-	-	13.605	14.587	982	94%	6%
	Total	3.240	2.027	300	9.075	553	2.791	13.605	31.591			
	Omisiones	87	-	66	3.490	481	13	-				
	RP(%)	2,61%	0,00%	18,03%	27,78%	46,52%	0,46%	0,00%				
	EP/%)	97%	100%	82%	72%	53%	100%	100%				

Índices Globales: Porcentaje de Acuerdo: 86,90% Coeficiente Kappa: 0,8178

Tabla 5– Matriz de confusión clasificación supervisada MSV utilizando el software ENVI

Según la tablas anteriores el clasificador desarrollado en esta investigación fue el que mejor eficiencia presento ya que tuvo el índice Kappa más alto entre las 3 clasificaciones propuestas, esto se ve reflejado tanto en la imagen como en su matriz de confusión, logrando de esta manera realizar una clasificación con un alto índice de exactitud temática. En consecuencia, el algoritmo desarrollado optimiza la funcionalidad general de cualquier clasificación que se realice, y cuyo funcionamiento se base en la identificación adecuada de los pixeles de las imágenes en sus clases correspondientes.

		Validación							Total	Comisión	EU(%)	RP(%)
		Caño Cristales	Cuerpos de Agua	Casco Urbano	Deforestación	Vías	Suelo Desnudo	Bosque Natural				
PREDICCION	Caño Cristales	3.114	-	-	30	-	58	-	3.202	88	97%	3%
	Cuerpos de Agua	-	2.001	-	-	-	-	-	2.001	-	100%	0%
	Casco Urbano	-	-	239	-	-	2	-	241	2	99%	1%
	Deforestación	132	-	88	5.601	4	2.278	985	9.088	3.487	72%	28%
	Vías	5	-	107	50	12	434	-	608	596	50%	50%
	Suelo Desnudo	8	-	-	8	-	2.727	-	2.743	16	99%	1%
	Bosque Natural	-	-	-	-	-	-	13.607	13.607	-	100%	0%
	Total	3.259	2.001	434	5.689	16	5.499	14.592	31.490			
	Omisiones	145	-	195	88	4	2.772	985				
	RP(%)	4,26%	0,00%	31,00%	1,52%	20,00%	33,51%	6,32%				
	EP/%)	96%	100%	69%	98%	80%	66%	94%				

Indices Globales: Porcentaje de Acuerdo: 86,69% Coeficiente Kappa: 0,8146

Tabla 6– Matriz de confusión clasificación supervisada MSV utilizando el software MONTEVERDI (ORFEO)

5. Conclusiones

En la actualidad no se encontró evidencia de investigaciones relacionados con la implementación de técnicas de procesamiento de imágenes almacenadas en una base de datos espacial. Durante el desarrollo de este trabajo de investigación, se ha implementado un algoritmo de clasificación de imágenes basado en MSV el cual logro alcanzar su objetivo principal de generar una clasificación supervisada de imágenes de usos y coberturas del suelo con altos estándares de exactitud temática procesando esta información desde una base de datos espacial.

La aplicación utilizo el lenguaje SQL para implementar consultas estructuradas de análisis espacial convencional a favor de mejorar la selección de pixeles al momento de intersectarlos con las clases de entrenamiento y de esta manera se definió que es más provechoso para el algoritmo utilizar una agrupación de pixeles si y solo si contienen al centroide del mismo, ya que de esta manera las probabilidades de asociar de manera correcta cada pixel a un hiperplano según la clase de entrenamiento se aumentan, generando unos mejores resultados de clasificación.

Los resultados obtenidos por parte del aplicativo han sido muy satisfactorios desde varios puntos de vista, ya que el algoritmo de MSV ha sido una técnica que proporciona muy buenos resultados dentro del proceso de clasificación de imágenes de satélite. En primer lugar, en relación con la comparación con un software comercial como ENVI, el algoritmo ha conseguido una tasa de acierto de pixeles bien clasificados, lo cual permite inferir el buen funcionamiento del aplicativo y el cumplimiento frente a la calidad de la clasificación en comparación con los estándares de calidad presentes en un software de uso comercial. En segundo lugar al compararlo con la clasificación obtenida por medio

del software libre Monteverdi y su extensión raster OrfeoToolbox se puede ver que la tasa de asertividad del aplicativo también superior, lo cual genera una evidencia de la competitividad del aplicativo desarrollado frente a herramientas Open Source.

Referencias

Alpaydin, E. (2010). Introduction to machine learning (2nd ed). Cambridge, Mass: MIT Press.

Anzola, N. S. (2016). Máquinas de soporte vectorial y redes neuronales artificiales en la predicción del movimiento USD/COP spot intradiario. ODEON, 0(9), 113–172.

Argañaraz, J. P., & Entraigas, I. (2011). Análisis comparativo entre las máquinas de vectores soporte y el clasificador de máxima probabilidad para la discriminación de cubiertas del suelo. Revista de Teledetección, 36, 26–39.

ASPRS. (1997). Manual of Photographic Interpretation. Bethesda, USA: American Society for Photogrammetry and Remote Sensing.

Ayala R., Becerra A., Flores I.M, J.F, Diaz J.R.,. (1999). Evaluation of greenhouse covered extensions and required resources with satellite images and GIS. Almeria's case. Second European Conference of the European Federation for Information Technology in Agriculture, Food and the Environment, Bonn (Germany), 27–30, 1999.

Borràs, J., Delegido, J., Pezzola, A., Pereira, M., Morassi, G., & Camps-Valls, G. (2017). Clasificación de usos del suelo a partir de imágenes Sentinel-2. Revista de Teledetección, (48), 55.

Canales, J. C., Zhang, X. L., & Liu, W. Y. (2009). Clasificación de grandes conjuntos de datos vía Máquinas de Vectores Soporte y aplicaciones en sistemas biológicos. Instituto Politécnico Nacional, México DF.

Castellon, J. (2015). Análisis comparativo entre ENVI y Orfeo Toolbox SVM. Teledeteccion, 1.

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement, 20(1), 37–46.

Cristianini, N., & Shawe-Taylor, J. (2000). An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Cambridge University Press.

Espinosa, A. T., Sánchez, J. C. M., & Castilla, J. S. R. (2014). Identificar cobertura vegetal de suelo clasificando pixeles en imágenes hiperespectrales con svm (máquina de soporte vectorial), 20.

Fletcher, T., Hussain, Z., & Shawe-Taylor, J. (2010). Multiple Kernel Learning on the Limit Order Book, 8.

Huang, W., Nakamori, Y., & Wang, S.-Y. (2005). Forecasting stock market movement direction with support vector machine. Computers & Operations Research, 32(10), 2513–2522.

Jaime, E., Larín, R., & Garea, E. (2011). Hacia métodos de análisis de datos espaciales raster en el nivel semántico. Computación y Sistemas, 15(1), 91–106.

Jensen J.R. (2005). Introductory Digital Image Processing – A Remote Sensing Perspective (3a ed). New Jersey, USA: Prentice Hall.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. Biometrics, 33(1), 159–174.

Lizarazo, I. (2008). Clasificación de la cobertura y del uso del suelo urbano usando imágenes de satélite y algoritmos supervisados de Inteligencia Artificial. UD y la GEOMÁTICA, (2), 4–18.

Mather P., Tso B.,. (2009). Classification methods for remotely sensed data (2 Edition). CRC Press.

Mountrakis, G., Im, J., & Ogole, C. (2011). Support vector machines in remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing, 66, 247–259.

Richards, J. A. y Jia X. (1999). Remote Sensing Digital Image Analysis. Springer-Verl.

Schowengerdt R.A. (1985). Techniques for image processing and classification in remote sensing. Academic Press.

Shekhar Sashi, C. S. (2003). Spatial Databases: A Tour, (Vol. 1). Minnesota.

V. Abraira. (2001). El índice kappa. Medicina de Familia. SEMERGEN, 247–249.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 25/09/2018
Aceitação/Acceptance: 11/11/2018

Evaluación Subjetiva de la Calidad de Vídeo para el Servicio YouTube Live en un Dispositivo Móvil

Alexander Sócola[1], Marta Solera[1], Luís Roberto Jiménez[1], Matías Toril[1]

apsocola@uma.es, msolera@ic.uma.es, lrjp@ic.uma.es, mtoril@ic.uma.es

[1] Departamento de Ingeniería de Comunicaciones, Universidad de Málaga, Málaga – España.

Pages: 552–562

Resumen: La plataforma de descarga de vídeo más popular es YouTube, que además ofrece el servicio de videostreaming en directo (YouTube Live). La demanda de nuevos recursos asociados a estos servicios ha provocado cambios en la manera de gestionar las redes de telecomunicación, pasando de la concepción clásica basada en la calidad de servicio (QoS) a otra más moderna basada en la calidad de experiencia (QoE). Por esta razón, es muy importante para los operadores entender el impacto del vídeo en sus redes y la calidad de experiencia que percibe el usuario. El presente artículo describe una metodología para evaluar la calidad de experiencia de vídeo del servicio de YouTube Live en un dispositivo móvil. En la evaluación subjetiva de la calidad del vídeo participaron 24 observadores, que calificaron 29 secuencias de vídeo a diferentes resoluciones. Posteriormente, se analizaron los resultados para obtener la Nota Media de Opinión (MOS) en función de la resolución y movimiento de las secuencias.

Palabras-clave: Calidad de experiencia, MOS, videostreaming, YouTube Live, método de evaluación de calidad de video subjetiva.

Subjective Evaluation of Video Quality for the YouTube Live Service on a Mobile Device

Abstract: The most popular video platform is YouTube, which also offers live video streaming service (YouTube Live). The demand for new resources causes changes in the way of managing networks, going from the classical conception based on the quality of service (QoS) to a more modern approach based on the quality of experience (QoE). For this reason, it is very important for operators to understand the impact of video on their networks and the quality of experience that the end-user perceives. This paper describes a subjective methodology for evaluating the video image quality for YouTube Live service on a mobile device. In the subjective assessment, 24 observers rated 29 sequences at different resolutions. Survey results were analyzed to obtain the Mean Opinion Score (MOS) according to the resolution and motion of the sequences.

Keywords*:* Quality of Experience, MOS, videostreaming, YouTube Live, Subjective video quality assessment method.

1. Introducción

La aparición de nuevas tecnologías como el Internet de las Cosas (IoT) y la implementación de la quinta generación (5G) de telefonía móvil ha provocado el desarrollo de nuevos servicios y aplicaciones que requieren de terminales cada vez más potentes, redes con más recursos y otra forma de gestionarlas, pasando de los indicadores de calidad de servicio (QoS) a otros más modernos basados en la opinión del usuario y la calidad de experiencia (QoE).

En la última década, uno de los servicios con mayor éxito y demanda ha sido la reproducción de vídeos a través de la descarga progresiva (videostreaming), que se ofrece en diferentes plataformas, como YouTube, Netflix, Twitch, etc. Este éxito ha causado un crecimiento del volumen de tráfico en las redes de comunicaciones móviles, debido a diferentes factores, como el incremento del tiempo de visualización, vídeos incrustados en una mayor cantidad de medios audiovisuales y la tendencia de los usuarios a observar los vídeos en una resolución mucho más alta. El estudio realizado por [1] establece que, del tráfico total de datos en el 2017, un 56% corresponde al servicio de vídeo y se estima que para el año 2023 el tráfico de vídeo crecerá hasta representar el 73% del tráfico total de datos.

La plataforma de visualización de vídeos en línea YouTube, desde sus inicios en el 2005 [2], ha tenido un rápido crecimiento hasta convertirse en el sitio de reproducción de vídeos más popular en el mundo [3]. En el año 2010, se implementó un nuevo servicio que consiste en la transmisión de vídeo en tiempo real que se denominó YouTube Live, causando un aumento aún mayor de tráfico en redes fijas y móviles. Por esta razón, para los operadores de redes es muy importante entender el impacto del vídeo en tiempo real en sus redes y entender la calidad de experiencia que percibe el usuario de este servicio.

Para evaluar la calidad de imagen percibida por un usuario, existen varias metodologías, que se dividen en objetivas y subjetivas. Las primeras se basan en analizar una imagen degradada y compararla píxel a píxel con su imagen de referencia, obteniendo métricas que cuantifican la diferencia media (p. ej., MSE, PSNR). Este esquema generalmente no es suficiente para estimar la calidad percibida, ya que el sistema visual humano es extremadamente complejo y dependiente de diversos factores, como el tipo de aplicación (TV, vídeo conferencia, etc.), el lugar de la imagen en donde se produce la degradación (zona de textura homogénea o compleja), el tipo de dispositivo utilizado (televisor de gran formato o teléfono móvil) y del tamaño del monitor.

El enfoque más preciso usado a lo largo de los años para evaluar la calidad percibida por un usuario es la evaluación subjetiva, ya que no existe mejor indicador que el proporcionado por el propio ser humano [4]. El método consiste en recoger la opinión del usuario a una serie de reproducciones de vídeo. El resultado del experimento se recoge en términos de un indicador denominado nota media de opinión (Mean Opinion Score, MOS), que refleja el grado de satisfacción sobre la percepción de la calidad del vídeo en una escala numérica que va del 1 al 5, siendo 1 una percepción mala y 5 excelente [5]. Como inconveniente, este método resulta largo, costoso e implica a los usuarios.

La metodología para el desarrollo de experimentos subjetivos se realiza en base a recomendaciones de la Unión Internacional de Telecomunicaciones (ITU). Las normas

ITU-R BT.500-13 [6] e ITU-T P.910 [7] se han utilizado con éxito durante muchos años para realizar evaluaciones subjetivas de calidad audiovisual y calidad de vídeo. Estas recomendaciones se diseñaron inicialmente para entornos de servicios fijos de vídeo para televisión, para ambientes silenciosos y sin distracciones. Con el pasar de los años estas recomendaciones se han actualizado y ampliado a medida que la tecnología ha avanzado para incluir nuevos paradigmas de observación de vídeo, cómo el servicio de vídeo por internet y la calidad de distribución de televisión ITU-T P.913 [8].

Diversos trabajos recogen métodos objetivos o subjetivos para evaluar la calidad de experiencia de vídeo. En [9], se presenta un método de evaluación subjetiva de calidad de vídeo para terminales móviles. En [10], Huawei publica un estudio de la importancia del servicio de vídeo en las actuales y futuras redes móviles. Incluye además los resultados de un estudio sobre la calidad de experiencia de vídeo en alta definición en distintos terminales móviles, realizados mediante evaluación subjetiva sobre miles de sus usuarios en diferentes países. Estos trabajos analizan el servicio de videostreaming convencional.

En este artículo se presenta un estudio realizado con observadores reales para evaluar la calidad de experiencia de vídeo del servicio YouTube Live en un dispositivo móvil. De esta forma, la principal novedad respecto a trabajos anteriores es el servicio de descarga progresiva en directo.

2. Diseño del experimento

En la siguiente sección se describen los elementos necesarios para evaluar la calidad subjetiva de imagen percibida por los usuarios.

2.1. Metodología de evaluación

Del universo de métodos disponibles en las recomendaciones para la evaluación subjetiva de vídeo, se eligió el Índice por Categorías Absolutas (ACR) [7]. Con dicho índice, el rango de calificaciones es una escala numérica con cinco niveles cuyos valores son: 1, 2, 3, 4 y 5, siendo la calidad percibida mala, mediocre, regular, buena y excelente, respectivamente [7]. Como refleja la figura 1, la presentación de las secuencias se hace una tras otra en la pantalla completa del terminal móvil, intercalando un periodo de 10 segundos para registrar la calificación de cada secuencia de vídeo.

Figura 1 – Diagrama de presentación de secuencias ITU-T P.910.

2.2. Descripción de la plataforma de obtención de vídeos

En la figura 2 se muestra la plataforma con la que se descargaron los vídeos del servidor YouTube Live. Dicha plataforma consta de un teléfono inteligente conectado a Internet a través de una red inalámbrica. A partir de la aplicación de YouTube que se ejecuta en el terminal móvil, se realizaron peticiones a un servidor de videostreaming en directo a diferentes resoluciones (144p, 240p, 360p, 480p, 720p y 1080p). Estas secuencias fueron grabadas y almacenadas en el terminal móvil para su posterior procesado y visualización.

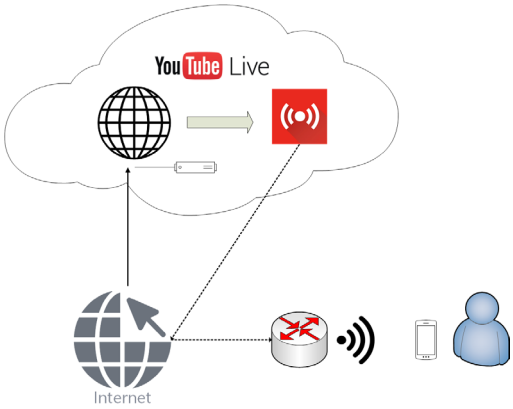


Figura 2 – Plataforma de obtención de secuencias de vídeo.

El dispositivo utilizado es un teléfono inteligente iPhone 6s, que dispone de una pantalla Retina HD 4.7" con 3D Touch y resolución de 1334 x 750 pixeles a 326 ppi [11], con la versión 11.2.1 del sistema operativo iOS. Para obtener las secuencias de vídeo, se utilizó la aplicación de YouTube móvil con versión 13.23.8 para iOS.

La grabación de las secuencias se hizo por medio de la aplicación nativa del sistema operativo iOS de grabar pantalla. Ésta permite grabar el audio y vídeo que en ese momento se proyecta en el dispositivo con la misma resolución a la que se reproduce.

2.3. Selección de secuencias

La Tabla 1 refleja las 17 secuencias de vídeo en directo seleccionadas para la prueba de calidad de experiencia. Cada vídeo dispone de seis resoluciones: 144p, 240p, 360p, 480p, 720p y 1080p, haciendo un total de 102 secuencias disponibles para el experimento.

Cada secuencia descargada tiene una duración de alrededor de 30 segundos, que se editaron posteriormente para hacer secuencias más cortas con una duración de 12 segundos. El procesado del vídeo se ha realizado con el Software Avidemux 2.7.1, que es una herramienta de uso libre para edición de vídeo.

Para cada una de las condiciones fijadas en la evaluación, se emplea contenido de vídeo diferente, evitando así dos efectos secundarios importantes asociados a la presentación de secuencias [12]:

URL de vídeo	Contenido	Cantidad de Movimiento
https://www.youtube.com/watch?v=RtU_mdL2vBM	NASA Live - Earth From Space	Bajo
https://www.youtube.com/watch?v=nmoQp7gyzIk	LiveWebcam From 4th Avenue	Bajo
https://www.youtube.com/watch?v=UmsfUJ8dF3U	Hamptons.com Shelter Island	Bajo
https://www.youtube.com/watch?v=vPbQcM4k1Ys	Venice Italy Live Cam	Bajo
https://www.youtube.com/watch?v=R4b_1rUNCUY	Decorah Eagles	Bajo
https://www.youtube.com/watch?v=qCI6y8kks8U	Canciones Infantiles	Alto
https://www.youtube.com/watch?v=dPhvx6KoVsY	African Wildlife Meditation	Alto
https://www.youtube.com/watch?v=-9SEk9-8CIU	ChillYourMind Radio	Alto
https://www.youtube.com/watch?v=90Ca1Ssnb0U	TN en Vivo	Alto
https://www.youtube.com/watch?v=Dx9juVmVR_Y	PUBG Sabadete	Alto
https://www.youtube.com/watch?v=WkXroQyvdTA	Soggy Dollar Bar Live Webcam	Bajo
https://www.youtube.com/watch?v=Ad9Eun4OxUg	Barker Ewing Whitewater	Bajo
https://www.youtube.com/watch?v=TIBX10DrtiE	Pacific Ocean Meditation	Bajo
https://www.youtube.com/watch?v=FoCVTWS0jxE	C5N EN VIVO	Alto
https://www.youtube.com/watch?v=kx0gRwi81sU	Canciones Infantiles - HeyKids	Alto
https://www.youtube.com/watch?v=aPu9mQ5FoWE	IAAF World Under 20	Alto
https://www.youtube.com/watch?v=6DJyNlbH9VU	Lithuania v Romania - FIBA	Alto

Tabla 1 – Fuentes de vídeo para la evaluación subjetiva.

- Efecto de aprendizaje: los observadores pueden reconocer objetos de una secuencia de baja calidad si los han visto anteriormente en una mejor calidad. Por lo tanto, su opinión podría ser parcial.
- Efecto de aburrimiento: incluso en sesiones cortas (menos de quince minutos), los observadores se aburren si los mismos contenidos se muestran repetidamente.

2.4. Observadores

Un total de 24 observadores formaron parte del experimento, 13 hombres y 11 mujeres, con un rango de edad comprendido entre los 21 y 59 años.

Antes de la sesión de evaluación se examinó a los observadores para determinar su agudeza visual y su visión normal de los colores mediante las pruebas de Snellen y de Ishihara, respectivamente [8].

2.5. Condiciones de evaluación

Cada condición de evaluación refleja una combinación de variables de prueba que caracteriza a una secuencia de vídeo que será calificada por los observadores. La prueba consistió en 29 secuencias de evaluación organizadas en tres bloques que se describen de la siguiente manera:

- Bloque 1: de entrenamiento y familiarización, para estabilizar las calificaciones de las secuencias que se presentan en los siguientes bloques (5 secuencias);
- Bloque 2: condiciones para evaluar el impacto de las diferentes resoluciones de vídeo en las secuencias con contenido de bajo y alto movimiento (12 secuencias).
- Bloque 3: condiciones para evaluar el impacto de las diferentes resoluciones de vídeo en las secuencias con contenido de bajo y alto movimiento. Las secuencias utilizadas en este bloque poseen una resolución diferente a la del bloque 2 (12 secuencias).

2.6. Procedimiento de evaluación

Bloque	Identificador	Resolución	Cantidad de movimiento de la secuencia
1	E1	720	Baja
	E2	144	Alta
	E3	480	Baja
	E4	1080	Alta
	E5	240	Alta
2	1	360	Baja
	2	720	Baja
	3	240	Baja
	4	1080	Baja
	5	480	Baja
	6	144	Baja
	7	360	Alta
	8	1080	Alta
	9	240	Alta
	10	480	Alta
	11	720	Alta
	12	144	Alta
3	1	144	Baja
	2	480	Baja
	3	1080	Baja
	4	360	Baja
	5	720	Baja
	6	240	Baja
	7	720	Alta
	8	240	Alta
	9	480	Alta
	10	144	Alta
	11	1080	Alta
	12	360	Alta

Tabla 2 – Clasificación de las condiciones de evaluación.

Los bloques y las secuencias se presentan en el mismo orden para todos los observadores, en la Tabla 2 se observan las diferentes condiciones de evaluación. Las pruebas se realizaron en un laboratorio de la ETSI de Telecomunicación de la Universidad de Málaga.

La distribución de tiempo de la prueba se hizo de la siguiente manera:

- 5 secuencias (12 segundos) de entrenamiento (E1, E2, E3, E4 y E5);
- 12 secuencias (12 segundos) de bloque 2;
- 12 secuencias (12 segundos) de bloque 3;
- total, 29 secuencias;
- tiempo de votación de secuencia = 10 segundos.
- para un total de 10 minutos y 38 segundos por observador.

3. Análisis de resultados

El principal indicador estudiado son los valores de MOS registrados por los observadores. También se han calculado otros indicadores (desviación típica, intervalo de confianza, proporción de votos buenos – GOB, proporción de votos aceptables – FAIR, proporción de votos malos – POW) como lo establece [7]. Además del valor promedio de MOS, se muestran los intervalos de confianza del 95%.

Para verificar la robustez de los resultados, se ha realizado un análisis de varianza (ANOVA).

3.1. Resultados

En la figura 3 se muestran las calificaciones obtenidas en función de las diferentes resoluciones de vídeo. Las barras ofrecen información del porcentaje de los votos buenos (%GOB), aceptables (%FAIR) y malos (%POW). Los valores promedio de MOS se aprecian con los marcadores en forma de cuadrado, mientras que las barras de error representan el intervalo de confianza del 95%.

La tendencia general muestra que conforme aumenta la resolución de las secuencias aumenta el porcentaje de votos buenos y disminuye el porcentaje de votos malos. En la gráfica también se observa esta tendencia para los valores de MOS. Con más detalle, se aprecia que este comportamiento se mantiene hasta la resolución de 720p.

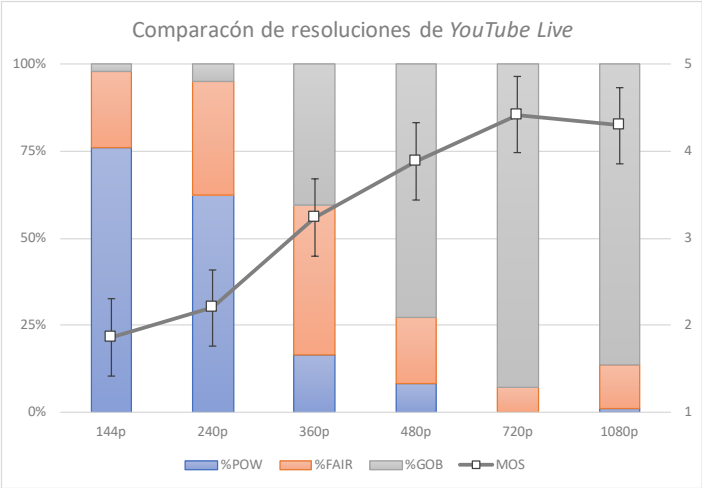


Figura 3 – Valores MOS según resolución de las secuencias

Inesperadamente, la mayor puntuación de MOS (4.42) es para la resolución de 720p. Por el contrario, la máxima resolución, 1080p, obtiene una calificación menor (4.29). Esto podría deberse a una peor calidad de imagen en 1080p si se hubiera utilizado un mayor factor de escala en el proceso de compresión de ese nivel de resolución. Para confirmar esta hipótesis, se debería verificar si la PSNR de 1080p es peor que la de 720p.

Otro análisis realizado desglosa los valores de MOS para secuencias de vídeo de alto y bajo movimiento en relación con la resolución (figura 4). En este caso, la tendencia que, en la mayoría de los casos, las secuencias de bajo movimiento tienen una mejor calificación que las secuencias de alto movimiento. De nuevo aparece la excepción en la resolución de 1080p, donde las secuencias de alto movimiento tienen una mejor calificación (4.56) que las secuencias de bajo movimiento (4.02).

La combinación con menor puntuación (1.77) es la resolución 144p con secuencias con contenido de alto movimiento y la combinación con mayor puntuación (4.56) es la resolución de 1080p con secuencias de alto movimiento.

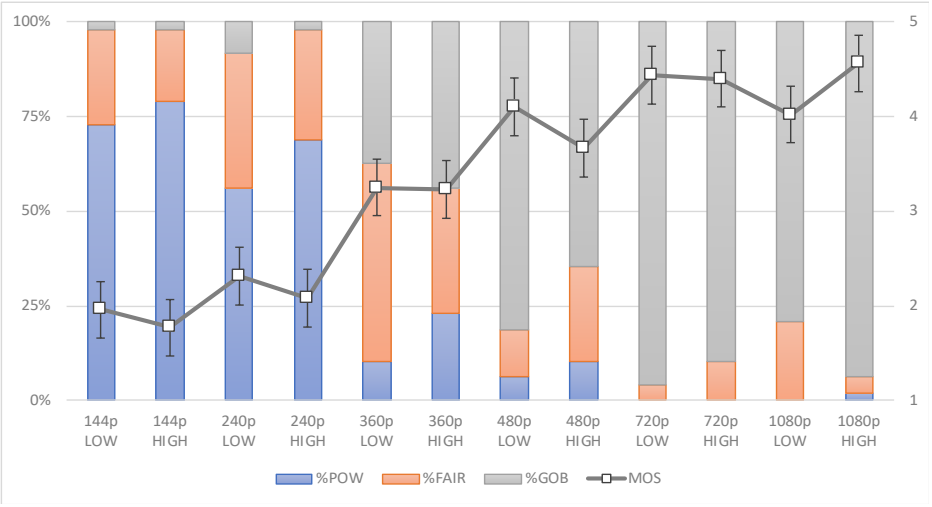


Figura 4 – Comparación de valores MOS según su movimiento y resolución

En la figura 5 se agrupan todas las resoluciones en función de la cantidad de movimiento. El valor de MOS para secuencias con bajo movimiento es de 3.35 mientras que para las de alto movimiento es de 3.28. De estos valores globales, se puede concluir que, en términos generales, no hay una diferencia significativa en la apreciación de la calidad de experiencia en función del contenido de las secuencias.

3.2. Validación de resultados

Para comprobar la significancia estadística de los valores de MOS, se ha utilizado la técnica de análisis de varianzas [13]. Para ello, se ha utilizado el Software IBM SPSS Statistics 23. El análisis realizado es de un sentido (de un factor) debido a que se tiene una variable dependiente que son los valores MOS obtenidos y la variable independiente es la resolución a la que se visualizan las secuencias de vídeo.

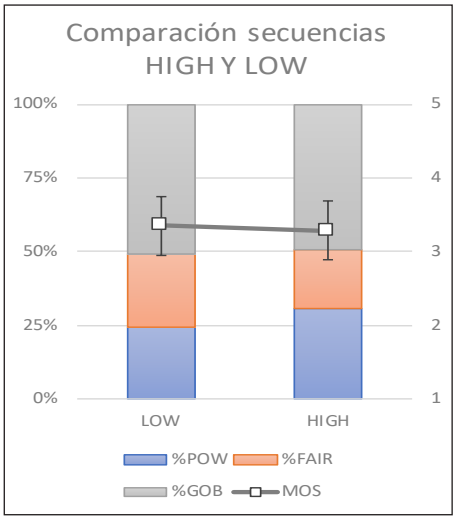


Figura 5 – Comparación de valores MOS globales de secuencias de bajo y alto movimiento

Realizada la prueba ANOVA, se obtiene los valores p, con el que se puede aceptar o rechazar la hipótesis de igualdad de medias. Cuando tenemos un valor p menor a 0.05, se rechaza la hipótesis de igualdad de medias y se puede concluir que las medias de tratamiento no son iguales, es decir, que la variable depende de los parámetros a los que se ha sometido la prueba.

En la Tabla 3, se presenta un resumen de los resultados obtenidos de la prueba ANOVA. La tabla contiene los principales resultados del análisis: suma de cuadrados, grados de libertad, media cuadrática, proporción F y significancia (p-valores).

MOS de todas las secuencias.					
	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Entre grupos	561,618	5	112,324	180,414	,000
Dentro de grupos	354,875	570	,623		
Total	916,493	575			
MOS de secuencias de bajo movimiento					
	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Entre grupos	250,778	5	50,156	85,981	,000
Dentro de grupos	164,500	282	,583		
Total	415,278	287			

MOS de secuencias de alto movimiento					
	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Entre grupos	324,069	5	64,814	103,506	,000
Dentro de grupos	176,583	282	,626		
Total	500,653	287			

Tabla 3 – Tabla ANOVA de secuencias

Las significancias calculadas (p-valores) para los valores MOS de todas las secuencias en relación con las diferentes resoluciones permite rechazar la hipótesis de igualdad de medias debido a que el valor p es menor a 0.05 (p = 0.000). Así, se puede concluir que los valores de MOS percibidos por los observadores se deben a los cambios en la resolución de las secuencias de vídeo.

4. Conclusiones

En este artículo se ha presentado un estudio de la calidad de experiencia medida de forma subjetiva para el servicio YouTube Live. El diseño del experimento se ha realizado siguiendo la normativa ITU-T P.913 [8] sobre métodos para la evaluación subjetiva de la calidad de vídeo.

Los resultados de los experimentos muestran que a una mayor resolución le corresponde una mejor calidad de experiencia. La menor puntuación MOS pertenece a la resolución de 144p y la mayor puntuación a la resolución de 720p. Numéricamente, no existe una diferencia significativa entre los valores MOS para las secuencias con contenido de bajo y alto movimiento.

Un caso de estudio en particular es la resolución de 1080p, para la que resulta una puntuación menor que para la resolución de 720p. Una potencial causa de este comportamiento puede estar relacionada al tamaño del dispositivo. para los usuarios el cambio de resolución de 720p a 1080p no supuso una gran diferencia para los observadores.

Los resultados obtenidos son coherentes con otros estudios del mismo tipo. Los mismos permitirán desarrollar un modelo de QoE para determinar la calidad que experimenta el usuario a partir de parámetros de red [14].

Como línea futura de trabajo, se propone la ampliación del estudio para una mayor cantidad de dispositivos, con diferentes tamaños y tecnologías de pantallas (p. ej., autoestereoscópica). Además, se plantea la comparación de los resultados subjetivos con medidas obtenidas por métodos objetivos (p. ej., basados en señal, paramétricos, etc.).

Referencias

Ericsson, A. (2018). Ericsson Mobility Report June 2018, pp. 3-13. Recuperado de: https://www.ericsson.com/assets/local/mobility-report/documents/2018/ericsson-mobility-report-june-2018.pdf (Accedido el 09/09/2018).

Davidson, J. Liebald, B. Liu, J. Nandy, P. Vleet, T.V. Gargi, U. Gupta, S. He, Y. Lambert, M.I. Livingston, B. and Sampath, D. (2010). The YouTube Video Recommendation System, *RecSys*.

VV.AA. (2016). La Sociedad de la Información en España 2016. Madrid, Fundación Telefóncia – Editorial Ariel.

ITU-T Report, (International Telecommunication Union). (2007). Definition of Quality of Experience (QoE), Liaison Statement, Ref.: TD 109rev2 (PLEN/12).

ITU-T P.800 (International Telecommunication Union). (1996). Methods for subjective determination of transmission quality.

ITU-R BT.513 (International Telecommunication Union). (2012). Methodology for the subjective assessment of the quality of television pictures.

ITU-T P.910 (International Telecommunication Union). (2008). Subjective video quality assessment methods for multimedia applications.

ITU-T P.9113 (International Telecommunication Union). (2016). Methods for the subjective assessment of video quality, audio quality and audiovisual quality of Internet video and distribution quality television in any environment.

Moorthy, A. K. Choi, L. K. Bovik, A. C. and De Veciana, G. (2012). Video Quality Assessment on Mobile Devices: Subjective, Behavioral and Objective Studies, in IEEE Journal of Selected Topics in Signal Processing, vol. 6, no. 6, pp. 652-671, Oct. 2012.

Huawei Technologies Co. Ltd. (2015). Mobile Video Service Performance Study. Shenzhen518129, P. R. China. Recuperado de: http://www.ctiforum.com/uploadfile/2015/0701/20150701091255294.pdf (Accedido el 09/09/2018).

Apple, "Especificaciones iPhone 6s", Apple Inc., Recuperado de: https://www.apple.com/la/iphone-6s/specs/ (Accedido el 07/11/2018).

Moreno-Roldán, J.M. (2015). Parametric Quality Assessment in Underwater Video. In Proceedings of the 2nd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 11–13 Mar. 2015; pp. 1270–1274.

Navidi, W. (2006). *Estadística para ingenieros y científicos*. México DF, McGraw-Hill/Interamericana Editores.

Jiménez, L. Solera, M. Toril, M. Oliver, P. (2017). Evaluación de la calidad de experiencia de YouTube Live en redes inalámbricas, en *XII Jornadas de Ingeniería Telemática, Sep*. Valencia – España

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 11/10/2018
Aceitação/Acceptance: 01/12/2018

Motivaciones hedónicas para el uso de sitios de redes sociales por adultos mayores: un estudio exploratorio en Chile

Patricio Ramírez-Correa[1], Giselle Painén-Aravena[1], Jorge Alfaro-Peréz[1], Ari Melo Mariano[2], Felipe Machorro-Ramos[3]

patricio.ramirez@ucn.cl, giselle.painen@alumnos.ucn.cl, jalfaro@ucn.cl, arimariano@unb.br, felipe.machorro@udlap.mx

[1] Universidad Católica del Norte, Larrondo 1281, 1781421, Coquimbo, Chile.

[2] Universidad de Brasilia, Campus Darcy Ribeiro, 70904970, Brasilia, Brasil.

[3] Universidad de las Américas Puebla, Ex Hacienda Sta. Catarina Mártir S/N, 72810, San Andrés Cholula, México.

Pages: 563–570

Resumen: El objetivo de este estudio es examinar las motivaciones hedónicas para el uso de redes sociales por adultos mayores en Chile. Con base en datos secundarios sobre acceso y uso de Internet en Chile, se realizan pruebas estadísticas para comprender este fenómeno. Los resultados del examen indican que aunque un 39,5% de chilenos adultos mayores tienen motivaciones hedónicas para el uso de las redes sociales, esta proporción disminuye significativamente a lo largo de los años. Por otro lado, no existen diferencias asociadas con el género ni en la proporción de uso ni en su intensidad. Finalmente, los resultados revelan que a lo largo de los años se mantiene constante la proporción de adultos mayores que utilizan de forma más intensa las redes sociales con motivación hedónica.

Palabras-clave: Motivaciones Hedónicas; Redes Sociales; Adultos Mayores; Chile.

Hedonic motivations for social networking sites use by elderly: an exploratory study in Chile

Abstract: The purpose of this study is to examine the hedonic motivations for social networks sites use by older adults in Chile. Based on secondary data on Internet access and use in Chile, statistical tests are carried out to understand this phenomenon. The results of the examination indicate that although a general 39.5% of the elderly persons have hedonic motivations for social networks sites use; this proportion decreases significantly over the years. On the other hand, there are no differences associated with gender neither in the proportion of use nor in its intensity. Finally, the results indicate that remains constant over the years the proportion of older adults who use social networks with hedonic motivation more intensely.

Keywords*:* Hedonic Motivations; SNS; Elderly; Chile.

1. Introducción

En forma creciente las tecnologías de información cumplen un papel hedónico en la sociedad actual (Turel, Serenko, & Bontis, 2010; Xu, Ryan, Prybutok, & Wen, 2012), y en consecuencia, esta dimensión de uso se ha convertido en foco de recientes estudios académicos (Pal & Triyason, 2018; Ramírez-Correa, Mello, & Mariano, 2018; Ramírez-Correa, Rondán-Cataluña, & Arenas-Gaitán, 2018). La literatura define a los sistemas de información hedónicos como aquellos que tienen como propósito entregar un valor auto realizable a sus usuarios finales, y en su forma más pura, la interacción con estos sistemas está diseñada para ser un fin en sí misma (Van der Heijden, 2004). En general, el valor de estos sistemas es una función del grado en que el usuario experimenta diversión al usarlo, y por tanto, si el usuario experimenta mayor diversión utilizándole, tendrá mayor intención de continuar su uso en el futuro. Las redes sociales (SRS, por sus siglas en inglés), definidas como servicios basados en web que permiten a las personas construir un perfil de usuario público o semi-público dentro de un sistema delimitado, articular una lista de otros usuarios con los que comparten una conexión, ver y recorrer su lista de conexiones y las hechas por otros usuarios dentro del sistema (Boyd & Ellison, 2007), se pueden considerar un ejemplo de estos sistemas de información hedónicos (Peter & Sherman, 2006; Xu et al., 2012).

Por otra parte, existe un importante interés en el envejecimiento de la población mundial. Según Naciones Unidas, la proporción de la población mundial de más de 60 años aumenta en forma acelerada, y se espera que se llegue a duplicar para el año 2050. De hecho, esta proporción pasó de un 8,5% en 1980 a un 12,7% en 2017, y se proyecta a un 21,3% para 2050 (United Nations, Department of Economic and Social Affairs, 2017). En particular en Chile, y de acuerdo al CENSO 2017, el 16,2% de los chilenos sobrepasan los 60 años. Esta realidad se asocia a una fuerte transformación de la pirámide poblacional del país, que ha pasado de una forma triangular en 1975 a una forma rectangular en nuestros días (ECLAC Demographic Observatory, 2016). De esa forma, el índice de envejecimiento chileno del 2015 señala que existía una relación de 86 personas mayores de 60 años por cada 100 menores de 15 años (CASEN, 2017), valor muy superior al promedio de América Latina (39,6), y cercano a los demás países de la Organización de Estados para la Cooperación y el Desarrollo (OECD, 2018). Adicionalmente, las estadísticas indican que los adultos mayores están viviendo cada vez más solos y sin la compañía de sus familiares cercanos (United Nations, Department of Economic and Social Affairs, 2017). Esto provoca problemas de depresión y, por consiguiente, un detrimento en la salud y bienestar general de este grupo etario. Por este motivo, y por tratarse de un fenómeno global, se han desarrollado una serie de tecnologías digitales con el fin de obtener beneficios para la salud. En particular en el caso de las SRS, la literatura médica enfatiza su efectividad en lograr mejoras en la calidad de vida de los adultos mayores (Khosravi & Ghapanchi, 2015).

Lamentablemente, y a pesar de que se ha estudiado el fenómeno de adopción de diversas tecnologías de información por parte de los adultos mayores, aun predominan suposiciones estereotipadas asociadas a la vejez (Beneito-Montagut, Cassián-Yde, & Begueria, 2018), por ejemplos, que las SRS son utilizadas solo por jóvenes, o que las motivaciones de su posible uso por adultos mayores son solo utilitarias. En este sentido,

y reconociendo los beneficios del uso de las SRS por los adultos mayores (Baker et al., 2018; Chopik, 2016; Zhou, 2018), es clara la necesidad de explorar el fenómeno en cuestión.

En este contexto, este artículo propone examinar las motivaciones hedónicas para el uso de SRS por adultos mayores en Chile. Esta primera exploración se realizará con datos secundarios sobre acceso y uso de Internet en Chile.

El texto restante del artículo se ha estructurado de la siguiente forma. Primero se presenta una breve revisión de antecedentes en la literatura, para luego entregar una reseña de la metodología de análisis. Seguido, se presentan los resultados y se realiza una discusión de ellos. Finalmente, se entregan las conclusiones del estudio.

2. Antecedentes

En paralelo con la extensión del uso de tecnologías de información a diversos ámbitos de nuestra vida, el estudio de los sistemas de información hedónicos se ha establecido como espacio de estudio relevante (Ramírez-Correa, Mello, et al., 2018). Hace más de una década y en su trabajo seminal sobre la aceptación de este tipo de sistemas, Hans van der Heijden apuntaba a la necesidad de entender y explicar los motivos de su utilización por parte de los consumidores en todo el mundo (Van der Heijden, 2004). Particularmente, Van der Heijden (2004) estudió la aceptación de los sistemas de información hedónicos en base al modelo de aceptación de tecnologías (TAM, por sus siglas en inglés) de Fred Davis (Davis, 1989). Los resultados de van der Heijden indican que tanto la facilidad de uso como la diversión percibida son los antecedentes más importantes para explicar la intención de usar este tipo de sistemas, cuando el usuario ya posee experiencia (van der Heijden, 2004).

Posteriormente, y también con base en TAM, Rosen y Sherman orientaron su trabajo a SRS (Peter & Sherman, 2006). Estos autores buscaron explicar el uso de SRS, considerando que este uso entorpece en la productividad personal, lo cual es inconsistente con uno de los antecedentes que explican la intención de uso de tecnologías según el TAM. Los resultados de estos autores señalan que tanto la diversión percibida como la facilidad de uso son los antecedentes del uso de SRS. En misma línea, Lin y Lu (2011) estudiaron la intención de uso continuo de SRS, sus hallazgos sugieren que conjuntamente a las opiniones de los amigos, la diversión es el factor más importante para explicar la intención del uso continuo de SRS.

Por su parte Xu y colegas, luego de un estudio empírico de múltiples factores antecedentes del uso de SRS por parte de estudiantes universitarios, encontraron como predictores de esta utilización tanto a las gratificaciones utilitarias (racionales y orientadas a objetivos) de acceso inmediato y coordinación, como a las gratificaciones hedónicas (asociadas al placer) de afecto, ocio y presencia social (Xu et al., 2012). Estos hallazgos han sido soportados por otros estudios (Ernst, Pfeiffer, & Rothlauf, 2015), y en particular, en el caso de los adultos mayores se ha reportado la existencia de motivaciones sociales, utilitarias y hedónicas para el uso de SRS (Gu, Suh, & Liu, 2015).

3. Metodología

Este estudio se basa en datos recogidos en la IX Encuesta de Acceso y Usos de Internet de Chile (SUBTEL, 2017), esta encuesta posee un diseño muestral probabilístico, estratificado por región y zona urbano/rural, y con selección polietápica. El cuestionario de la encuesta fue aplicado a 3.600 hogares entre septiembre y octubre de 2017, abarcando las 15 regiones de Chile.

Los datos de la encuesta fueron procesados para obtener a los adultos mayores[1] que reportaron usar SRS con motivaciones hedónicas. En total fue de 595 usuarios, 276 hombres (46%) y 319 mujeres (54%).

Para el análisis se utilizó las respuestas múltiples a la pregunta ¿Cuál o cuáles de estas redes sociales usa cuando quiere entretenerse?, se consideró el uso de SRS si el adulto mayor indicaba al menos una red social como respuesta. Para efectos de aproximar el uso intensivo de SRS se utilizó la suma de las SRS que reporto el usuario, este valor se clasifico en dos segmentos: bajo y sobre el percentil 50.

Por otra parte, y para efectos de análisis, las edades de los usuarios se clasificaron en terciles.

4. Resultados

La Tabla 1 muestra el porcentaje de uso por motivos hedónicos de distintas SRS, clasificadas en terciles de edad y ordenadas de mayor a menor según el uso global.

Red Social	Edad			Total (595)
	60-64 (195)	65-70 (217)	71-90 (183)	
WhatsApp	40,29%	19,23%	17,31%	24,74%
Facebook	37,32%	21,91%	10,91%	22,68%
YouTube	29,14%	16,04%	12,27%	18,76%
Google	14,71%	14,21%	7,65%	12,26%
Instagram	1,56%	0,93%	2,81%	1,71%
Twitter	2,09%	0,46%	1,10%	1,19%
Pinterest	2,09%	0,00%	0,55%	0,85%
LinkedIn	0,52%	0,00%	0,00%	0,17%

Tabla 1 – Uso de SNS por motivos hedónicos

Claramente las SRS WhatsApp, Facebook, YouTube y Google sobresalen por sobre las restantes, tanto en forma global como por terciles de edad. Como se puede observar, WhatsApp es usado por más del 40% de los adultos entre 60 y 64 años, pero a partir de los 65 años baja abruptamente su uso a menos de la mitad de esta cifra. También sucede este cambio, pero menos notoriamente, en Facebook y YouTube. Por otra parte,

[1] Acorde a la Ley 19.828 de Chile un adulto mayor es una persona de 60 años y más.

si bien Instagram y Twitter tienen globalmente un bajo porcentaje de uso por los adultos mayores, comparten su nivel más bajo de uso entre los 65 y 70 años.

Se realizaron pruebas χ^2 para estudiar si existen diferencias asociadas a la edad y al género en el uso por motivos hedónicos de SRS en los adultos mayores. La Tabla 2 muestra el resultado de estas pruebas. Como se puede ver en este resultado, tanto para hombres y mujeres por separado como en forma agregada, el uso de SRS por motivos hedónicos disminuye significativamente con el paso de los años (ver últimas dos columnas de la Tabla 2). Por otra parte, no existen diferencias estadísticamente significativas entre hombres y mujeres en el uso de SRS por motivos hedónicos, ni en forma totalizada ni en ninguno de los terciles analizados (ver últimas dos filas de la Tabla 2).

Género	**Edad**			**Total**	χ^2	**Sig.**
	60-64 (195)	**65-70 (217)**	**71-90 (183)**			
Hombres (276)	57,14%	32,41%	27,38%	38,41%	18,43	0,00
Mujeres (319)	53,15%	41,28%	25,25%	40,44%	16,96	0,00
Total (595)	54,87%	36,87%	26,23%	39,50%	33,40	0,00
χ^2	0,31	1,84	0,11	0,26		
Sig.	0,58	0,18	0,74	0,61		

Tabla 2 – Análisis de diferencias en uso de SNS por motivos hedónicos

Finalmente, se realizaron pruebas χ^2 para examinar diferencias asociadas a la edad y al género en la intensidad de uso por motivos hedónicos de SRS en los adultos mayores. La Tabla 3 muestra el resultado de este examen. Como se aprecia en este resultado, tanto para hombres y mujeres por separado como en forma agregada, la intensidad de uso de SRS por motivos hedónicos no cambia significativamente con el paso de los años (ver últimas dos columnas de la Tabla 3). Por otra parte, y al igual que en el caso del uso, no existen diferencias estadísticamente significativas entre hombres y mujeres en la intensidad de uso de SRS por motivos hedónicos, ni en forma totalizada ni en ninguno de los terciles analizados (ver últimas dos filas de la Tabla 3).

Género	**Edad**			**Total**	χ^2	**Sig.**
	60-64 (195)	**65-70 (217)**	**71-90 (183)**			
Hombres (106)	43,75%	42,86%	39,13%	42,45%	0,14	0,93
Mujeres (129)	45,76%	48,89%	44,00%	46,51%	0,18	0,91
Total (235)	44,86%	46,25%	41,67%	44,68%	0,26	0,88
χ2	0,04	0,29	0,12	0,39		
Sig.	0,84	0,59	0,73	0,53		

Tabla 3 – Análisis de diferencias en intensidad de uso de SNS por motivos hedónicos

A partir de estos resultados podemos esbozar tres ideas relevantes. Primero, el efecto de disminución del uso hedónico de una SRS a medida que avanza la edad del adulto mayor

podría estar asociado a la edad de jubilación[2], pues dado el retiro del trabajo la red de contactos laborales disminuye, y con ello el uso hedonista asociado a consumir y producir información de valor para los integrantes de la red. Es decir, con la jubilación va desarrollando un proceso inverso a la socialización iniciada en la niñez. En base a la teoría de roles de Rosow, a este fenómeno se le ha llamado la desaparición social del anciano (Rosow, 1967). Por tanto, este es un llamado a las empresas para establecer sistemas que vinculación con trabajadores retirados y así aumentar potencialmente el bienestar de este grupo etario.

Segundo, no existen diferencias en la intensidad de uso SRS por motivos hedónicos a medida que pasan los años, es decir, un poco más de la mitad de los adultos mayores usuario de SRS por motivos hedónicos utilizan esta tecnología en forma poco intensiva, y por tanto, pudiera existir una brecha importante a ser cubierta por nuevos SRS orientados específicamente a este segmento. Y como lo destaca la literatura, es probable que muchas personas mayores no adopten estos servicios simplemente porque no están diseñados para sus necesidades (Berenguer et al., 2017).

Y tercero, si bien en muchos estudios del uso de tecnología de información las diferencias de género han sido un punto importante de análisis (Ramirez-Correa, Arenas-Gaitan, & Rondan-Cataluna, 2015; Venkatesh & Morris, 2000), nuestros resultados señalan que de existir estas diferencias no están en el uso por motivos hedónicos de SRS.

4. Conclusiones

Este estudio se exploró el uso por motivos hedónicos de SRS por parte de adultos mayores en Chile. La principal conclusión señala que existe una importante proporción de uso de SRS con estos fines, pero esta disminuye significativamente a lo largo de los años. Creemos que futuros estudios sobre este fenómeno deberían estudiar a través de modelos de aceptación de tecnologías tanto las creencias como las actitudes que predicen el uso de SRS como sistemas hedónicos.

Referencias

Baker, S., Warburton, J., Waycott, J., Batchelor, F., Hoang, T., Dow, B., … Vetere, F. (2018). Combatting social isolation and increasing social participation of older adults through the use of technology: A systematic review of existing evidence. *Australasian Journal on Ageing*. https://doi.org/10.1111/ajag.12572

Beneito-Montagut, R., Cassián-Yde, N., & Begueria, A. (2018). What do we know about the relationship between internet-mediated interaction and social isolation and loneliness in later life? *Quality in Ageing and Older Adults*, *19*(1), 14–30. https://doi.org/10.1108/QAOA-03-2017-0008

Berenguer, A., Goncalves, J., Hosio, S., Ferreira, D., Anagnostopoulos, T., & Kostakos, V. (2017). Are Smartphones Ubiquitous?: An in-depth survey of smartphone adoption by seniors. *IEEE Consumer Electronics Magazine*, *6*(1), 104–110. https://doi.org/10.1109/MCE.2016.2614524

[2] En Chile, desde los 60 años para las mujeres y desde los 65 años para los hombres.

Boyd, D. M., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, *13*(1). https://doi.org/11

CASEN. (2017). *Informe Desarrollo Social 2017*.

Chopik, W. J. (2016). The Benefits of Social Technology Use Among Older Adults Are Mediated by Reduced Loneliness. *Cyberpsychology, Behavior, and Social Networking*, *19*(9), 551–556. https://doi.org/10.1089/cyber.2016.0151

ECLAC Demographic Observatory. (2016). *Population projections*.

Ernst, C. P. H., Pfeiffer, J., & Rothlauf, F. (2015). Hedonic and utilitarian motivations of social network site usage. In *Factors Driving Social Network Site Usage* (pp. 11–28). https://doi.org/10.1007/978-3-658-09918-3_2

Gu, Y., Suh, A., & Liu, L. (2015). Motivations and obstacles for seniors to adopt social network services. In *Pacific Asia Conference on Information Systems, PACIS 2015 - Proceedings*.

Khosravi, P., & Ghapanchi, A. H. (2015). Investigating the effectiveness of technologies applied to assist seniors: A systematic literature review. *International Journal of Medical Informatics*, *85*(1), 17–26. https://doi.org/10.1016/j.ijmedinf.2015.05.014

Lin, K. Y., & Lu, H. P. (2011). Why people use social networking sites: An empirical study integrating network externalities and motivation theory. *Computers in Human Behavior*, *27*(3), 1152–1161. https://doi.org/10.1016/j.chb.2010.12.009

OECD. (2018). Elderly population (indicator). https://doi.org/10.1787/8d805ea1-en

Pal, D., & Triyason, T. (2018). User Intention towards a Music Streaming Service: A Thailand Case Study. *KnE Social Sciences*, *3*(1), 1–16.

Peter, A. R., & Sherman, P. (2006). Hedonic information systems: Acceptance of social networking websites. *Association for Information Systems - 12th Americas Conference On Information Systems, AMCIS 2006*, *2*, 1208–1213.

Ramirez-Correa, P. E., Arenas-Gaitan, J., & Rondan-Cataluna, F. J. (2015). Gender and Acceptance of E-Learning: A Multi-Group Analysis Based on a Structural Equation Model among College Students in Chile and Spain. *Plos One*, *10*(10). https://doi.org/10.1371/journal.pone.0140460

Ramírez-Correa, P. E., Mello, T. M., & Mariano, A. M. (2018). A aceitação da Netflix: um estudo utilizando equações estruturais. *Revista Tecnologia e Sociedade*, *14*(30).

Ramírez-Correa, P. E., Rondán-Cataluña, F. J., & Arenas-Gaitán, J. (2018). A Posteriori Segmentation of Personal Profiles of Online Video Games' Players. *Games and Culture*. https://doi.org/10.1177/1555412018766786

Rosow, I. (1967). *Social integration of the aged*. Free Press.

SUBTEL. (2017). *IX Encuesta Accesos y Usos de Internet*.

Turel, O., Serenko, A., & Bontis, N. (2010). User acceptance of hedonic digital artifacts: A theory of consumption values perspective. *Information and Management*, *47*(1), 53–59. https://doi.org/10.1016/j.im.2009.10.002

United Nations, Department of Economic and Social Affairs, P. D. (2017). World Population Ageing 2017. *United Nations*, 124. https://doi.org/ST/ESA/SER.A/348

Van der Heijden, H. (2004). User acceptance of hedonic information systems. *MIS Quarterly*, 695–704.

Venkatesh, V., & Morris, M. G. (2000). Why don't men ever stop to ask for directions? Gender, social influence, and their role in technology acceptance and usage behavior. *Mis Quarterly*, *24*(1), 115–139. https://doi.org/10.2307/3250981

Xu, C., Ryan, S., Prybutok, V., & Wen, C. (2012). It is not for fun: An examination of social network site usage. *Information and Management*, *49*(5), 210–217. https://doi.org/10.1016/j.im.2012.05.001

Zhou, J. (2018). Improving older people's life satisfaction via social networking site use: Evidence from China. *Australasian Journal on Ageing*, *37*(1), E23–E28. https://doi.org/10.1111/ajag.12499

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 15/09/2018
Aceitação/Acceptance: 03/12/2018

O Papel da Aceitação da Tecnologia nas Cidades Inteligentes: Um estudo das percepções dos usuários do Uber Brasil

Ari Melo Mariano[1], Patricio Ramírez-Correa[2], Jorge Alfaro-Pérez[2] , Giselle Painén-Aravena[2] , Felipe Machorro-Ramos[3]

arimariano@unb.br, patricio.ramirez@ucn.cl, jalfaro@ucn.cl, giselle.painen@alumnos.ucn.cl, felipe.machorro@udlap.mx

[1] Universidade de Brasília, Campus Darcy Ribeiro, 70904970, Brasília, Distrito Federal, Brasil.

[2] Universidad Católica del Norte, Larrondo 1281, 1781421, Coquimbo, Chile

[3] Universidad de las Américas Puebla. Ex Hacienda Sta. Catarina Mártir S/N. San Andrés Cholula, Puebla, México. C.P. 72810.

Pages: 571–583

Resumo: O crescimento de estudos a respeito de cidades inteligentes tem apresentado novas perspectivas para sua materialização. Embora ainda não exista um conceito consolidado, pode-se perceber que o conceito cidade inteligente em realidade é um *proxy*, formada pela aceitação e compartilhamento da tecnologia de maneira conjunta, que por sua vez está representada pela união de vários aspectos de aceitação e compartilhamento a nível individual. Embora falar de cidades inteligentes pareça um tópico muito distante para o Brasil, algumas ações já estão em funcionamento, como o Uber. Compreender as percepções a nível individual do usuário contribui em antecipar possíveis obstáculos para uma adoção de outras iniciativas das cidades inteligentes. Assim, o objetivo deste estudo é conhecer os principais discursos dos usuários de Uber no Brasil, atentando para o papel dos sistemas de informação. Para alcançar este objetivo é realizada uma análise textual com 142 registros da plataforma reclame aqui, a fim de conhecer as principais queixas dos usuários adotando como abordagem a Teoria Unificada de Aceitação e Uso da Tecnologia (UTAUT). Os resultados revelaram cinco diferentes discursos, sendo três deles diretamente ligados à UTAUT.

Palavras-chave: Cidades Inteligentes, Aceitação e Uso da Tecnologia, UTAUT, Análise Textual, Iramuteq.

The Role of Technology Acceptance in Smart Cities: A study of Uber Brazil users' perceptions.

Abstract: The growth of studies on Smart Cities has presented new perspectives for its materialization. Although there is still no consolidated concept, the concept of intelligent city is a proxy, formed by the acceptance and sharing of technology in a joint way, which in turn is represented by the union of several aspects of acceptance and sharing to individual level. Although speaking of smart cities seems

a very distant topic for Brazil, some actions are already in operation, such as the Uber. Understanding individual perceptions of the user contributes to anticipate potential obstacles to an adoption of other Smart Cities initiatives. Thus, the objective of this study is to know the main discourses of Uber users in Brazil, paying attention to the role of information systems. To achieve this goal, a textual analysis is carried out with 142 records of the complaint platform here, to know the main complaints of the users adopting as approach the Unified Theory of Acceptance and Use of Technology (UTAUT). The results revealed five different discourses, three of them directly related to UTAUT.

***Keywords**:* Smart Cities, Acceptance and Use of Technology, UTAUT, Textual Analysis, Iramuteq.

1. Introdução

O conceito de uma cidade inteligente é desafiador, pois o próprio conceito de "cidade" isolado é complexo, pois envolve muitas icognitas, dificultando a capacidade de estimar seu futuro (Cosgrave, 2013). Segundo Norman (2018), à medida que o mundo avança para uma população global de quase 10 bilhões até 2050, aumenta a pressão de como se vai lidar com os desafios das áreas urbanizadas como saúde urbana, mudança climática, justiça social e governança urbana. Em uma tentativa de contribuir com respostas a estes problemas, as pesquisas sobre cidades inteligentes amadurecem e novas abordagens interdisciplinares aparecem progressivamente com uma linha marcada na aplicação e usabilidade das Tecnologias de Informação e Comunicação (TIC) (Visvizi & Lytras, 2018).

Segundo Capdevila & Zarlenga (2015), embora não exista uma definição única do que são Cidades Inteligentes, geralmente se aceita como sendo políticas urbanas "inteligentes" por meio de iniciativas de governos locais que usam tecnologias de informação e comunicação para aumentar a qualidade de vida de seus habitantes, contribuindo para um desenvolvimento sustentável. Porém, apesar de todo investimento em tecnologia, as cidadas são formadas por indivíduos com diferentes níveis de interação e aceitação da tecnologia. Atentar para a usabilidade assim como para aceitação da tecnologia é uma pauta importante, pois todo processo de difusão começa a nível individual. Rolfstam, Phillips & Bakker (2011), explicam que é por meio desta percepção individual crescente que ocorre a difusão, ou seja, se necessita inicialmente compreender as motivações do individuo e assim pensar no conceito de grupo. Rogers (1995), explica que os indivíduos são vistos como possuidores de diferentes graus de disposição para adotar inovações e, portanto, é geralmente observado que a porção da população que adota uma inovação é aproximadamente normalmente distribuída ao longo do tempo, sendo necessário quebrar esta distribuição normal em segmentos, respeitando suas características no processo de adoção de uma nova tecnologia. Este processo de conhecer os consumidores segundo suas características é explicado pelo avanço rápido da tecnologia, principalmente as de informação, que gerou diferentes reações na interação do ser humano com a tecnologia. Guelzim, Obaidat & Sadoun (2016), explicam que muitos usuários passaram de consumidor de informações para principais contribuintes e produtores de dados, tornando o processo de conhecer a aceitação a nível individual necessário.

Assim, em um contexto tão cambiante e múltiplo, a adoção de cidades inteligentes torna-se um desafio para diferentes nações, seus distintos níveis de tecnologia e seus consumidores

finais com distintos níveis de familiaridade com a mesma. No Brasil, uma das representações mais significativas das cidades inteligentes é o aplicativo de mobilidade Uber. O aplicativo Uber foi escolhido como centro deste estudo pois é uma iniciativa que envolve economia compartilhada, sistemas de informação, sustentabilidade, ao diminuir o número de veículos e suas emissões, assim como a necessidade de ter a propriedade de um veículo.

Os efeitos deste primeiro contato com futuro das Cidades Inteligentes (*Smart Cities*) geraram uma série de dúvidas por parte dos indivíduos, grupos e governos, sobre aceitação, uso, regulação e impactos desta nova economia (Uber, 2018). Defossez (2017), explica em seu estudo os principais desafios da regulação de um aplicativo que surge com um conceito de economia compartilhada no Brasil e Europa, ocasionando no leitor a dúvida sobre a preparação dos países para a chegada das *Smart Cities*. Deste modo esta pesquisa busca responder quais são os principais desafios do Uber na expansão do aplicativo no Brasil? Conhecer estes desafios é importante para saber as debilidades enfrentadas no país para adoção de novas práticas apoiada em sistemas de informação (SI), contribuindo com uma visão de como a sociedade brasileira está preparada para migrar para uma sociedade baseada em SI, para possuir cidades inteligentes. Assim, o objetivo deste estudo é conhecer os principais discursos dos usuários de Uber no Brasil, atentando para o papel dos SI.

2. Cidades Inteligentes – *Smart Cities*

Segundo Silva, Khan & Han (2018), muito mais do que um conceito único, as cidades inteligentes são a consolidação de múltiplos atributos, como sustentabilidade, qualidade de vida, urbanização, e inteligência, definida como o desejo de melhorar as referências sociais, ambientais e econômicas da cidade e seus habitantes. Tuominen (2018) , explica que ainda existe confusão sobre o uso do termo *Smart Cities,* porém um fator comum liga todos eles: a tecnologia. A Figura 1 explica a evolução do conceito e nomeclatura até a atual legenda *Smart Cities*.

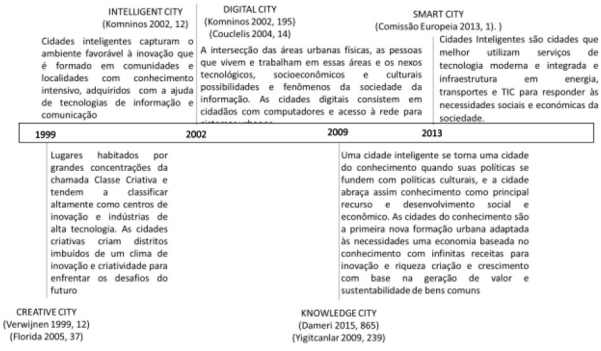


Figura 1 – Evolução do conceito da *Smart Cities*. Fonte: Adaptado de Tuominen (2018)

Embora os conceitos foram se desenvolvendo ao logo do tempo, em realidade as abordagens coexistem, sendo o termo *Smart Cities* atualmente mais conhecido. Pode-

se perceber que todas as colaborações na conceitualização de cidades inteligentes compartilham a necessidade de tecnologia e informação. Tuominen (2018), explica que o processo de desenvovimento das cidades para seu papel *smart*, depende não apenas do compromisso dos investimentos em tecnologia, infraestrutura, mas sobretudo do envolvimento do cidadão. Assim, o estudo da aceitação da tecnologia a nível individual é pré-requisito para aplicar-se o conceito de cidades inteligentes.

2.1. Aceitação em Sistemas de Informação

Embora o contexto mundial tenha se alterado substancialmente, muitas decisões continuam sendo tomadas com base em paradigmas já superados, e mesmo que existam empresas investindo fortemente em tecnologia, o capital despendido por estas não garantem o sucesso por si só (Graeml, 1998). Segundo Oliveira, Baptista & Campos (2016), no universo das tecnologias digitais, compreender as razões que levam indivíduos a adotar ou rejeitar a Tecnologia de Informação se tornou uma necessidade, pois é a partir dos indivíduos que a tecnologia se difunde. Nos últimos anos, muitos modelos têm sido criados e estudados pela optica de SI, como a Teoria da Ação Racionalizada, o Modelo de Aceitação da Tecnologia, o Modelo Motivacional, a Teoria do Comportamento Planejado, Modelo de Utilização do PC, a Teoria da Difusão da Inovação, a Teoria Social Cognitiva, a Teoria Unificada de Aceitação e Uso da Tecnologia (UTAUT), e a Teoria Unificada de Aceitação e Uso da Tecnologia 2 (UTAUT2) (Bobsin et al.; 2009; Rondan-Cataluña et al., 2015).

A integração exponencial da tecnologia nas organizações ocasionou em uma novo modelo organizacional. Para compreender e atuar neste novo modelo nasce os SI como a combinação entre pessoas, hardware, ,software, redes de comunicação, recursos de dados e políticas e procedimentos que armazenam, restauram, transformam e disseminam informações em uma organização (O'Brien & Marakas, 2013). Neste contexto, avaliar a aceitação da tecnologia por parte dos usuários torna-se necessário para que se possa conhecer os erros e acertos na aplicação de SI, o comportamento individual e as variáveis que auxiliam na mensuraração do grau de sucesso (Mathieson & Ryan, 1994; Davis & Warshaw, 1989; Bobsin, Visentini & Rech, 2009). Porém, a implementação de SI é dispendiosa e tem uma taxa de sucesso relativamente baixa, sendo estudos nesta área importantes. Legris, Ingham & Collerette, (2003), explicam que desde os anos setenta, a pesquisa de SI contribui para uma melhor compreensão deste processo e seus resultados. As aplicações dos Modelos de Aceitação e Uso da Tecnologia têm sido amplamente usados para compreender a experiência dos usuários relacionados à diversas tecnologias. Williams (2015) faz uma revisão da literatura do modelo UTAUT e encontra 174 artigos, de 18 Universidades diferentes. Estados Unidos (45), China (19) e Taiwan (17), são os locais onde mais estudos foram feitos. No total, estes artigos analisaram 25 sistemas de comunicação, 90 de uso geral, 11 sistemas de escritório e 48 de negócios especializados.

Para avaliar a aceitação a nível individual foi escolhido o modelo ofertado pela a UTAUT, por ser um modelo consolidado na literatura, alcançando valores de explicação de 69% a 70%(Mariano & Diaz, 2017). O UTAUT está formado pelas variáveis i. Expectativa de Performance, explicado pela crença do indivíduo no aumento de seu desempenho ao adotar determinada tecnologia,ii. Expectativa de esforço, explicado pelo grau

de dificuldade de uso que encontra o indivíduo ao interagir com a tecnologia, iii. as Condições Facilitadoras, explicadas pela infraestrutura de apoio ao indivíduo para uso da tecnologia e finalmente iv. Influência Social, que explica como o indivíduo reage a influência da coletividade sobre o uso de uma tecnologia.

2.2. Aplicativo Uber

A **Uber Technologies Inc**. foi fundada em 2010 e hoje está em mais de 600 cidades em 65 países (Uber,2018). Em um contexto de Cidades Inteligentes a Uber pode ser considerada como um passo a mais no processo de desenvolvimento. A empresa Uber possui dois clientes, o usuário final e o motorista, tendo uma comunicação específica para cada um deles. Apesar de ser considerado como um modelo de negócio, ao se tratar de um SI, existem diversos desafios para seu bom funcionamento. Como os dados da Uber são privados, foi consultada as principais reclamações a respeito da Uber no Brasil nestes 4 anos de operação. No Brasil, muitos consumidores utilizam-se de a plataforma "Reclame aqui" para realizar reclamações a respeitos de serviços e produtos. Atualmente "Reclame aqui" possui 42 milhões de visualizações mês, com 56% do público masculino e 44% feminino (Reclame aqui, 2018). Buscando as reclamações a respeito do Uber, foram encontradas um total de 114298 reclamações, sendo 65713 respondidas pela empresa e 48585 não respondidas.

3. Método

Os dados foram coletados a partir de registros na plataforma "reclame aqui" (site https://www.reclameaqui.com.br/), a respeito do Uber em duas datas. 07 de setembro de 2016 e 07 de setembro de 2018. Esta data foi escolhida por se tratar de um feriado nacional referente a independência do país em que as pessoas saem para ver os desfiles e optam por usar um transporte alternativo ao seu próprio veículo, uma vez que as ruas principais mudam seu fluxo para receber as festividades de Independência. Foram escolhidos os anos de 2016 e 2018 para uma comparação. O ano de 2018 foi escolhido por ser o presente ano, favorecendo uma compreensão da situação atual. Foram encontradas 71 registros a respeito da organização em diferentes cidades do Brasil. O ano de 2016 foi escolhido pois foi o segundo ano de registros no site escolhido, pois em 2015 se tinham poucos registros, não favorecendo o critério de equilíbrio da amostra. Foram encontrados 152 registros, porém para igualar a amostra foi realizado um sorteio para compor uma amostra de 71 registros, e assim garantir o equilibrio da amostra final. Os 142 dados (71 de 2016 e 71 de 2018) foram classificados de acordo com 5 critérios: a. Cidade do reclamante, b. Estado do reclamante, c. Ano, d. Tipo de reclamação, e. Classificação da reclamação.

A classificação das reclamações foram organizadas segundo as seguintes classes: i. Expansão de Linha, referente a novos serviços prestados pela empresa, como entrega de encomendas, comida, veículo compartilhado, ii. Alcance do serviço, referente as dificuldades de acesso ao serviço fora dos grandes centros, iii. Financeiro, todas aquelas reclamações sobre tarifas indevidas e cobranças abusivas ou equivocadas, iv. SI, reclamações sobre o funcionamento do aplicativo quanto a cadastro, erros, bugs, v. Atendimento ao cliente, reclamações de mau atendimento por conta do motorista ao cliente, vi. Condições do veículo, aqueles registros sobre o uso de veículos velhos

por motorista, vii.Propaganda enganosa, referente aos descontos e parceria para valores especiais anunciados e não cumpridos e finalmente viii. Objetos perdidos, reclemação sobre a não devolução de itens perdidos.

O tratamento de dados foi dividido em três partes. A primeira parte foi o tratamento para adequar os registros ao formato de entrada dos programas utilizados. As coletas foram realizadas entre 11 e 12 de setembro de 2018. Os dados foram registrados em excel e word para serem posteriormente analisados. Foram corrigidos erros e abreviações contidas nos registros. A segunda parte foram as análises por meio dos programas informáticos utilizado os dados tratados. Os dados em word foram a entrada do programa Iramuteq 0.7. IRAMUTEQ (Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires), um programa livre baseado no ambiente estatístico do software R e na linguagem *python* (www.python.org), que possibilita diferentes processamentos e análises estatísticas de textos produzidos. O programa facilta o organziação do texto para posterior análise textual. Os 142 registros foram analisados por meio de Classificação Hierárquica Descendente, Análise Fatorial de Correspondência e de similitude e depois foram comparados os anos de 2016 e 2018, além dos três Estados com maior número de reclamações: São Paulo, Rio de Janeiro e Minas Gerais. O segundo programa utilizado foi o *Power* BI, para gerar combinações via planilha excel com dados coletados e assim criar mapas e gráficos objetivando melhor visualização de alguns padrões. Uma vez que os resultados foram organizados, observou-se desde a perspectiva da UTAUT, procurando associar os principais problemas ao não cumprimento das variáveis antecessoras Expectativa de Performance, Expectativa de Esforço, Condições Facilitadoras e Influência Social na Intenção e uso do aplicativo Uber.

4. Resultados e Análises

Segundo Almico & Faro (2014), o Iramuteq é uma solução informática para análise de textos, que busca apreender a estrutura e a organização do discurso, informando as relações entre os mundos lexicais mais frequentemente enunciados pelo sujeito. Por meio dos registros contidos no word, o Iramuteq realizou 5 análises sobre os registros. Três delas com toda a amostra de 142 e duas delas comparando os anos de 2016 e 2018.

4.1. Classificação Hierárquica Descendente

Método da Classificação Hierárquica Descendente (CHD) o texto dividido em segmentos de aproximadamente três linhas (ST), são classificados em função dos seus respectivos vocabulários, e o conjunto deles é repartido em função da frequência das formas reduzidas. A partir de matrizes cruzando formas reduzidas e ST (em repetidos testes do tipo x^2), aplica-se o método de CHD e obtém-se uma classificação definitiva. Segundo Vizeu Camargo & Justo (2013), a CHD permite a análise de segmentos de textos que apresentam vocabulário semelhante entre si e ao mesmo tempo diferente de outros segmentos de texto, validando a formação das classes calculando distancias e proximidades a partir do teste Chi-Quadrado Vizeu Camargo & Justo (2013). Esta análise visa obter classes de ST que, ao mesmo tempo, apresentam vocabulário semelhante entre si, e vocabulário diferente dos segmentos das outras classes. A partir dessas análises o software organiza a análise dos dados em um dendograma que ilustra as relações entre as classes. Para Reinert (1990), quando analisou a literatura, numa visão interpretativa,

reconheceu nas classes a "noção de mundo", enquanto um quadro perceptivo cognitivo com certa estabilidade temporal associada a um ambiente complexo.

Assim, foram analisados os 142 registros e foram encontrados 437 segmentos de texto, com um aproveitamento de 77,38%, validando a análise. Santos et al. (2017), explica que para uma análise CHD ser válida necessita de ao menos 70% de aproveitamento. Os discursos encontrados foram divididos em 5 classes (Figura 2).

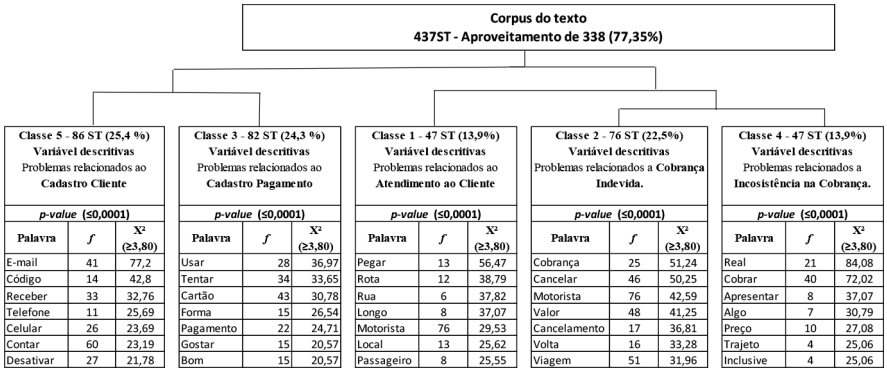


Figura 2 – Análsie de CHD

As classes foram organizadas segundo o discurso e estão ordenados pelo x^2. Na Figura 3 pode-se observar dois grupos, um formado pelas classes 5 e 3 e outro formado pelas classes 1,2 e 4. Isto significa que o discurso da classe 3 e 5 possuem proximidade, enquanto no segundo grupo , os discursos 2 e 4 possuem uma proximidade entre si, sendo parte do discurso do grupo 1. Vizeu Camargo & Justo (2013), explicam que esta organziação deve respeitar as palavras com $x^2 \geq 3{,}84$ e p-value < 0,05. Isso garante que as classes são organizadas pelo discurso contido e não pela casualidade.

A Classe 5 está formada por palavras como E-mail, Código, Receber, Telefone, Celular e Contar. A análise textual revela que está classe representa os problemas relacionados ao acesso ao sistema, onde os clientes reclamam dos códigos não recebidos, e-mails de reclamação sobre o não funcinamento do aplicativo, desativações de conta do usuário sem motivo para o cliente, demonstrando problemas relacionados ao cadastro dos clientes, por isso esta classe foi nomeada: **Cadastro Cliente**

falha ao entrar a sua conta foi desativada por favor contate support uber com para uma resolução já mandei um e-mail para o suporte mas até agora não recebi uma resposta

Cliente_77 *Respondida *Belo_Horizonte *MG *070916

Na classe 3 estão palavras como: Usar, Tentar, Cartão, Forma, Pagamento, Gostar e Bom. Aqui estão descritos os problemas relacionados ao cadastro e uso de formas de pagamento, falta de comunicação entre aplicativo e rede bancária do cliente. As palavras gostaria e boa vem da expressão de "gostaria de uma explicação" e "boa noite", usadas para iniciar o registro. Esta classe foi nomeada: **Cadastro Pagamento**

"Cartão inválido. Boa noite, gostaria de saber porque meu cartão consta como inválido da noite pro dia estava usando normalmente estou com limite disponível já tentei cadastrar até uma nova conta e ainda assim meu cartão está inválido"

Cliente_137 *Respondida *Guarulhos *SP *070916

Assim a Classe 5 e 3 formam um mesmo grupo de problemas com origem no SI do Aplicativo, revelando que as duas classes mais fortes dizem respeito a uma necessidade de melhoria do aplicativo. Pode-se perceber que para estes dois grupos que possuem o discurso mais forte nos registros, ainda existe um gap nas condições facilitadoras (Venkatesh et al., 2003), pois o cliente encontra problemas com o aplicativo e não encontra soluções via a empresa, tendo que recorrer a plataforma "Reclame aqui".

A Classe 1 possui as palavras Pegar, Rota, Rua, Longo, Motorista, Local, Passageiro. A análise das palavras no texto revela que esta classe se trata das reclamações dos clientes ao motorista, que faz rotas mais longas, além da falta de preparação em lidar com o cliente. Assim que esta classe foi nomeada: **Atendimento ao Cliente**

"Na rua haddock lobo nos jardins cidade de São Paulo meu destino era o aeroporto internacional de Guarulhos o motorista estava seguindo o waze que deu como caminho mais rápido uma rota pelo centro de São Paulo".

Cliente_128 *Respondida *Goiânia *GO *070916

Os problemas de atendimento (Classe 1) relatados devem ter uma atenção especial, pois com o advento das redes sociais um mau atendimento pode gerar uma campanha negativa sobre a empresa, impactando na Influência Social do indivíduo(Venkatesh et al., 2003).

A Classe 2 está representada pelas palavras Cobrança, Cancelar, Motorista, Valor, Cancelamento, Volta e Viagem. Analisando os registros pode-se perceber que aqui estão contidas as reclamações a respeito das cobranças indevidas, muitas delas relacionadas a não fechamento da corrida pelo motorista e outras com cobrança em duplividade. Também a aparecem problemas relacionados a cancelamento da corrida e cobrança aplicada. Assim esta classefoi chamada **Cobrança Indevida.**

"Só descobri porque recebi uma notificação avisando que o motorista estava ok cancelei a viagem logicamente ou aceitava a corrida ou cancelava ai depois recebo essa cobrança por e mail do cancelamento no valor de R$ 7,00".

Cliente_132 *Respondida *São_Paulo *SP *070916

E finalmente a Classe 4, formada pelas palavras: Real, Cobrar, Apresentar, Algo, Preço, Trajeto e Inclusive. Estas palavras aparecem nos discursos explicando problemas que os clientes tiveram em relação a cobrança de um trajeto em especial, que usualmente está ligado a tarifa dinâmica que pode alterar o valor de um mesmo trajeto a depender da demanda. Alguns clientes também reportam que o valor expresso no aplicativo antes de subir ao veículo e o valor cobrado após não conferem. Assim esta classe é chamada **Incosistência na Cobrança.**

"Ligamos não atendeu quando cancelamos cobrou uma taxa de 10 00 reais gostaríamos do reembolso por falta e de respeito ao cliente não concordo pagar esse valor tenho os print de tela como prova"

Cliente_6 *Não_respondida * Guarulhos*SP *070918

Pode-se perceber que os discursos das classes 2 e 4 estão muito próximos e ambos falam de problemas no faturamento. A grande diferença é qua na classe 2 a raiz do problema está no motorista, enquanto na classe 4 a raíz do problema está no aplicativo e na usabilidade do cliente. Embora tenham raízes diferentes, ambos os problemas descritos podem ser corrigidos ou amenizados com a melhoria da percepção das condições facilitadoras (Venkatesh et al., 2003).

4.2. Análise Fatorial de Correspondência

A Análise Fatorial de Correspondência (AFC) associa textos com variáveis, ou seja, possibilita a análise da produção textual em função das variáveis de caracterização. Ao realizar a AFC e colocar em um plano cartesiano pode-se perceber o seguinte resultado (Figura 3).

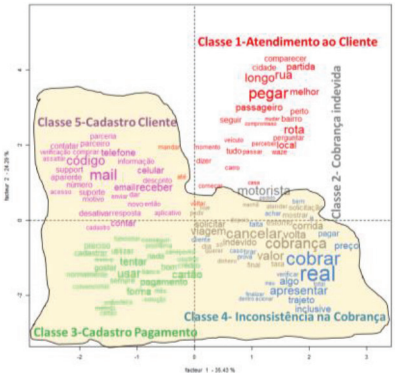


Figura 3 – Análise Fatorial de Correpondência (AFC)

Os discursos mais centrais são aqueles mais comuns e a medida que se afastam do centro vão ganhando um formato mais diferencido. Das 5 classes separadas representadas no plano cartesiano da AFC, pode-se perceber que as Classes 3, 5 e 4 estão diretamente relacionadas a problemas no aplicativo, seja na arquitetura do aplicativo, seja na não compreensão da Usabilidade, revelando o papel funamental das pesquisas de aceitação de SI para uma iniciativa como Uber.

Na perspectiva da usabilidade pode-se perceber que existe uma grande prerrogativa centrada nas Condições Facilitadoras (Venkatesh et al., 2003), materializada nas opções que a organização oferece Analisando quais categorias receberam mais reclamações, SI aparece em primeiro lugar com 79 reclamações registradas, seguido de reclamações sobre os aspectos financeiros, com 34 e Atendimento ao cliente, com 18. Na Figura 5 aparecem as principais categorias dos problemas encontrados e uma análise isolada da principal categoria (SI) comparando o ano de 2016 e 2018**.**

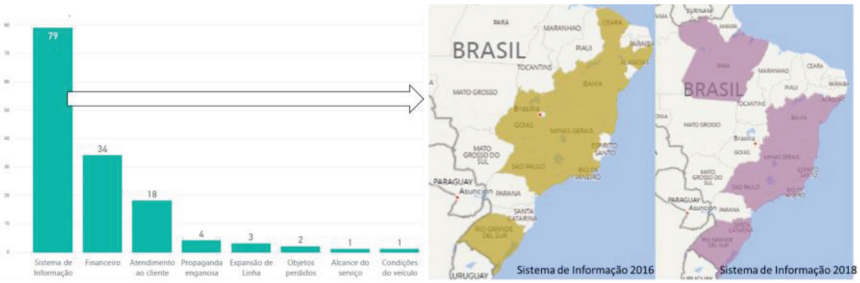


Figura 4 – Análise por categoria e comparação da categoria SI 2016/ 2018

Figura 5 – Nuvem de palavra comparada entre 2016/ 2018

Quando comparado os anos de 2016 e 2018, encontrou-se que em 2016 a quantidade de problemas referentes a SI, como cadastro do consumidor, cadastro da forma de pagamento, funcionamento do aplicativo, entre outros, foi de 52 registros, enquanto em 2018 passou para 27. Pode-se perceber muitos estados continuaram tendo problemas relacionados ao SI e outros Estados apareceram na amostra, como Pará, Santa Catarina e Espírito Santo, enquanto outros despareceram como Ceará, Brasília e Goiás. Na Categoria SI, o principal tipo de reclamação são aquelas relacionadas ao cadastro de cliente ou meio de pagamento, com 61 registros. Separando o ano de 2016 e 2018 pode-se perceber a construção da Figura 4. As palavras mais utilizadas em 2016 foram desativar, e-mail, contar sempre, deixando em evidência os problemas relacionados a cancelamentos de conta de usuário sem motivos aparentes, enquanto em 2018 aparecem paravras como cancelar, pedido, atendimento, cpf, motorista e valor, exemplificando as principais reclamações referentes a cancelamento das viagens, má conduta do motorista, cadastro não aceitando o CPF (cadastro de pessoas físicas) e problemas relacionados a valor cobrado. As análises via Figuras 4 e 5 permitem compreender a expansão da Organização Uber, pois em 2016, prevaleciam problemas de cancelamento espontâneo das contas dos usuários, muitas vezes sem explicação, demonstrando que o sistema ainda precisava de melhorias, em 2018 os problemas foram substituídos por problemas relacionado ao uso do aplicativo, como cancelamento de viagens e política de cobrança. Comparando os anos percebe-se que inicialmente existiam problemas mais críticos quanto ao aplicativo, como desativações das contas de maneira espontânea, já em 2018 os problemas estão muito voltados a interação humana quanto ao uso ou mal uso do aplicativo, algumas vezes por

desconhecimento e outras vezês por má fé do motorista, como visto nos trechos citados. Assim, mesmo que em cada ano (2016 e 2018) existam problemas diferentes um ajuste no aplicativo promovendo maior condição facilitadora (Venkatesh et al., 2003), podem inibir ações humanas seja do mal uso por desconhecimento ou intencional.

4.3. Implicações práticas

Os resultados deste estudo contribuem na afirmação que as empresas devem adotar como etapa em seu processo de aquisição de tecnologia, o grau de aceitação e uso de seus usuários. Este é um passo decisivo das empresas para caminheram para a materialização de uma cidade inteligente. Gonçalves & Paiva (2014), explicam que a smart city é uma construção coletiva onde a alma não é a tecnologia. Deste modo, pode-se perceber que o conceito cidade inteligente em realidade um Proxy de uma aceitação da tecnologia de maneira conjunta, assim como seu compartilhamento de informação, representando a união de vários aspectos de aceitação e compartilhamento. Capdevila & Zarlenga (2015) ratificam esta visão explicando que a necessidade do processo de difusão da tecnologia inicialmente a nível indivíduo e apenas depois pensar em cidade inteligentes como conceito.

5. Conclusões

Este estudo teve como problema conhecer quais são os principais desafios do Uber na expansão do aplicativo no Brasil. Foi observado mediante os discursos que os principais problemas estão relacionados ao Cadastro do Cliente e Formas de Pagemento, Atendimento, Inconsistência no pagamento e Cobrança indevida. Embora a formação de recurso humano seja importante, sabe-se que a melhoria do aplicativo ameniza e corrige a maioria dos problemas levantados. Os cinco discursos encontrados foram classificados e ratificaram a necessidade do estudo da tecnologia a nível individual antes de se pensar em conceitos maiores como o de Cidades Inteligentes (*Smart Cities*). Assim, o objetivo deste estudo que foi conhecer os principais discursos dos usuários de Uber no Brasil, atentando para o papel dos SI foi cumprido, sendo observado que três (Cadastro do Cliente, Cadastro das formas de Pagamento e Incosistencia no pagamento) estão ligadas a problemas no aplicativo, enquanto o Atendimento ao cliente e a Cobrança indevida estão relacionados a conduta humana. O Uber representa materialmente um dos melhores exemplos do futuro das cidades, conhecer seus principais desafios foi importante para sugerir possibilidades de avançar neste campo. Para futuros estudos aconselha-se ampliar a pesquisa para uma amostra mais representativa a fim de ampliar os achados e contribuir com o tema.

Referências

Almico, T., & Faro, A. (2014). Enfrentamento de cuidadores de crianças com câncer em processo de quimioterapia. Psicologia, saúde & doenças, 15(3), 723–737.

Bobsin, D., Visentini, M. S., & Rech, I. (2009). Em busca do estado da arte do UTAUT: ampliando as considerações sobre o uso da tecnologia. RAI-Revista de Administração e Inovação, 6(2), 99–118.

Capdevila, I., & Zarlenga, M. I. (2015). Smart city or smart citizens? The Barcelona case. Journal of Strategy and Management, 8(3), 266–282.

Cosgrave, E., Arbuthnot, K., & Tryfonas, T. (2013). Living labs, innovation districts and information marketplaces: A systems approach for smart cities. Procedia Computer Science, 16, 668–677.

Couclelis, Helen (2004) The Construction of the digital city. Environment and Planning B: Planning and Design, Vol. 31, Iss. 1, 5–19.

Dameri, Renata Paola (2016) Smart City Implementation. Creating Economic and Public Value in Innovative Urban Systems. Springer International Publishing AG, Cham, Switzerland

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: a comparison of two theoretical models. Management science, 35(8), 982–1003.

Defossez, D. A. L. (2017). The regulation of a project of the deregulation: UBER in Brazil and the European Union. Brasília: Universidade de Brasilia (UnB). Economia

Dusi, L. D. A. (2017). O uso de aplicativos para smartphone no transporte individual: 99Taxis e Uber. Monografia em projeto final em Transportes.

Florida, Richard (2005) Cities and the Creative Class. Routledge, New York and London.

Gonçalves, R., & de Paiva, A. (2014). Smart cities, smarters citizens. Conjuntura da Construção, 12(3), 17–18.

Graeml, A. R. (1998). O valor da tecnologia da informação. Simpósio de Administração da Produção, Logística e Operações Industriais, 1.

Guelzim, T., Obaidat, M. S., & Sadoun, B. (2016). Introduction and overview of key enabling technologies for smart cities and homes. In Smart Cities and Homes (pp. 1–16).

Komninos, Nicos (2002) Intelligent Cities: Innovation, Knowledge Systems and Digital Spaces. Spon Press, London and New York

Legris, P., Ingham, J. e Collerette, P. (2003). Por que as pessoas usam a tecnologia da informação? Uma revisão crítica do modelo de aceitação de tecnologia. Informação e gestão , 40 (3), 191–204.

Mariano, A. M., & Diaz, L. F. A. (2017) A importância da aceitação e uso da tecnologia em aplicativos de mobilidade urbana: contribuições da literatura científica. VII Congresso Brasileiro CONBREPRO. Ponta Grossa. PR-Brasil

Mathieson, K., & Ryan, T. (1994). The effect of definitional variations on users' evaluations of information systems. ACM SIGMIS Database: the DATABASE for Advances in Information Systems, 25(2), 37–48.

Norman, B. (2018). Are autonomous cities our urban future?. Nature communications, 9.

O'Brien, J. A., & Marakas, G. M. (2013). Administração de sistemas de informação. AMGH.

Oliveira, T., Thomas, M., Baptista, G., & Campos, F. (2016). Mobile payment: Understanding the determinants of customer adoption and intention to recommend the technology. Computers in Human Behavior, 61, 404–414.

Reinert, M. (1990). Alceste une méthodologie d'analyse des données textuelles et une application: Aurelia De Gerard De Nerval. Bulletin of Sociological Methodology/ Bulletin de Méthodologie Sociologique, 26(1), 24–54.

Reclame aqui. (2018). Página de internet. Disponível em https://www.reclameaqui.com.br/, acesso em 14 de setembro de 2018.

Rogers, E. M. (1995). Lessons for guidelines from the diffusion of innovations. Joint Commission Journal on Quality and Patient Safety, 21(7), 324–328.

Rolfstam, M., Phillips, W., & Bakker, E. (2011). Public procurement of innovations, diffusion and endogenous institutions. International Journal of Public Sector Management, 24(5), 452–468.

Rondan-Cataluña, F. J., Arenas-Gaitán, J., & Ramírez-Correa, P. E. (2015). A comparison of the different versions of popular technology acceptance models: A non-linear perspective. Kybernetes, 44(5), 788–805.

Santos, V.E.P. ; Salvador, P.T.C.O. ; Gomes, A.T.L. ; Chiavone, F. B. T. ; Alves, K.Y.A ; Bezerril, M.S. (2017) . IRAMUTEQ nas pesquisas qualitativas brasileiras da área da saúde: scoping review. In: 6º Congresso Ibero-americano em investigação qualitativa, 2017, Salamanca. Atas CIAIQ 2017, v. 2

Silva, B. N., Khan, M., & Han, K. (2018). Towards sustainable smart cities: A review of trends, architectures, components, and open challenges in smart cities. Sustainable Cities and Society, 38, 697–713.

Tuominen, E. (2018). Diffusion of innovations in a smart city context: Facilitation of innovation-oriented smart city development in Kalasatama.

Uber Technologies Inc. | Uber do Brasil Tecnologia Ltda (2018). Página de internet. Disponível em https://www.uber.com/pt-BR/br/, acesso em 14 de setembro de 2018.

Venkatesh, V., Morris, MG, Davis, GB, & Davis, FD (2003) User acceptance of information technology: Toward a unified view. MIS quarterly, 425–478.

Verwijnen, Jan – Lehtovuori, Panu (1999) The Creative City's New Field Condition. Can Urban Innovation and Creativity Overcome Bureaucracy and Technocracy? In: Creative Cities. Cultural Industries, Urban Development and the Information Society. ed. by: Jan Verwijnen – Panu Lehtovuori, 12–35. University of Art and Design Helsinki UIAH Publishing Unit, Helsinki.

Visvizi, A., & Lytras, M. (2018). It's Not a Fad: Smart Cities and Smart Villages Research in European and Global Contexts.

Vizeu Camargo, B., & Justo, A. M. (2013). IRAMUTEQ: um software gratuito para análise de dados textuais. Temas em psicologia, 21(2).

Williams, M. D., Rana, N. P., & Dwivedi, Y. K. (2015). The unified theory of acceptance and use of technology (UTAUT): a literature review. Journal of Enterprise Information Management, 28(3), 443–488.

Yigitcanlar, Tan (2009) Planning for knowledge-based urban development: global perspectives. Journal of Knowledge Management, Vol. 13, Iss. 5, 228–242.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 09/09/2018
Aceitação/Acceptance: 01/12/2018

Regresión con polinomios ortogonales completos de alto grado para mejorar la predicción de la distancia mediante RSSI

Diego Ordóñez-Camacho [1], Héctor Ávalos [2], Estevan Gómez [3]

dordonez@ute.edu.ec, havalos@ute.edu.ec, estevan.gomez@ute.edu.ec

[1,2,3] Universidad UTE, Grupo de Investigación en Informática (GrIInf), Quito, Ecuador.

Pages: 584–594

Resumen: Un sistema de posicionamiento en interiores permite determinar la ubicación de una persona o equipo, donde el uso del GPS no es posible. El uso de técnicas de radiofrecuencia y RSSI presenta ventajas a nivel de costo e infraestructura; sin embargo, la señal es muy inestable y debe preprocesarse para poderla usar con éxito en la inferencia de la distancia; mientras más exacta sea la distancia predicha ente receptor y emisores, más exacta será la determinación de la ubicación. Este trabajo construye sobre estudios previos de filtrado de la señal, y los mejora estudiando la aplicación de técnicas de regresión no lineales, más específicamente regresiones polinomiales de alto grado, que proporcionen una predicción más ajustada al patrón que sigue la señal RSSI al variar la distancia entre emisor y receptor.

Palabras-clave: sistema de posicionamiento en interiores; RSSI; Bluetooth; filtros de señal; regresiones.

High degree complete orthogonal polynomials to improve distance prediction based on RSSI

Abstract: An indoor positioning system allows to determine the location of a person or device, when using the GPS is not possible. The use of radiofrequency and RSSI techniques presents advantages regarding cost and infrastructure; however, the signal is very unstable and must be preprocessed to use it successfully in the inference of distance; The more accurate the predicted distance between the receiver and the emitters, the more accurate the determination of the location will be. This work builds on previous studies of signal filtering, improving them by studying the application of non-linear regression techniques, more specifically high-degree polynomial regressions, which provide a better prediction of the pattern followed by the RSSI signal as distance between sender and receiver varies.

Keywords*:* indoor positioning systems; Bluetooth; RSSI; signal filters; regressions.

1. Introducción

Los sistemas de posicionamiento en interiores (SPI) permiten determinar la ubicación de una persona o equipo cuando las condiciones del entorno impiden el uso de técnicas clásicas de ubicación en exteriores como el GPS (*Global Positioning System*); bajo techo, o en interiores, puede ser infructuoso tratar de fijar la señal de los satélites GPS (Li, Lam, & Wang, 2018).

Entre las técnicas más utilizadas para un SPI pueden mencionarse el uso de visión computarizada para reconocer el entorno (Zhang & Zhou, 2018), el uso de iluminación inteligente que permite detectar el cambio de frecuencia de la luz (US9964630B2, 2018), o el uso de emisores de radiofrecuencia, como por ejemplo balizas Bluetooth (Chu & Chien, 2019), en todos los casos para inferir la ubicación del interesado. Es esta última técnica aquella tomada en cuenta en este estudio por considerarse la menos invasiva a nivel de infraestructura y la menos exigente a nivel computacional, al compararla con la iluminación inteligente y con la visión computarizada, respectivamente.

El uso de radiofrecuencia, Bluetooth en el caso de esta investigación, requiere que se instalen emisores en posiciones bien conocidas y con un mínimo de obstrucciones. En el caso de Bluetooth, y más específicamente de balizas Bluetooth Low Energy (BLE), estos emisores suelen ser muy pequeños y operan con baterías de una larga duración (varios años). Las balizas emiten una señal que debe ser capturada y procesada por los receptores quienes deberán calcular su posición mediante ella. De la señal de radiofrecuencia se suele utilizar la *Received Signal Strength Indicator* (RSSI), que indica la potencia de la señal, en decibeles, al momento de la recepción. La potencia de la RSSI disminuye con la distancia ("IEEE Standard for Information technology–Telecommunications and information exchange between systems Local and metropolitan area networks–Specific requirements Part 11", 2012), razón por la cual se puede usar como el primer paso en detectar la ubicación de un equipo receptor: en primer lugar se infiere la distancia a los emisores utilizando la RSSI y luego se calcula la posición mediante técnicas de triangulación.

Es en la inferencia de la distancia donde radica el éxito de la técnica. Mientras más exacta sea la distancia predicha ente receptor y emisores, más exacta será la determinación de la ubicación. No es de extrañarse entonces, que sea el paso más complejo; esto se debe principalmente a la alta volatilidad de la señal (Wu, Lee, Tseng, Jan, & Chuang, 2008), ya que es muy susceptible a cambios bruscos debido a las condiciones en el entorno: las personas que cruzan el área y atenúan la señal, las paredes donde la señal puede rebotar, la cantidad de metal en las estructuras que distorsiona la señal, entre otras.

La raíz del problema tiene más que ver con condiciones físicas que parecerían inherentes a la tecnología de radiofrecuencia; sin embargo, se han estudiado métodos para paliar el problema, en general relacionados con técnicas de filtrado y suavizado de la señal, muchos con resultados sumamente alentadores. Este trabajo construye sobre estos estudios previos de filtrado de la señal, y los mejora estudiando la aplicación de técnicas de regresión no lineales, más específicamente regresiones polinomiales de alto grado, que proporcionen una predicción más ajustada al patrón que sigue la señal RSSI al variar la distancia entre emisor y receptor.

1.1. Trabajos relacionados

Estudios previos han demostrado que ciertos filtros digitales de suavizado, pueden tener un efecto positivo al ser aplicados a la señal RSSI, aumentando de esta manera la calidad de la estimación de la distancia. Entre los casos más representativos tenemos el uso de filtros de media móvil (Bellavista, Corradi, & Giannelli, 2006; Ordóñez-Camacho, Gómez, & Ávalos, 2018), métodos iterativos de mínimos cuadrados (Lin, Lu, Chen, Li, & Chuang, 2017), sistemas distribuidos de trazas de señal (Luo et al., 2016), tomas múltiples de datos (Xue, Qiu, Hua, & Yu, 2017) y filtros dinámicos adaptativos (Ordóñez-Camacho & Cabrera-Goyes, 2018). Otros trabajos recurren a soporte adicional de hardware, como en el caso del giroscopio (Heo & Kwon, 2017) o utilizando antenas sectoriales (Nagaraju, Gudino, Kadam, Ookalkar, & Udeshi, 2016).

Otro campo en el cual se investiga activamente para mejorar la precisión de la estimación es el relacionado con las técnicas de regresión o predicción; esto puede venir ligado también a filtros de señal como los ya mencionados, los cuales se aplican en un paso previo. En (E. Cabrera-Goyes & Ordóñez-Camacho, 2017; Edwin Cabrera-Goyes & Ordóñez-Camacho, 2018) se aplica primero un filtro de media móvil simple y luego se trabaja con regresiones lineales para la estimación final; En (Onofre, Caseiro, Pimentão, & Sousa, 2016) se utiliza lógica difusa, en (Jadidi, Patel, & Miro, 2017) se utilizan procedimientos gaussianos, en (Yu & Guo, 2017) se trabaja con filtros de Kalman, mientras que en (Aykaç, Erçelebi, & Aldin, 2017) se emplean filtros de partículas.

En cuanto al uso de regresiones polinomiales podemos citar a (Finlayson, Mackiewicz, & Hurlbert, 2015; Lei et al., 2016; Sallouha, Chiumento, & Pollin, 2017) que se aproximan al objetivo del presente trabajo. En general estos trabajos suelen utilizar polinomios de orden no tan alto (dos a seis), lo cual en general se justifica dado que los polinomios de alto grado tienden a ser muy sensibles a los *outliers*, a sobreajustar los datos de entrenamiento y a incrementar el error en los datos de test (Gelman & Imbens, 2018). Sin embargo, esto puede evitarse, como se destaca en (Hall & Racine, 2015), al utilizar una cantidad de datos suficientemente grande, filtrando los *outliers* y con un método sólido de test que permita minimizar el sobreajuste. De la misma manera, el uso de polinomios puros puede ocasionar una excesiva correlación entre los términos, e inestabilidad numérica, por lo cual se considera recomendable trabajar con polinomios ortogonales (Gautschi, 2018).

2. Metodología

El primer paso en esta experimentación consistió en recurrir a la base de datos utilizada y descrita en (Ordóñez-Camacho et al., 2018) de manera que sea posible establecer *benchmarks* y comparaciones estables con respecto a otros métodos previamente aplicados. La base consiste de más de 1700 muestras de RSSI Bluetooth, capturadas en un corredor de 28 metros. Las muestras se tomaron con dos receptores que se desplazaron en el corredor con intervalos de un metro, recolectando la información de dos emisores ubicados a los extremos del corredor; cuatro lotes en total.

La información fue recolectada gracias a una aplicación Android específicamente diseñada para el efecto; los datos en bruto fueron luego brevemente preprocesados en Excel para integrarlos en una base de datos indicando el lote, el ordinal del momento

de la toma de la muestra, el valor RSSI de la señal y la distancia real entre el emisor y el receptor.

Luego se trabajaron los datos en el *software* estadístico R[1], en el cual en primer lugar se aplicó una serie de filtros de suavizado tipo *Exponential Moving Average*, EMA (*Ecuación 1*), variando el peso de la última toma con respecto al acumulado previo con una relación alpha entre 0.001 y 0.05. Los datasets filtrados se convirtieron en polinomios de grado variable entre 1 y 23; se utilizaron polinomios completos, calculados mediante el paquete *stats*[2], tanto puros (*Ecuación 2*), como ortogonales (*Ecuación 3*), estos últimos resueltos por *poly*[3] mediante el método de Gram-Schmidt (Gentle, 2017).

$$S_t = \alpha Y_t + (1-\alpha) S_{t-1} \tag{1}$$

$$P(x) = \sum_{i=0}^{n} a_i x^i \tag{2}$$

$$\deg P_m = n, \langle P_m, P_n \rangle = 0 \ \ for \ \ m \neq n \tag{3}$$

Los datasets resultantes se sometieron al cálculo de regresión mediante *lm*[4], utilizando como método de test el de validación cruzada de 10 iteraciones (Refaeilzadeh, Tang, & Liu, 2016; Ribeiro, Singh, & Guestrin, 2016). Las tablas resultantes se graficaron y analizaron en busca de la mejor relación de error. Finalmente se verificó si las diferencias en la reducción del MAE eran significativas mediante un análisis de Students, efectuado con *t.test*[5].

3. Resultados

El caso de base para el análisis de la predicción de la distancia con base en RSSI comprende el uso de la regresión lineal para construir un modelo predictivo. En este experimento se trabajó inicialmente esta técnica y los valores resultantes se presentan en la *Figura 1*. En primer lugar se construyó una regresión lineal con los datos sin filtrar, cuyo error medio absoluto, mejor conocido como MAE (*Mean Absolute Error*) es de 4.8076 m, marcado por la línea roja en la figura; luego se construyó una serie de modelos lineales simples sobre los datos filtrados mediante EMA, con alpha variable desde 0.001 hasta 0.05, en intervalos de 0.01; la progresión de estos valores está representada por la serie en azul de la figura, donde el error mínimo de 3.5539 m, se consigue con un alpha de 0.006; la reducción del error es del 26 %.

El siguiente experimento, ligado más directamente al objetivo específico de este trabajo, consistió en combinar los filtros EMA, tal como en el caso anterior, pero ahora con modelos de regresión polinomial para romper la linealidad. Se construyeron polinomios

1 https://www.r-project.org

2 https://www.rdocumentation.org/packages/stats

3 https://www.rdocumentation.org/packages/stats/versions/3.5.1/topics/poly

4 https://www.rdocumentation.org/packages/stats/versions/3.5.1/topics/lm

5 https://www.rdocumentation.org/packages/stats/versions/3.5.1/topics/t.test

completos de grado variable entre 1 y 23 (grado máximo permitido, con estos datos, por la librería utilizada), cuyos resultados se pueden ver en la *Figura 2*, donde la región de interés, aquella con el error mínimo, aparece en azul; el error mínimo es de 3.3715 m, conseguido con polinomios de grado 21 y con alpha de 0.006. La reducción del error con respecto al caso de base filtrado es del 5 %, y del 30 % con respecto a la regresión lineal sin ningún filtro.

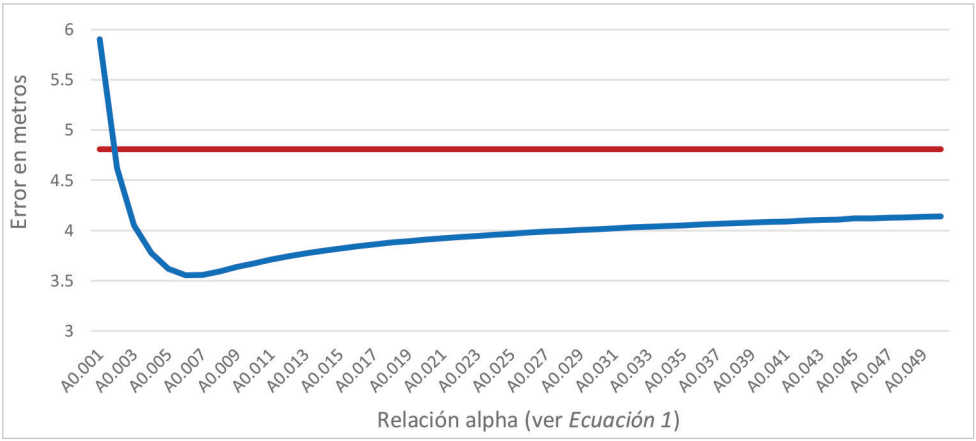


Figura 1 – Error al aplicar regresión lineal. La línea roja en 4.8076 marca el error al no filtrar los datos; la línea azul representa el error al variar alpha en un filtro EMA; el error mínimo es de 3.5539 m, con alpha de 0.006.

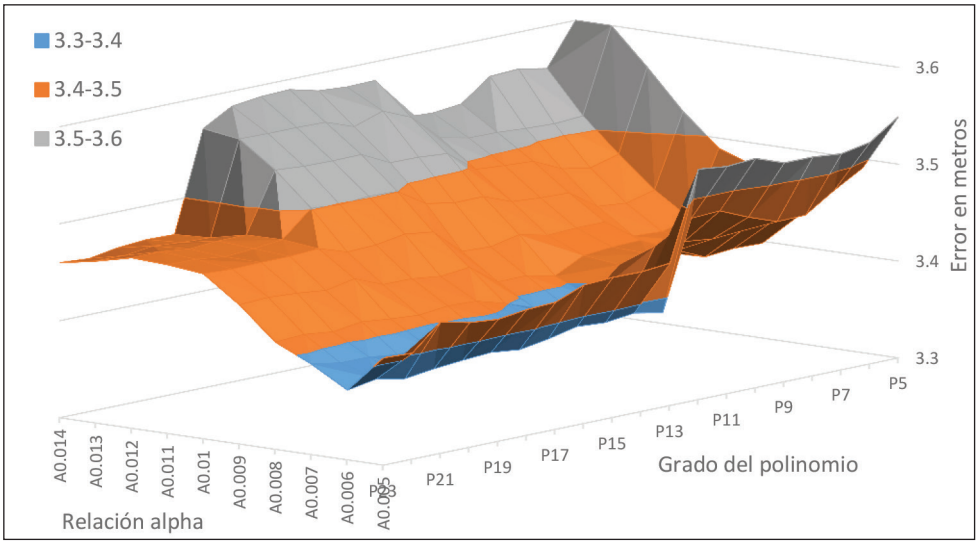


Figura 2 – Error mínimo, marcado por la región en azul, utilizando regresión polinomial en bruto; el error mínimo de 3.3715 m, se consigue con un polinomio de grado 21 y un alpha de 0.006.

Como experimento final se aplicaron los mismos filtros EMA ya utilizados, contrastándolos ahora con modelos de regresión construidos mediante polinomios ortogonales, los cuales evitan la inestabilidad numérica ya descrita en la literatura referenciada. Igual que en el caso anterior se construyeron polinomios completos de grado variable entre 1 y 23, y los resultados se aprecian en la *Figura 3*; el error mínimo se encuentra en la región azul y es de 3.1698 m, conseguido con polinomios de grado 23 y con alpha de 0.019. Se reduce aún más el error con respecto al caso anterior con un 6 %, y con respecto al caso de base la reducción total es del 34 %. El coeficiente alpha requerido por el filtro de suavizado para alcanzar el error mínimo, corresponde a 52 tomas de datos, las cuales requirieron aproximadamente 1 minuto para que el receptor las recolecte de un promedio de 3 emisores. Los coeficientes obtenidos para el modelo calculado con polinomios ortogonales se pueden ver en la *Tabla 1*.

Intercept	**C1**	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**
13.4367	-240.1335	-26.0742	-28.8681	38.009	36.3965	-33.6732	-16.5246
C8	**C9**	**C10**	**C11**	**C12**	**C13**	**C14**	**C15**
31.6995	22.0149	-10.9396	-58.3266	7.816	56.9486	-3.7032	-34.3443
C16	**C17**	**C18**	**C19**	**C20**	**C21**	**C22**	**C23**
-14.3089	18.667	15.989	-4.8117	-15.756	0.8135	17.8359	-8.9568

Tabla 1 – Coeficientes del modelo calculado con polinomios ortogonales de grado 23

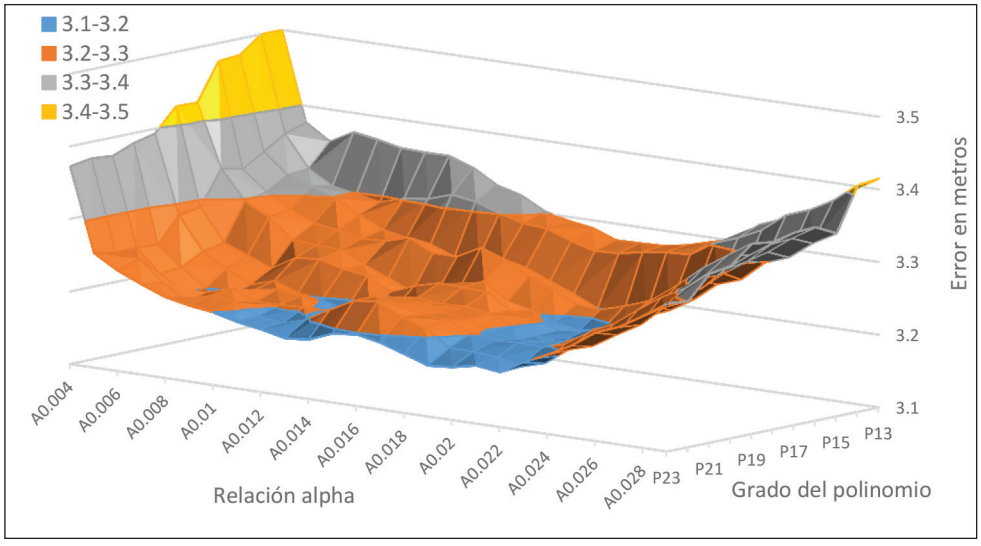


Figura 3 – Error mínimo, marcado por la región en azul, utilizando regresión polinomial ortogonal; el error mínimo de 3.1698 m, se consigue con un polinomio de grado 23 y un alpha de 0.019.

Las diferencias existentes entre las reducciones de error son en todos los casos altamente significativas, como se puede apreciar en la *Tabla 2*, donde en cada celda se tiene, si se compara la técnica anotada en la columna versus la técnica anotada en la fila, la reducción del MAE en porcentaje, y el valor *p* de significancia estadística de dicha reducción.

Tabla 2 – Análisis de la significancia en la disminución del error mediante una prueba *t*. Los encabezados de fila y columna muestran el código de la técnica y el error mínimo. Las celdas muestran la reducción relativa del error (%) y el valor *p* de significancia

	LF (3.5539)	**P** (3.3715)	**O** (3.1698)
LSF (4.8076)	26 % / 2.2e-16	30 % / 2.2e-16	34 % / 2.2e-16
LF (3.5539)		5 % / 1.724e-06	11% / 2.806e-09
P (3.3715)			6 % / 0.0004225
LSF: lineal sin filtro; LF: lineal filtrado; P: polinomial; O: ortogonal			

4. Discusión

Los resultados de este trabajo, al compararlos con otros previos, resultan muy alentadores. Con respecto a (E. Cabrera-Goyes & Ordóñez-Camacho, 2017; Edwin Cabrera-Goyes & Ordóñez-Camacho, 2018) se obtiene una mejora del 30 % en cuanto a la exactitud de la predicción. La experimentación realizada en (Ordóñez-Camacho & Cabrera-Goyes, 2018) reporta una reducción del 20 % en el error, mediante el uso del filtro adaptativo y regresión lineal; con respecto a dicho estudio, el presente trabajo obtiene una reducción adicional del 14 %. Las técnicas de filtrado utilizadas en (Ordóñez-Camacho et al., 2018) permiten obtener una reducción del error del 12 %, lo cual da a la técnica presentada en este estudio una ventaja de más del 20 % en dicho indicador.

El estudio de (Tóth, 2016) trabaja con una base de datos de similares características a la que se utilizó aquí y, si bien no presenta datos que puedan compararse, si recalca que el error puede llegar a los 5 m. Si bien el estudio de (Dahlgren & Mahmood, 2014) tampoco presenta datos que puedan compararse directamente con estos, sí permite, en sus conclusiones, confirmar la alta variabilidad que presenta el RSSI, lo cual justifica que, especialmente en distancias mayores a 5 m, el error sea significativo.

En una comparación más interna a la presente experimentación puede claramente observarse que la disminución del error es siempre significativa a medida que se van introduciendo nuevas técnicas de procesamiento. Si bien el solo hecho de aplicar el filtrado EMA aporta con la disminución más importante del error (26 %), la subsiguiente aplicación de las técnicas de regresión polinomial continúa produciendo una reducción destacable del MAE hasta alcanzar la reducción total del 34 %.

5. Conclusiones

En este trabajo se analizó el efecto que tiene sobre el cálculo de la distancia mediante la señal RSSI, la aplicación de técnicas de filtrado de la señal combinadas con distintos tipos de regresión. Se aplicaron filtros EMA con una relación variable del valor acumulado (alpha) y se aplicaron regresiones lineal, polinomial y polinomial ortogonal.

Los resultados confirmaron que el uso de los filtros de suavizado de la señal tiene una gran incidencia en la reducción del error, mediando una configuración apropiada de la relación entre la última muestra y el valor acumulado. Se verificó también que el uso de regresiones polinomiales, especialmente al utilizar polinomios ortogonales completos de alto grado, supera en este caso al uso de la regresión lineal, proporcionando finalmente un incremento en la precisión de más de un metro y medio en la predicción de la distancia.

Referencias

Aykaç, M., Erçelebi, E., & Aldin, N. B. (2017). ZigBee-based indoor localization system with the personal dynamic positioning method and modified particle filter estimation. *Analog Integrated Circuits and Signal Processing*, *92*(2), 263–279. https://doi.org/10.1007/s10470-017-0969-4

Bellavista, P., Corradi, A., & Giannelli, C. (2006). Evaluating filtering strategies for decentralized handover prediction in the wireless internet. En *Computers and Communications, 2006. ISCC'06. Proceedings. 11th IEEE Symposium on* (pp. 167–174). IEEE.

Cabrera-Goyes, E., & Ordóñez-Camacho, D. (2017). Towards a Bluetooth Indoor Positioning System with Android Consumer Devices. En *2017 International Conference on Information Systems and Computer Science (INCISCOS)* (pp. 56–59). https://doi.org/10.1109/INCISCOS.2017.14

Cabrera-Goyes, Edwin, & Ordóñez-Camacho, D. (2018). Posicionamiento en espacios interiores con Android, Bluetooth y RSSI. *Enfoque UTE*, *9*(1), 118–126. https://doi.org/10.29019/enfoqueute.v9n1.238

Chu, H.-C., & Chien, M.-F. (2019). An Adaptive Bluetooth Low Energy Positioning System with Distance Measurement Compensation. En K. J. Kim & H. Kim (Eds.), *Mobile and Wireless Technology 2018* (pp. 223–234). Springer Singapore.

Dahlgren, E., & Mahmood, H. (2014). *Evaluation of indoor positioning based on Bluetooth Smart technology*. Chalmers University of Technology, Göteborg, Sweden.

Finlayson, G. D., Mackiewicz, M., & Hurlbert, A. (2015). Color Correction Using Root-Polynomial Regression. *IEEE Transactions on Image Processing*, *24*(5), 1460–1470. https://doi.org/10.1109/TIP.2015.2405336

Gautschi, W. (2018). *A Software Repository for Orthogonal Polynomials*. SIAM.

Gelman, A., & Imbens, G. (2018). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, 1–10.

Gentle, J. E. (2017). Matrix Transformations and Factorizations. En J. E. Gentle (Ed.), *Matrix Algebra: Theory, Computations and Applications in Statistics* (pp. 227–263). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-64867-5_5

Hall, P. G., & Racine, J. S. (2015). Infinite order cross-validated local polynomial regression. *Journal of Econometrics*, *185*(2), 510–525. https://doi.org/10.1016/j.jeconom.2014.06.003

Heo, J., & Kwon, Y. (2017). Improved Indoor Positioning System Using BLE Beacons and a Compensated Gyroscope Sensor. En *Intelligent Robotics and Applications* (pp. 69–76). Springer, Cham. https://doi.org/10.1007/978-3-319-65298-6_7

IEEE Standard for Information technology–Telecommunications and information exchange between systems Local and metropolitan area networks–Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. (2012). *IEEE Std 802.11-2012 (Revision of IEEE Std 802.11-2007)*, 1–2793. https://doi.org/10.1109/IEEESTD.2012.6178212

Jadidi, M. G., Patel, M., & Miro, J. V. (2017). Gaussian processes online observation classification for RSSI-based low-cost indoor positioning systems. En *2017 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 6269–6275). https://doi.org/10.1109/ICRA.2017.7989743

Lei, T., Wen, X., Lu, Z., Jing, W., Zhang, B., & Cao, G. (2016). Handoff management scheme based on frame loss rate and RSSI prediction for IEEE 802.11 networks. En *2016 International Symposium on Wireless Communication Systems (ISWCS)* (pp. 555–559). https://doi.org/10.1109/ISWCS.2016.7600966

Li, F., Lam, K.-Y., & Wang, L. (2018). Global positioning system (GPS) on smart phones is often either unavailable or inaccurate in urban and natural canyons, indoors, and in forests. Hence it is augmented by use of other available radio frequency (RF) signals, such as WiFi with radio frequency identification and cellular, using their received signal strength, time difference of arrival and even angle of arrival. A novel methodology developed... *Wireless Networks*, *24*(7), 2735–2748.

Lin, K. H., Lu, C. C., Chen, H. M., Li, H. F., & Chuang, C. F. (2017). A modified least squares iteration for indoor positioning system. En *2017 IEEE International Conference on Consumer Electronics - Taiwan (ICCE-TW)* (pp. 109–110). https://doi.org/10.1109/ICCE-China.2017.7991019

Luo, Q., Yan, X., Li, J., Peng, Y., Tang, Y., Wang, J., & Wang, D. (2016). DEDF: lightweight WSN distance estimation using RSSI data distribution-based fingerprinting. *Neural Computing and Applications*, *27*(6), 1567–1575. https://doi.org/10.1007/s00521-015-1956-2

Lydecker, S. H., Klitenik, K., Ryan, D., Malandrakis, E. P., & Abou-Rizk, M. J. (2018). *US9964630B2*. United States. Recuperado de https://patents.google.com/patent/US9964630B2/en

Nagaraju, S., Gudino, L. J., Kadam, B. V., Ookalkar, R., & Udeshi, S. (2016). RSSI based indoor localization with interference avoidance for Wireless Sensor Networks using anchor node with sector antennas. En *2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)* (pp. 2233–2237). https://doi.org/10.1109/WiSPNET.2016.7566539

Onofre, S., Caseiro, B., Pimentão, J. P., & Sousa, P. (2016). Using Fuzzy Logic to Improve BLE Indoor Positioning System. En *Technological Innovation for Cyber-Physical Systems* (pp. 169–177). Springer, Cham. https://doi.org/10.1007/978-3-319-31165-4_18

Ordóñez-Camacho, D., & Cabrera-Goyes, E. (2018). An Adaptive-Bounds Band-Pass Moving-Average Filter to Increase Precision on Distance Estimation from Bluetooth RSSI. En *Proceedings of the International Conference on Information Technology & Systems (ICITS 2018)* (pp. 823–832). Springer, Cham. https://doi.org/10.1007/978-3-319-73450-7_78

Ordóñez-Camacho, D., Gómez, E., & Ávalos, H. (2018). Tamaño óptimo de la ventana en un filtro de media móvil para RSSI: mejorando la estimación de la distancia con Bluetooth. *Congreso de Ciencia y Tecnología ESPE*, *13*(1). https://doi.org/10.24133/cctespe.v13i1.709

Refaeilzadeh, P., Tang, L., & Liu, H. (2016). Cross-Validation. En *Encyclopedia of Database Systems* (pp. 1–7). Springer, New York, NY. https://doi.org/10.1007/978-1-4899-7993-3_565-2

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why Should I Trust You?": Explaining the Predictions of Any Classifier. En *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1135–1144). New York, NY, USA: ACM. https://doi.org/10.1145/2939672.2939778

Sallouha, H., Chiumento, A., & Pollin, S. (2017). Localization in long-range ultra narrow band IoT networks using RSSI. En *2017 IEEE International Conference on Communications (ICC)* (pp. 1–6). https://doi.org/10.1109/ICC.2017.7997195

Tóth, Z. (2016). ILONA: indoor localization and navigation system. *Journal of Location Based Services*. Recuperado de https://www.tandfonline.com/doi/abs/10.1080/17489725.2017.1283453

Wu, R.-H., Lee, Y.-H., Tseng, H.-W., Jan, Y.-G., & Chuang, M.-H. (2008). Study of characteristics of RSSI signal. En *Industrial Technology, 2008. ICIT 2008. IEEE International Conference on* (pp. 1–3). IEEE.

Xue, W., Qiu, W., Hua, X., & Yu, K. (2017). Improved Wi-Fi RSSI Measurement for Indoor Localization. *IEEE Sensors Journal*, *17*(7), 2224–2230. https://doi.org/10.1109/JSEN.2017.2660522

Yu, Z., & Guo, G. (2017). Improvement of Positioning Technology Based on RSSI in ZigBee Networks. *Wireless Personal Communications*, *95*(3), 1943–1962. https://doi.org/10.1007/s11277-016-3860-1

Zhang, J., & Zhou, P. (2018). Integrating low-resolution surveillance camera and smartphone inertial sensors for indoor positioning. En *2018 IEEE/ION Position, Location and Navigation Symposium (PLANS)* (pp. 410–416). https://doi.org/10.1109/PLANS.2018.8373408

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 13/09/2018
Aceitação/Acceptance: 11/11/2018

Un estudio comparativo de técnicas de minería de datos y aprendizaje máquina para la estimación del esfuerzo utilizando puntos de función

Christian Quesada López[1], Juan Murillo-Morera[1,2], Marcelo Jenkins[1]

{cristian.quesadalopez;marcelo.jenkins}@ucr.ac.cr, juan.murillo.morera@una.cr

[1] Universidad de Costa Rica, San Pedro, Costa Rica.

[2] Universidad Nacional de Costa Rica, Heredia, Costa Rica.

Pages: 595–609

Resumen: En los últimos años, una gran cantidad de técnicas de minería de datos y de aprendizaje máquina han sido utilizadas para la construcción de modelos de estimación del esfuerzo de desarrollo del software. La literatura ha reportado resultados inconsistentes acerca de la efectividad de los modelos dependiendo de los conjuntos de datos. En este estudio utilizamos un procedimiento automatizado para la comparación exhaustiva de modelos de estimación de esfuerzo y presentamos los resultados del análisis comparativo a partir de la combinación de un conjunto de técnicas de pre-procesamiento de datos, selección de atributos y algoritmos de aprendizaje aplicado a distintos sub conjuntos de datos del repositorio ISBSG. Los resultados indican que las técnicas con mejores resultados para los modelos basados en los puntos de función IFPUG-FPA fueron *LeastMedSq*, *SMOreg* y *GaussianProcesses* y para COSMIC-FFP fueron *SMOreg*, *MP5* y *AdditiveRegression*. Las técnicas que incorporan estrategias de regresión son las que mejores resultados obtienen. Asimismo, la combinación de técnicas de pre procesamiento y selección de atributos mejoran los resultados de exactitud. Los modelos de estimación evaluados alcanzaron valores en la exactitud estandarizada entre el 49.94% y 64.05% para IFPUG-FPA y entre el 80.30% y el 67.31% para COSMIC-FFP. Con nuestro procedimiento de evaluación es posible analizar la exactitud de distintos modelos de estimación, cuáles técnicas obtienen los mejores resultados de exactitud a partir de cada conjunto de datos y la combinación de técnicas que puede mejorar el desempeño de los modelos.

Palabras-clave: Estimación de esfuerzo de desarrollo del software; puntos de función, COSMIC-FFP IFPUG-FPA; minería de datos; aprendizaje máquina; estudio empírico.

A comparative study of data mining and machine learning techniques for software effort estimation using function points

Abstract: In recent years, a large number of data mining and machine learning techniques have been used for the construction of software development effort estimation models. The literature has reported inconsistent results about the

effectiveness of the models because of their dependence on the data sets. In this study, we use an automated procedure for the exhaustive comparison of effort estimation models, and present the results of a comparative analysis derived from the combination of a set of data pre-processing, attribute selection techniques and learning algorithms applied to different sub sets of data from the ISBSG repository. The results indicate that the techniques with the best results for models based on IFPUG-FPA function points were LeastMedSq, SMOreg, and GaussianProcesses, but for COSMIC-FFP were SMOreg, MP5, and AdditiveRegression. The techniques that incorporate regression strategies yield the best results obtained for these data sets. In addition, the combination of pre-processing techniques and attributes selection improve the accuracy results. The estimation models reached a standardized accuracy between 49.94% and 64.05% for IFPUG FPA and between 80.30% and 67.31% for COSMIC FFP. With our evaluation procedure, it is possible to analyze the accuracy of different estimation models, which techniques obtain the best accuracy results from each data set, as well as the combination of techniques that can improve the performance of the models.

***Keywords**:* Software effort estimation; function points COSMIC-FFP IFPUG-FPA; data mining; machine learning; empirical study.

1. Introducción

La necesidad de modelos de estimación de esfuerzo y productividad precisos es uno de los aspectos más importantes en la ingeniería del software (Abran, 2015). En los últimos años, la evaluación de estos modelos de estimación ha sido un área de investigación activa en la comunidad de la ingeniería del software (Dejaeger, Verbeke, Martens & Baesens, 2012). Sin embargo, la estimación del esfuerzo utilizando el tamaño funcional como predictor es aún un reto para los investigadores y profesionales (Jorgensen & Shepperd, 2007; Gencel, 2008). Una gran cantidad de técnicas de minería de datos y de aprendizaje máquina han sido utilizadas para la construcción de los modelos de estimación (Wen, Li, Lin, Hu & Huang, 2012; Dejaeger et al, 2012; Huang, Li & Xie, 2015), pero los resultados reportados acerca de la efectividad de las técnicas utilizadas para la construcción de los modelos han sido inconsistentes dada su dependencia con los conjuntos de datos utilizados como entrada (Dejaeger et al, 2012).Por lo tanto, son necesarios estudios empíricos que permitan realizar comparaciones sin sesgo (Kitchenham & Mendes, 2009).

La evaluación de combinaciones de técnicas para distintos conjuntos de datos es un campo de investigación abierto (Keung, Kocaguneli & Menzies, 2013). Uno de los principales desafíos en la evaluación de los modelos de estimación de esfuerzo se relaciona con la configuración de las técnicas, la selección de los datos y las métricas de evaluación, dado que una validación incompleta o inapropiada puede presentar sobreajuste o sub ajuste de los resultados. Un procedimiento de evaluación de modelos debe implementar las recomendaciones de los marcos de trabajo de estimación propuestos en la literatura para lograr un proceso sistemático y sin sesgo que permita la comparación de los resultados de la exactitud de las técnicas utilizadas para la estimación de esfuerzo (Song, Jia, Shepperd, Ying & Liu, 2011; Menzies & Shepperd, 2012; Shepperd & MacDonell, 2012; Song, Minku & Yao, 2013; Keung et al. 2013; Huang et al, 2015; Langdon, Dolado, Sarro & Harman, 2016). Las evaluaciones comparativas que permitan una selección

acertada de los modelos de estimación en un contexto determinado es clave para tomar decisiones acerca de su efectividad (Shepperd, 2007).

El objetivo del estudio es realizar una comparación exhaustiva de modelos de estimación de esfuerzo a partir de la combinación de un conjunto de técnicas de pre-procesamiento de datos, selección de atributos y algoritmos de aprendizaje aplicado a distintos sub conjuntos de datos del repositorio de proyectos del *International Software Benchmarking Standards Group* (ISBSG). El repositorio de proyectos ISBSG proporciona información de proyectos de software de múltiples compañías, países, industrias y ambientes de desarrollo. En su versión (R12), el repositorio contiene 6 mil proyectos con 105 características entre las cuales se encuentra el tamaño funcional, el esfuerzo de desarrollo y los factores de contexto (ISBSG, 2013).

Para realizar la evaluación, utilizamos un procedimiento automatizado que recibe como entrada las métricas de tamaño funcional y un conjunto de factores de contexto de proyectos de desarrollo de aplicaciones de negocios y calcula las métricas de exactitud. En total, se evalúan 600 distintos modelos de estimación producto de la combinación de las técnicas seleccionadas que se aplican a treinta grupos de datos escogidos a partir del repositorio ISBSG R12. Este estudio compara algoritmos de las familias de funciones y regresiones, bayesianos, árboles, reglas y redes. Los proyectos seleccionados son aplicaciones medidas bajo los métodos de medición de tamaño funcional IFPUG FPA y COSMIC FFP, ya que el método IFPUG FPA es el más utilizado en la industria (Fingerman, 2011) y el método COSMIC FFP es el de mayor crecimiento en adopción en los últimos años (Dumke & Abran, 2011).

El resto del artículo se estructura de la siguiente manera: la Sección 2 presenta los trabajos relacionados, la Sección 3 describe el procedimiento para la comparación de los modelos de estimación, la Sección 4 describe el estudio empírico, la Sección 5 discute los resultados y finalmente, la Sección 6 presenta las conclusiones y trabajo futuro.

2. Trabajo Relacionado

En el área de la estimación del esfuerzo de desarrollo del software se ha estudiado como mejorar la estabilidad de los resultados de los modelos de estimación, la selección y configuración de las técnicas utilizadas, y la selección de los conjuntos de datos para obtener resultados confiables en los estudios comparativos. La investigación sobre la estimación plantea distintos enfoques para la evaluación comparativa de modelos o combinación de estos, estrategias de línea base y una correcta selección de métricas de evaluación que permitan realizar comparaciones con el menor sesgo posible (Whigham, Owen & Macdonell, 2015; Saeed, Butt, Kazmi & Arif, 2018). Estos se dirigen, principalmente, a lograr la estabilidad de los resultados de evaluación a través de diferentes conjuntos de datos. Dejaeger et al. (2012) establecen que la selección y combinación de diferentes técnicas de minería de datos e inteligencia artificial permiten mejorar los resultados de la evaluación de los métodos de estimación para conjuntos de datos con características específicas. La motivación detrás de la evaluación de estas combinaciones es depender lo mínimo posible de las características de los datos cuando se realiza el entrenamiento de los modelos de estimación. Los enfoques de ingeniería de software basados en búsquedas (SBSE) pueden apoyar los procesos de selección de

técnicas y parámetros (Ferrucci, Harman & Sarro, 2014; Xia, Mathew, Shen & Menzies, 2018) dado que estos modelos dependen de las características de los datos y no existen reglas únicas para decidir cuál es la mejor combinación y configuración de las técnicas.

Oliveira, Braga, Lima & Cornélio (2010) realizaron investigaciones comparativas para evaluar el efecto de la configuración de parámetros en el desempeño de los modelos de estimación de esfuerzo construidos mediante técnicas de regresión. Del mismo modo, Song et al. (2013) realizaron comparaciones para evaluar el efecto de los de parámetros en el desempeño de los modelos construidos a partir de técnicas de aprendizaje máquina. Por su parte, Dejaegeret al. (2012) realizaron una comparación empírica a gran escala usando diferentes tipos de técnicas de inteligencia artificial y minería de datos para analizar los aspectos relacionados con la selección de las características de los conjuntos de datos, al igual que Liu, Xiao & Zhu. (2018) quienes desarrollaron una propuesta utilizando como principales técnicas la selección de atributos y el razonamiento basado en casos para proyectos con características semejantes. Keung et al. (2013) evaluaron noventa predictores con veinte conjuntos de datos y siete métricas de evaluación para determinar la estabilidad según la selección de los predictores. Finalmente, Huang et al. (2015) estudiaron un conjunto de técnicas de pre procesamiento y de aprendizaje máquina y analizaron las características de los conjuntos de datos y algoritmos de acuerdo a su efectividad.

Con el fin de establecer marcos de trabajo para la evaluación comparativa de modelos de estimación de esfuerzo, múltiples estudios en el área han propuesto enfoques para mejorar las comparaciones. Shepperd & MacDonell (2012) propusieron un marco de trabajo para evaluación de modelos de estimación que reduce la inconsistencia, contempla el uso de métricas que disminuyen el sesgo, y permite la agregación de resultados entre estudios y futuros meta análisis. Menzies & Shepperd (2012) y Keung et al. (2013) discuten el problema de la inestabilidad de los sistemas de estimación y analizan las fuentes de inestabilidad. Azhar, Riddle, Mendes, Mittas & Angelis (2013) proponen un marco de trabajo estadístico basado en un algoritmo de múltiples comparaciones para clasificar la efectividad de los modelos de estimación de esfuerzo. Whigham et al. (2015) presentaron un método para establecer una línea base de comparación para los modelos de estimación de esfuerzo basado en regresiones lineales múltiples. Dolado, Rodriguez, Harman, Langdon & Sarro (2016) proponen una medida para comparar modelos de estimación basada en un enfoque de pruebas del mínimo intervalo de equivalencia para el error absoluto y una estimación aleatoria como línea base de referencia. Langdon et al. (2016) proponen un método para establecer líneas base de selección aleatoria (*random guessing*) para la comparación de los modelos de estimación de esfuerzo. Finalmente, Lavazza & Morasca (2017) proponen un marco de trabajo para la construcción de indicadores de exactitud sobre modelos de estimación de esfuerzo. Estos indicadores son basados en los cuadrados y los valores absolutos de los residuos de los resultados obtenidos por los modelos.

3. Procedimiento automatizado para la comparación de modelos de estimación de esfuerzo

En esta sección se presenta el procedimiento automatizado para la comparación exhaustiva de modelos de estimación de esfuerzo. Este presenta una estrategia

sistemática de evaluación de modelos que permite realizar una comparación reduciendo el sesgo entre los distintos modelos basados en los métodos IFPUG FPA y COSMIC FFP. Este procedimiento presenta una extensión de los trabajos previos realizados en Quesada López & Jenkins (2016) y Murillo-Morera, Quesada-López, Castro-Herrera & Jenkins (2017) donde se han refinado las estrategias de construcción y evaluación de los modelos de estimación, el procesamiento de los conjuntos de datos, y se han agregado las funcionalidades de parametrización. La estrategia de evaluación se basa principalmente en las recomendaciones del marco de trabajo propuesto por Song et al. (2011) y trabajos en el área de modelos de estimación de esfuerzo tales como (Dejaeger et al., 2012; Langdon et al., 2016; Menzies, Yang, Mathew, Boehm & Hihn, 2017; Lavazza & Morasca, 2017). El procedimiento para la evaluación construye los modelos de estimación de esfuerzo y los evalúa para determinar cuál combinación de técnicas genera los mejores resultados a partir de las características de los conjuntos de datos.

A continuación, describimos la estrategia de evaluación que utiliza las mediciones de tamaño funcional y los factores de contexto como predictores, y obtiene como salida la estimación del esfuerzo de desarrollo de un proyecto de software. Con este trabajo definimos, implementamos y evaluamos un marco de trabajo exhaustivo de comparación de modelos de estimación de esfuerzo el cual tiene la flexibilidad para incorporar nuevas técnicas de pre procesamiento, selección de atributos, algoritmos de aprendizaje y su configuración de parámetros, que permite determinar los mejores modelos para un conjunto de datos específico.

La Figura 1 presenta los componentes de la estrategia de evaluación de modelos. Primero, el componente de evaluación permite analizar el desempeño de los modelos a partir de los datos históricos de proyectos y segundo, el componente de estimación de esfuerzo de desarrollo construye los modelos a partir de las técnicas que mejor resultado presentaron en el contexto y realiza la estimación de esfuerzo a partir de los nuevos datos. En el proceso de construcción de los modelos se combinan distintas técnicas para el pre procesamiento de datos, la selección de atributos o características y los algoritmos de aprendizaje. Cada combinación de los tres tipos de técnicas representa un esquema de aprendizaje que es comparado durante la aplicación de la estrategia de evaluación. A continuación describimos las fases de la estrategia:

Construcción de los modelos de estimación de esfuerzo: en esta fase se analizan los modelos de estimación construidos a partir de cada uno de los esquemas de aprendizaje para determinar su desempeño. Los modelos se construyen para cada uno de los esquemas de aprendizaje y los datos históricos de los proyectos desarrollados que funcionan como los conjuntos de entrenamiento. Cada esquema de aprendizaje se construye a partir de la selección de una técnica de pre procesamiento de datos (DP), una de selección de atributos (AS) y un algoritmo de aprendizaje (LA). Los parámetros de las técnicas o algoritmos de aprendizaje son configurados para determinar cuáles ofrecen un mejor resultado en las métricas de desempeño. El proceso de construcción contempla los siguientes aspectos:

a. Los datos históricos se dividen en dos conjuntos de datos, un conjunto de entrenamiento (90%) que permite la construcción de los modelos a partir de los distintos esquemas de aprendizaje y un conjunto de prueba (10%) que permite evaluar el desempeño de los modelos.

b. Para la separación de los conjuntos de datos se realiza un proceso de una ronda de validación cruzada MxN. Se realizan M repeticiones donde se aplica una aleatorización al conjunto de datos históricos y se generan N particiones de las cuales se asigna un 10% de las instancias para el conjunto de datos de prueba y un 90% de las instancias para el conjunto de datos de entrenamiento. La selección se realiza aleatoriamente donde se hace la separación de N partes, el modelo es construido a partir de las N-1 partes y la parte restante es utilizada para las pruebas. Esto se repite para las N rondas lo que permite que cada parte sea utilizada para el entrenamiento y las pruebas lo que minimiza el sesgo de la selección. El proceso se repite M veces aplicando una aleatorización para reducir el efecto del ordenamiento. El conjunto de datos de prueba es independiente del proceso de construcción de los modelos y es utilizado para evaluación.
c. El proceso de validación cruzada permite la construcción y evaluación de MxN modelos por cada conjunto de datos y para cada uno de los esquemas de aprendizaje. La estrategia de múltiples rondas de validación cruzada para cada uno de los conjuntos de datos se realiza para reducir la variabilidad de los resultados del desempeño de cada uno de los modelos construidos. Los resultados de cada una de las rondas son promediados para obtener el valor total del desempeño.
d. Los modelos se construyen y evalúan para cada uno de los esquemas de aprendizaje a partir del conjunto de datos de entrenamiento realizando tres pasos: Primero, se aplica la técnica de pre procesamiento de datos para la transformación, la creación de variables ficticias en escala [0, 1], eliminación de valores atípicos y el tratamiento de los valores faltantes. Segundo, se realiza la selección de atributos para la reducción de la dimensión del conjunto de datos aplicando métodos de filtrado o envoltura y evaluando el desempeño de cada uno de los subconjuntos seleccionados mediante una validación cruzada que permite determinar el subconjunto de atributos que muestra mejor desempeño. Tercero, se realiza la construcción del modelo utilizando las técnicas de aprendizaje de máquina y de regresión a partir del conjunto de datos reducido. En este caso se evalúa el desempeño de distintas configuraciones de parámetros mediante una validación cruzada que permite determinar la configuración con mejor desempeño.
e. Para realizar el proceso de validación, el conjunto de datos de prueba es tratado del mismo modo que el conjunto de entrenamiento y los valores reales y de estimación son calculados a partir de las métricas de desempeño. Cada uno de los resultados de evaluación se almacena para calcular los promedios y varianzas durante cada uno de los procesos de validación cruzada de MxN rondas.

El proceso de evaluación de esquemas de aprendizaje permite construir los modelos y determinar los mejores esquemas de aprendizaje que pueden ser utilizados para la estimación del esfuerzo de desarrollo para los nuevos datos.

Evaluación de modelos de estimación de esfuerzo*:* en la fase de evaluación y a partir de los resultados de la fase de construcción de modelos se selecciona el o los esquemas de aprendizaje para construir los modelos que permiten realizar las estimaciones de esfuerzo a partir de los nuevos datos y evaluar su desempeño. El

proceso obtiene la estimación de esfuerzo para los nuevos datos y reporta los resultados de estimación y desempeño. El proceso de evaluación contempla:

a. Para realizar la estimación de esfuerzo se seleccionan los mejores esquemas de aprendizaje y se construyen los modelos a partir dell conjunto de datos históricos.
b. El proceso de creación de los modelos de estimación se realiza de acuerdo a la estrategia descrita en la primera fase. Se pre procesan los datos, se realiza la selección de los mejores atributos para los nuevos datos y se selecciona la mejor configuración de parámetros de acuerdo a los resultados de los modelos analizados en la fase anterior.

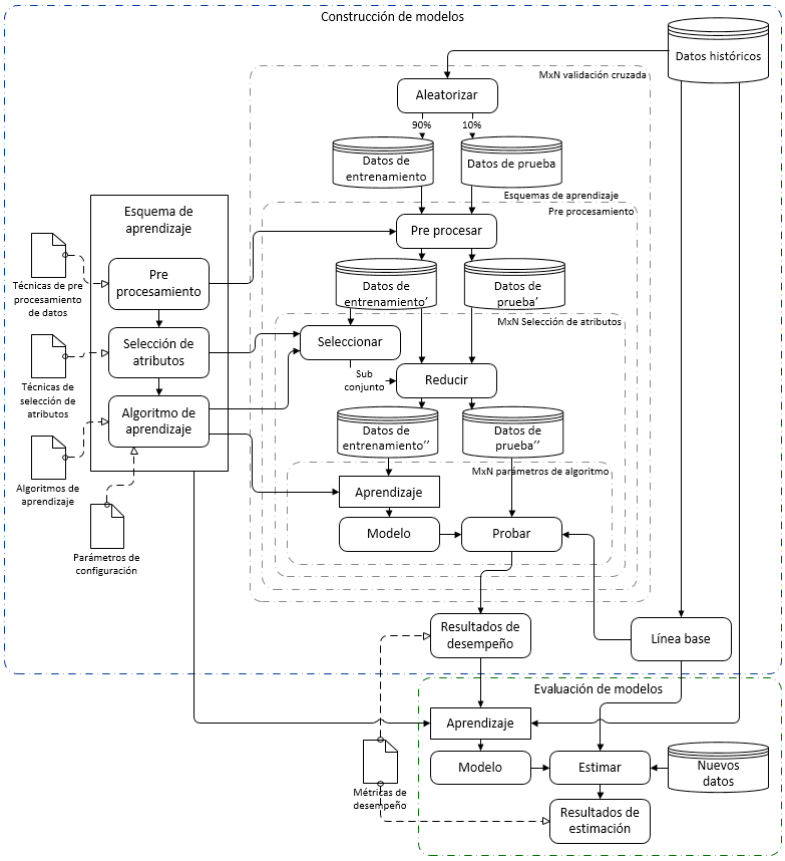


Figura 1 – Componentes de la estrategia de evaluación de modelos de estimación.

La estrategia busca la estabilización de los modelos y reducir el riesgo de que los mismos presenten problemas de sobre ajuste o sub ajuste que puede producir resultados no confiables sobre el desempeño. Asimismo, mediante la implementación de las estrategias de modelos de línea base, bajo el mismo conjunto de datos, es posible

realizar comparaciones más objetivas sobre el desempeño de los distintos modelos de estimación. Esto es posible porque la estrategia de evaluación plantea la comparación de los resultados de los modelos construidos para cada uno de los esquemas de aprendizaje contra los modelos construidos a partir de la línea base, ambos a partir de los mismos datos en el contexto evaluado. En este caso, se comparan los modelos generados por la estrategia exhaustiva contra los modelos basados en una estrategia de selección aleatoria de acuerdo al marco de trabajo de línea base propuesto por Langdon et al. (2016). Las técnicas específicas para cada uno de los componentes de los esquemas de aprendizaje se listan en La Sección 4.

La Figura 2 muestra los compontes de la herramienta prototipo para la evaluación de los modelos de estimación de esfuerzo. Esta incluye los siguientes componentes: (1) El componente de construcción de modelos se encarga de construir y evaluar el desempeño de los modelos creados a partir de cada uno de los esquemas de aprendizaje y los datos históricos. Este componente incluye el pre procesamiento de datos, la selección de atributos y la configuración de los parámetros de los algoritmos de aprendizaje. El proceso de evaluación construye los modelos y determina los mejores esquemas de aprendizaje que son utilizados para la estimación de los nuevos datos. (2) El componente de evaluación permite aplicar los mejores esquemas de aprendizaje para la construcción de los modelos utilizados en la estimación del esfuerzo de desarrollo del software. Este componente incluye las mejores técnicas de pre procesamiento, selección de atributos, configuración de atributos y algoritmo de aprendizaje que son aplicados en la estimación del esfuerzo de los nuevos datos. (3) El componente de reporte se encarga de realizar el proceso de presentación de los resultados de evaluación de los modelos y de la estimación del esfuerzo de desarrollo. Este componente incluye el reporte de los resultados de desempeño para cada uno de los esquemas de aprendizaje y el resultado del proceso de estimación.

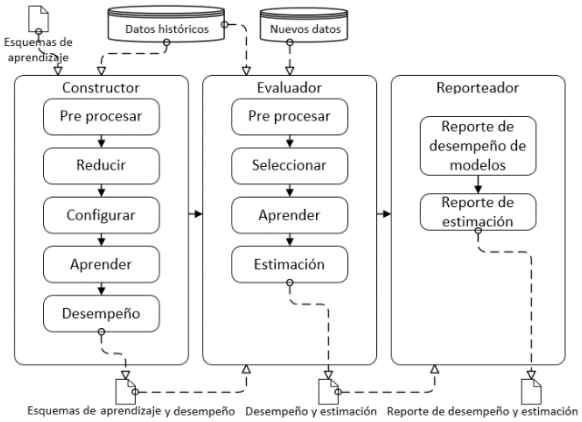


Figura 2 – Herramienta prototipo de evaluación de modelos.

La herramienta que implementa la estrategia de evaluación se desarrolla en Java. Esta utiliza como base las técnicas implementadas en las librerías de *Waikato Environment for Knowledge Analysis* (Witten, Frank, Hall & Pal, 2016).

4. Descripción del estudio empírico

La estructura del estudio empírico se realiza de acuerdo a los lineamientos por Wohlin, Runeson, Höst, Ohlsson, Regnell & Wesslén (2012). El estudio realiza un análisis empírico exhaustivo donde (a) se valida la estabilidad de los resultados del procedimiento para la construcción de modelos de estimación de esfuerzo basados en el tamaño funcional, (b) se estudian modelos simplificados de estimación de esfuerzo y (c) se reportan las técnicas que presentan los mejores resultados para los conjuntos de datos estudiados. Las preguntas de investigación analizadas son:

RQ1. ¿Cuál es la consistencia de los resultados de la exactitud obtenidas por el procedimiento de evaluación?

RQ2. ¿Cuál es la exactitud de los modelos construidos a partir de la combinación de técnicas para los conjuntos de datos?

El proceso general del estudio se lleva a cabo en dos fases. Primero, realizamos la selección de datos a partir del repositorio de proyectos ISBSG R12 para obtener los sub conjuntos de proyectos y variables utilizadas en el experimento. Segundo, se evalúan los modelos de estimación de esfuerzo y se determina el nivel de exactitud de las estimaciones. Para la recolección de datos se aplica la estrategia de evaluación descrita en la Sección 3.

Selección del conjunto de datos. Los conjuntos de datos (DS) utilizados en este estudio son aplicaciones de negocio del dominio de sistemas de información administrativo (MIS) medidas bajo los métodos de medición IFPUG FPA y COSMIC FFP. La Tabla 1 muestra los parámetros de selección utilizados. Se seleccionan los proyectos con calidad A o B, del método de medición IFPUG FPA y COSMIC FFP y del año 2005 o superior. Este estudio incorpora las variables de contexto de los proyectos tales como el tipo de desarrollo (TD), el tipo de lenguaje (TL), el lenguaje programación (LP), la plataforma de desarrollo (PD), la arquitectura (AR) y el tamaño del equipo de desarrollo (TE).

DS	Método	Calidad datos	Calidad conteo	Periodo	N	Grupo	Nivel de recursos	Dominio	Tipo
DS202	IFPUG FPA 4+	A	A|B	2005-2011	202	BA	1	BA	Multi
DS053	COSMIC FFP 2+	A|B	A|B	2008-2012	53	BA	1	BA	Multi

Tabla 1 – Conjuntos de datos seleccionados del ISBSG R12.

Cada uno de los conjuntos de datos (DS202 y DS53) es dividido manualmente para obtener distintos grupos que contengan subconjuntos de los componentes funcionales básicos (BFC) del tamaño funcional y variables de contexto que permitan realizar

comparaciones de efectividad de los modelos. Por ejemplo, conjuntos de datos que solo contengan el tamaño funcional total, sub conjuntos de BFC y/o variables de contexto. Los conjuntos de datos del método de medición IFPUG FPA (DS202) se dividen en 21 grupos y los conjuntos de datos del método de medición COSMIC FFP (DS53) se dividen en 9 grupos, todos con distintas variables predictoras de tamaño funcional y contexto.

Las variables de contexto (atributos nominales) son pre procesadas para codificarlas con valores entre 0 y 1 (*dummy coding*). No se aplica ningún tratamiento para los valores faltantes de las variables de contexto. Se aplican técnicas de distancia para la identificación de los valores atípicos.

Técnicas utilizadas para la construcción de modelos. En total se utilizan 2 técnicas de pre procesamiento de datos (DP), 5 técnicas de selección de atributos (AS) y 15 métodos de algoritmos de aprendizaje (LA) que son combinadas para la construcción de los modelos de estimación (DP×AS×LA). En el caso de la técnica de transformación de BoxCox se utilizan 6 parámetros para *lambda*. Las técnicas de DP permiten limpiar el ruido, eliminar datos faltantes y valores atípicos y transformar las variables. Las técnicas de AS permiten seleccionar los atributos predictores más apropiados que explican las variaciones para predecir la variable de salida. Finalmente, los LA son los modelos que permiten realizar las estimaciones. Las principales categorías de algoritmos utilizados son las funciones y regresiones, bayesianos, árboles, reglas y redes. Los algoritmos utilizados en la evaluación son los implementados por las librerías de WEKA para valores continuos (Witten, Frank, Hall & Pal, 2016). Se aplica el enfoque de *filter* para la reducción de las variables de los conjuntos de datos. Se selecciona este enfoque dado que, aunque no ofrece resultados de exactitud tan buenos como un enfoque de *wrapper*, es menos costoso computacionalmente.

Las técnicas de DP son la transformación logaritmo (LG) y *BoxCox* (BC), las técnicas de DS son *GeneticSearch* (GS), *BestFirst* (BF), *LinearForwardSelection* (LFS), *BackwardElimination* (BE), *ForwardSelection* (FS), y finalmente, los algoritmos LA son *GaussianProcesses* (GP), *LeastMedSq* (LMS), *LinearRegression* (LR), *MultilayerPerceptron* (MP), *RBFNetwork* (RBFN), *SMOreg* (SMO), *AdditiveRegression* (AR), *Bagging* (BGN), *ConjunctiveRule* (CR), *DecisionTable* (DT), *M5Rules* (M5R), *ZeroR* (ZR), *DecisionStump* (DS), *M5P* (M5P) y *REPTree* (RT). Las configuraciones de parámetros específicas utilizadas para cada una de las técnicas implementadas por los LA se seleccionan a partir de las recomendaciones presentadas por la documentación de las librerías WEKA (Witten et al., 2016).

Por cada uno de los sub conjuntos de datos se construye un modelo para cada uno de los esquemas de aprendizaje (DP×AS×LA). En esta evaluación, para cada uno de los sub conjuntos de datos se ejecuta un N_PASS=10 donde para cada pasada de la evaluación se calculan y promedian las métricas de exactitud. En total se realizan mil ejecuciones (N_PASS=10, M×N donde M=10 y N=10) por esquema de aprendizaje (DP×AS×LA).

Métricas de evaluación de exactitud. La exactitud de los modelos de estimación de esfuerzo se analiza basada en las métricas de exactitud: magnitud del error relativo (MRE), coeficiente de correlación de Spearman (SP), exactitud estandarizada (SA) y número de predicciones de los valores reales (Pred25) de acuerdo a las recomendaciones de estudios en el área.

Amenazas a la validez. El tamaño y las características limitadas de los conjuntos de datos utilizados en los análisis puede ser una amenaza a la validez interna. Los datos fueron filtrados para asegurar que solo se utilizó información de alta calidad. Fueron utilizadas técnicas, estrategias y métricas de evaluación recomendados en estudios previos del área. Los conjuntos de datos son del dominio de aplicaciones de negocios del ISBSG por lo que los resultados solo pueden generalizarse a este contexto y de acuerdo a los paramentos de selección de proyectos aplicados en el estudio.

5. Análisis de Resultados

En esta sección se reportan los resultados de la evaluación.

Consistencia de los resultados de la exactitud (RQ1). Para verificar la consistencia de la estrategia de evaluación se recolectan los resultados de las métricas de exactitud para cada uno de los modelos de estimación de esfuerzo mediante dos corridas independientes. Para cada uno de los grupos de los conjuntos de datos DS202 y DS053 se construyen los 600 modelos a partir de la combinación de las técnicas de DP (n=8), las técnicas de AS (n=5) y los LA (n=15) y se comparan los resultados de las métricas MRE, SP, SA y Pred25. Se aplica la prueba estadística *Wilcoxon signed Rank* para determinar la consistencia entre los resultados obtenidos por cada modelo de cada uno de los grupos de datos. Los resultados de las métricas de evaluación no pertenecen a una distribución normal, por lo que se aplica esta prueba no paramétrica pareada entre los resultados de las dos corridas.

Los resultados de las pruebas para cada grupo de datos y las métricas indican que no se encuentra una diferencia significativa entre los resultados de exactitud entre las dos corridas realizadas con la estrategia de evaluación. Esto implica que la estrategia de evaluación utilizada permite la obtención de resultados consistentes. En la práctica, la consistencia de la estrategia de evaluación permite realizar comparaciones exhaustivas y sin sesgo entre los modelos de estimación de esfuerzo para determinar el modelo que mejor desempeño produce en el contexto de evaluación. Del mismo modo, al permitir aplicar una estrategia sin sesgo de selección de proyectos es posible mejorar la confiablidad en cuanto al sobre ajuste en los resultados de los modelos. Por tanto, a partir de los resultados obtenidos en la evaluación de los modelos, es posible analizar la exactitud de los modelos de estimación de esfuerzo, las técnicas que obtienen los mejores resultados de exactitud y los aspectos que pueden impactar el desempeño.

Exactitud de los modelos de estimación de esfuerzo (RQ2). Para determinar los mejores modelos se utilizan los resultados de la métrica de exactitud estandarizada (SA). Los resultados de los 30 grupos de los dos conjuntos de datos analizados, 21 para los basados en el método IFPUG FPA (DS202) y 9 para los basados en el método COSMIC FFP (DS053) son seleccionados a partir de los mejores resultados de exactitud. Los modelos de estimación evaluados alcanzaron valores entre el 49.94% y el 64.05% para el método de medición IFPUG FPA y entre el 67.31% y el 80.30% para el método de medición COSMIC FFP con base en la exactitud estandarizada (SA). Esto significa que existen modelos basados en los proyectos del método IFPUG FPA que podrían superar la línea base de selección aleatoria hasta por un 54.05% y hasta un 80.30% en el caso del método COSMIC FFP. Los resultados obtenidos son similares a los reportados en

estudios previos en el área que han alcanzado valores que oscilan entre el 30% y 60% con valores de hasta el 80% (Whigham et al., 2015; Dolado et al., 2016).

Considerando los modelos con los mejores resultados (1% superior del total de 600 modelos por grupo), la mayor frecuencia de técnicas (LA) para el conjunto de datos del método IFPUG FPA son las de la regresión *LeastMedSq* que representa un 54.0% (n=68) del total y *SMOreg* que representa un 41.3% (n=52) del total. En el caso de los conjuntos de datos del método COSMIC FFP son las de *SMOreg* que representa un 29.6% (n=16) del total, *MP5* que representa un 29.6% (n=16) del total y *AdditiveRegression* que representa un 24.1% (n=13) del total. Todas estas técnicas (LA) incorporan estrategias de regresión para su implementación.

Las mejores combinaciones de técnicas de DP×AS×LA para los grupos del conjunto del método IFPUG FPA son las de BC×BF×LMS que representa un 10.3% (n=13) del total, BC×BE×SMO que representa un 9.5% (n=12), BC×BE×LMS que representa un 8.7% (n=11) y BC×LFS×LMS que representa un 7.9% (n=10). En el caso del método de medición COSMIC FFP fueron las de BC×LFS×SMO que representa un 7.4% (n=4) del total, BC×FS×AR que representa un 7.4% (n=4) y None×BF×M5P que representa un 5.6% (n=3). En general, los mejores resultados las obtuvieron las técnicas *LeastMedSq*, *SMOreg* y *GaussianProcesses* para los conjuntos de datos del método IFPUG FPA y *SMOreg*, *MP5* y *AdditiveRegression* para COSMIC FFP. Asimismo, los resultados indicaron que la combinación con técnicas de pre procesamiento y selección de atributos mejoraron los resultados de exactitud. Las combinaciones que mejores resultados obtuvieron para IFPUG FPA fueron las de BC×BF×LMS, BC×BE×SMO, BC×BE×LMS, y BC×LFS×LMS. En el caso de COSMIC FFP fueron las combinaciones de BC×LFS×SMO, BC×FS×AR y BF×M5P.

6. Conclusiones

Los resultados del estudio indican que el procedimiento de evaluación presentó resultados consistentes a través de múltiples corridas evitando el sobreajuste o sub ajuste de los modelos. Los modelos de estimación evaluados alcanzaron valores entre el 49.94% y el 64.05% para el método de medición IFPUG FPA y entre el 67.31% y el 80.30% para el método de medición COSMIC FFP con base en la exactitud estandarizada (SA).

Las técnicas de construcción de modelos (LA) que mejores resultados obtuvieron para los conjuntos de datos del método IFPUG FPA fueron: *LeastMedSq*, *SMOreg* y *GaussianProcesses*. En el caso de los conjuntos de datos del método COSMIC FFP los mejores resultados los obtuvieron las técnicas: *SMOreg*, *MP5* y *AdditiveRegression*. Los resultados mostraron que las técnicas que incorporaron estrategias de regresión para su implementación son las que mejores resultados obtuvieron para los conjuntos de datos. Asimismo, la combinación con técnicas de pre procesamiento y selección de atributos mejoran los resultados de exactitud.

La selección de las técnicas de construcción de modelos de estimación de esfuerzo depende de los conjuntos de datos para los cuales son creadas, por lo que el procedimiento de evaluación podría permitir a los profesionales identificar los modelos que mejor se ajustan a sus proyectos particulares.

Como trabajo futuro, se desea determinar el impacto de las distintas técnicas de pre procesamiento de datos, selección de atributos y algoritmos de aprendizaje en los resultados de exactitud de los modelos a partir de los parámetros de configuración. Asimismo, es de interés replicar este estudio utilizando el repositorio ISBSG R16.

Agradecimientos. Este estudio fue apoyado por la Universidad de Costa Rica No. 834-B8-A27. Nuestro agradecimiento al *International Software Benchmarking Standards Group*.

Referencias

Abran, A. (2015). Software project estimation: the fundamentals for providing high quality information to decision makers. John Wiley & Sons.

Azhar, D., Riddle, P., Mendes, E., Mittas, N., & Angelis, L. (2013). Using ensembles for web effort estimation. In 2013 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (pp. 173–182). IEEE.

Dejaeger, K., Verbeke, W., Martens, D., & Baesens, B. (2012). Data mining techniques for software effort estimation: a comparative study. IEEE transactions on software engineering, 38(2), 375–397. doi: 10.1109/TSE.2011.55

Dolado, J., Rodriguez, D., Harman, M., Langdon, W., & Sarro, F. (2016). Evaluation of estimation models using Minimum Interval of Equivalence. App Soft Comp, 49, 956–967.

Dumke, R., & Abran, A. (2011). COSMIC Function Points: Theory and Advanced Practices. CRC Press.

Ferrucci, F., Harman, M., & Sarro, F. (2014). Search-based software project management. In Software Project Management in a Changing World (pp. 373–399). Springer, Berlin.

Fingerman, S. (2011). Practical software project estimation; a toolkit for estimating software development effort & duration. Sci-Tech News (Vol. 65). McGraw Hill Professional.

Gencel, C. (2008). How to Use COSMIC Functional Size in Effort Estimation Models?. In Software Process and Product Measurement (pp. 196–207). Springer, Berlin, Heidelberg.

Huang, J., Li, Y. F., & Xie, M. (2015). An empirical analysis of data preprocessing for machine learning-based software cost estimation. Information and soft. Technology, 67, 108–127.

ISBSG. (2013). The ISBSG Development & Enhancement project data. ISBSG, Ed., R 12.

Jorgensen, M., & Shepperd, M. (2007). A Systematic Review of Software Development Cost Estimation Studies. IEEE Transactions on Software Engineering, 33(1), 33–53.

Keung, J., Kocaguneli, E., & Menzies, T. (2013). Finding conclusion stability for selecting the best effort predictor in software effort estimation. Automated Soft Eng, 20(4), 543–567.

Kitchenham, B., & Mendes, E. (2009). Why comparative effort prediction studies may be invalid. In Proceedings of the 5th international Conference on Predictor Models in Software Engineering (p. 4). ACM. doi: 10.1145/1540438.1540444

Langdon, W. B., Dolado, J., Sarro, F., & Harman, M. (2016). Exact mean absolute error of baseline predictor, MARP0. Information and Software Technology, 73, 16–18.

Lavazza, L., & Morasca, S. (2017, June). On the evaluation of effort estimation models. In Proceedings of the 21st International Conference on Evaluation and Assessment in Software Engineering (pp. 41–50). ACM.

Liu, Q., Xiao, J., & Zhu, H. (2018). Feature selection for software effort estimation with localized neighborhood mutual information. *Cluster Computing*, 1-9.

Menzies, T., & Shepperd, M. (2012). Special issue on repeatable results in software engineering prediction. doi: 10.1007/s10664-011-9193-5

Menzies, T., Yang, Y., Mathew, G., Boehm, B., & Hihn, J. (2017). Negative results for software effort estimation. Empirical Software Engineering, 22(5), 2658–2683.

Murillo-Morera, J., Quesada-López, C., Castro-Herrera, C., & Jenkins, M. (2017). A genetic algorithm based framework for software effort prediction. Journal of Software Engineering Research and Development, 5(1), 4.

Oliveira, A. L., Braga, P. L., Lima, R. M., & Cornélio, M. L. (2010). GA-based method for feature selection and parameters optimization for machine learning regression applied to software effort estimation. Information and Software Technology, 52(11), 1155–1166.

Quesada López, C., & Jenkins Coronas, M. (2016). Function point structure and applicability: A replicated study. Journal of Object Technology (15).

Saeed, A., Butt, W. H., Kazmi, F., & Arif, M. (2018, February). Survey of Software Development Effort Estimation Techniques. In Proceedings of the 2018 7th International Conference on Software and Computer Applications (pp. 82–86). ACM.

Shepperd, M., & MacDonell, S. (2012). Evaluating prediction systems in software project estimation. Information and Software Technology, 54(8), 820–827.

Song, Q., Jia, Z., Shepperd, M., Ying, S., & Liu, J. (2011). A general software defect-proneness prediction framework. IEEE Transactions on Software Engineering, 37(3), 356–370.

Song, L., Minku, L. L., & Yao, X. (2013). The impact of parameter tuning on software effort estimation using learning machines. In Proceedings of the 9th international conference on predictive models in software engineering (p. 9). ACM.

Wen, J., Li, S., Lin, Z., Hu, Y., & Huang, C. (2012). Systematic literature review of machine learning based software development effort estimation models. Information and Software Technology, 54(1), 41-59. doi: 10.1016/j.infsof.2011.09.002

Whigham, P. A., Owen, C. A., & Macdonell, S. G. (2015). A baseline model for software effort estimation. ACM Transactions on Software Engineering and Methodology, 24(3), 20.

Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., & Wesslén, A. (2012). Experimentation in software engineering. Springer Science & Business Media.

Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann.

Xia, T., Chen, J., Mathew, G., Shen, X., & Menzies, T. (2018). Why Software Effort Estimation Needs SBSE. arXiv preprint arXiv:1804.00626.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 09/10/2018
Aceitação/Acceptance: 03/11/2018

Pedagogia Transformadora e o Reaprender nas Humanidades Digitais

João Carlos Wiziack[1], Vitor Duarte dos Santos[2]

jwiziack2@gmail.com, vsantos@novaims.unl.pt

1 Escola de Comunicação e Artes da Universidade de São Paulo/Nova IMS, Av. Prof. Lúcio Martins Rodrigues, 443, Butantã, São Paulo, Brasil

2 IMS Information Management School, Campos de Campolide, 1070-312, Lisboa, Portugal

Pages: 610–625

Resumo: Ensinar e aprender são partes do mesmo processo, antes e principalmente no contexto da cultura digital contemporânea. Como mediadores, disseminadores do uso das tecnologias digitais nos ambientes de aprendizagem, cabe aos docentes a aplicação pedagógica adequada das tecnologias digitais e desenvolver novas competências para si e para os alunos frente as novas formas de aprender. Esse artigo sugere reflexões sobre o processo pedagógico educacional contemporâneo avaliando a necessidade de ajustes à velocidade de aprendizagem com as mudanças culturais na era da Sociedade da Informação. Com metodologia exploratória utiliza autores que oportunizam a reflexão sobre geração de ideias pedagógicas em sintonia com essa transformação social. O diálogo com as ideias levantadas, suportado por Grupos Focais com docentes especialistas, nos leva a conclusão que o atual ambiente de Ensino-Aprendizagem requer aprofundamento cognitivo face a interação computacional e novas competências para um dinâmico papel docente na interface tecno relacional de humanização contemporânea.

Palavras-chave: Humanidades digitais; Pedagogia transformadora; Educação significativa; Ecossistemas reticulares; Ensino- aprendizagem

Transforming Pedagogy and the Digital Humanities Relearning

Abstract: Teaching - learning are part of the same process, before and especially in the context of contemporary digital culture. As mediators and disseminators of digital technology effective use in learning environments, teachers should apply appropriate pedagogical use of digital technologies in school, as well as to seek development of new skills for themselves and for students, facing new forms of learn. The article suggests reflections on the contemporary educational pedagogical process, evaluating adjustments to the speed of learning, due to cultural changes of the Information Society era, dealing with exploratory methodology offering reflection on the generation of pedagogical ideas attending this social transformation. This dialogue among ideas was supported by Focal Groups with specialized teachers, which have led us to the conclusion that Teaching, and Learning environment requires cognitive deepening face to computational interaction and competences

development for a dynamic teaching role in the techno relational interface of contemporary humanization.

***Keywords**:* Digital Humanities; Transforming Pedagogy; Significant Education; Reticular Ecosystems; Teaching and Learning

1. Introdução

Em Humanidades Digitais, podemos entender que o termo "digital" modifica ou qualifica de alguma forma o substantivo estabelecendo uma relação direta com os efeitos da tecnologia digital. O aspecto digital se incorpora assim ao campo das possibilidades heurísticas de encontro de soluções para as questões por meio desses recursos tecnológicos. Assim nos alinhamos aos conceitos do texto *Manifesto for the Digital Humanities* como termo que identifica o conjunto de Ciências Humanas e Sociais, além de Artes e Letras (Dacos, 2010).

O processo de humanização desejado e necessário está profundamente afetado hoje pela forma como nos comunicamos e vivemos na era da sociedade da informação (Castells, 2000), do conhecimento (Hargreaves, 2003) e da aprendizagem (Hargreaves, 2003 e Pozzo, 2004), tempos em que se verifica grandes dificuldades em acompanhar o desenvolvimento veloz da tecnologia digital e ao mesmo tempo assimilar e converter informação em conhecimento.

A educação e o reaprender para a vida contemporânea e futura ocupam posição de centralidade de conteúdo nessa discussão, tendo em vista a condição de processo de humanização e prática social que insere o indivíduo no convívio com os coletivos. Contudo, o estudo tem seu foco em aprofundar o conhecimento dos desafios na adoção adequada de recursos tecnológicos de forma *pedagogicamente eficaz e atualizada*, em linha com a revolução tecnológica em constante avanço.

Ao justificarmos a reflexão crítica sugerida sobre pedagogia transformadora, buscamos em Saviani (2008, p.422) um conceito muito profundo e em linha com a problemática de nossa realidade ao dizer que " o trabalho educativo é o ato de produzir, direta e intencionalmente, em cada indivíduo singular, a humanidade que é produzida histórica e coletivamente pelo conjunto dos homens".

É uma visão holográfica em que um modelo intencional centrado na escola atua sobre os elementos culturais que precisam ser assimilados pelos indivíduos a fim de que eles se tornem humanos, por meio das melhores formas de atingimento desse objetivo. Interessante essa primeira apreensão de como a tecnologia pode ajudar a *nos tornarmos humanos*.

Observa-se que essas ideias se referem a conceitos pedagógicos sobre educação que valem para as várias técnicas e metodologias de aprendizagem atualmente aplicadas, independentemente do caráter político-ideológico do contexto em que nasceram, bastando para tanto o encontro da intencionalidade política-filosófica com as melhores alternativas instrucionais oferecidas pelo novo contexto tecno relacional que vivemos.

Isto significa que da mesma forma como foi antes e que continuará persistindo em direção ao futuro a pedagogia será a formatação de conceitos, ideias e instrumentos

que responderão da melhor forma a pergunta sobre que tipo de humanidade queremos construir e, para tanto, que educação precisamos planejar para o tipo de cidadãos que desejamos ser. (Saviani,2008; Giroux,1999)

O artigo valida por meio de autores e especialistas críticos e atuantes na educação brasileira e internacional as questões que se colocam na busca de nova identidade profissional para o docente, bem como apresenta um modelo alternativo para melhor entendimento e aplicação do uso tecnológico nos processos de ensino e aprendizagem discutido e avaliado por dois Grupos Focais com especialistas educacionais.

Nas sessões que se seguem discute-se o lugar da Pedagogia no processo de humanização sob influência do ambiente da sociedade da informação, do conhecimento e da aprendizagem caracterizados pelo comportamento informacional da cognição integrada e das novas competências já em exercício na era da tecnologia digital.

Finalmente essas discussões e contribuições obtidas com os Grupos Focais estão apresentadas nas conclusões finais registrando os principais aspectos de atenção e valorização do papel da Pedagogia nas Humanidades Digitais frente as novidades tecnológicas e interação computacional que se apresentam como novos elementos a se incorporar aos conteúdos de aprendizagem.

2. O lugar da Pedagogia nas Humanidades

De acordo com Hirsch (2012), a pedagogia era vista, como um pensamento tardio e entre parêntesis no contexto das pesquisas em Digital Humanities e que seria preciso leva-la além, tirando-a da marginalização e exclusão para coloca-la à frente, o que de fato somente aconteceu a partir de 2011/2012.

O autor explica que a quase exclusão da pedagogia na agenda acadêmica se deveu ao grande antagonismo existente entre ensino e pesquisa com nítida vantagem e melhores chances para a obtenção de fundos para a pesquisa em humanidades, em função do grande enfoque dado para as publicações acadêmicas.

De acordo com Giroux (1994) citado em Hirsch, (2012, p.6) a "importância da pedagogia não recai simplesmente sobre como os estudantes aprendem, mas também sobre como os educadores (no amplo sentido do termo) constroem as posições ideológicas e políticas das quais eles falam." Esse é um ponto cuja discussão consideramos fundamental para o entendimento do significado da Pedagogia no contexto educacional do ensino e aprendizagem da forma como apresentamos.

Saviani (2005) explicita as relações *entre educação e seus condicionamentos sociais* [grifo nosso], entendendo a própria educação como uma das modalidades dessa prática, ou seja humanizar é integrar o homem à sociedade em linha com suas conformações definidas e aceitas no contrato social.

Freire (2003) tambem exorta sobre o pensar na perspectiva política sobre a forma de estar no mundo conformando o homem, desintegrado da realidade, sufocado pelo pensamento hegemônico impossibilitado de transpor os limites que lhes são impostos pelo próprio mundo. O analfabeto da era digital também pode ser considerado hoje um ser imerso no mundo e vivendo sem consciência de sua realidade.

Para Freire,(1982, pp.27-28) a educação" é a resposta da finitude da infinitude e é possível para o homem porque este é inacabado e sabe-se inacabado". Interessante entender o saber-se inacabado como consciência e conhecimento de limites e não sentimento de fraqueza.

Essa visão deixa claro que o alinhamento pedagógico com o momento tecnológico da contemporaneidade passa pelo significado de expansão e superação de nossos limites com o desenvolvimento de novas competências para a sustentabilidade do processo civilizatório.

As tecnologias digitais parecem conspirar junto com autores nascidos muito antes delas ao permitirem que o ambiente democrático , desterritorializado (Levy 1977), e acessível à informação em tempo real determinem uma pedagogia para o aprender e viver na nova realidade. Apenas reforçando o conceito crítico invocado, "a prática social põe-se , portanto, como o ponto de partida e o ponto de chegada da prática educativa" (Saviani, 2008, p. 422), em que ensinar e aprender também marcam o encontro entre professor e aluno para a mesma caminhada

Segundo Hirsch (2012), graças a essa visão crítica sobre pedagogia, com a especialização e desenvolvimento tecnológico foi possível evoluir, recuperando e incorporando as próprias descobertas e histórias que melhor esclarecem as questões passadas e assim permitindo começar a contá-las de outras formas.[1]

Diante disso, prossegue Hirsch, o desenvolvimento curricular, o crescente apoio de instituições e organizações profissionais além de agências de financiamento , sem dúvida reformataram a paisagem da pedagogia nas Humanidades Digitais (p.10). Hirsch (2012, p.13) afirma também, que a Pedagogia é o melhor caminho de estabilização de um campo de conhecimento em função de seus cânones que representam os fundamentos de qualquer disciplina ou seja, "aquilo que é ensinado",em linha com o aforismo cunhado por Roland Barthes. Os cânones têm um papel fundamental em reformatar os campos de estudo e que considera prudente que se pense criticamente nas implicações institucionais e consequências políticas ao ensinar Humanidades Digitais.

Na lógica das redes, (Castells, 2000) onde várias vozes se juntam para buscar, alterar e reconfigurar a informação, como melhorar nossa capacidade crítica permitindo construir uma cultura aprendente que possibilite a construção de caminhos para a sociedade do conhecimento? Como fazer do docente um monitor e desenvolvedor das melhores posturas nessa direção constituindo-se em protagonista dessa pedagogia crítica?

3. Informação, Conhecimento e Aprendizagem

A Sociedade contemporânea decididamente tem uma nova maneira de ser e de se comportar como similaridades quase que globais em muitos aspectos. De acordo com Castells (2000, p. 25), "a tecnologia não determina a sociedade e nem a sociedade escreve o curso da transformação tecnológica", mas tudo é resultado de uma interatividade muito

[1] Vide, a título de exemplo, a riqueza oferecida por um site de extensa variedade de informações sobre arte no link http://www.etty.edu/research/publications/all.html

complexa entre criatividade e iniciativa empreendedora, inovação e aplicações por parte da sociedade e com certeza, com as influências e consequências da vida em rede

Nesse modelo, cada vez mais o comportamento social vem sofrendo transformações na medida em que a sociedade planetária vai se embrenhando na vida em rede (Castells, 2000), claramente afetando o processo de aprendizagem pelo contato virtual e conhecimento em tempo real de fatos e informações. McLuhan já observava sobre a função social das mídias que "as sociedades sempre foram influenciadas mais pela natureza dos media, através dos quais os homens se comunicam, do que pelo conteúdo da comunicação" (McLuhan,2007, p.22). Tal como com a máquina, não foi ela que nos transformou, mas sim o que se fez com ela e todo seu conteúdo.

Da mesma forma, o "conteúdo da escrita é a fala, assim como a palavra escrita é o conteúdo da imprensa e a palavra impressa é o conteúdo do telégrafo", (Mcluhan,2007, p. 22). Esclarece o autor que a mensagem de qualquer meio ou tecnologia é o que estas provocam nas coisas humanas quaisquer sejam seus impactos. Assim, é justa uma reflexão sobre os inúmeros efeitos que tecnologia digital vem causando na forma de pesquisar, aprender e construir conhecimentos, abrindo novas perspectivas sobre releituras de nossa trajetória bem como permitindo avançar pelas fronteiras onde estacionamos.

A tecnologia simplesmente acelera uma ideia que é posta em relevo desde a década de 90 que já destacava a "integração de diferentes tipos de conhecimento - habilidades práticas e cognitivas, conhecimento factual e conceitual, motivação, valores, atitudes, emoções, etc. – como um aspecto essencial dos enfoques baseados em competências" (Coll, 1999, p.36).[2]

No contexto dessa nova complexidade educacional em que a adoção das TIC, a interdisciplinaridade entre o saber pedagógico, o conteúdo do conhecimento, as habilidades tecnológicas surgem como motivações docentes influenciando novos comportamentos destaca-se ainda outro grande desafio para a educação contemporânea: *a competência de pensar e agir com criticidade*.

A sociedade somente se transforma se as pessoas individualmente se transformam e para tanto os educadores precisam encarar as escolas como estruturas sociais, fazendo-as funcionar para "educar as potencialidades que as pessoas têm, para pensar, para agir, para ser indivíduos e para ser capazes de compreender os limites de seus compromissos ideológicos"(Giroux, 1999, p. 21).

De acordo com Silva, (2008), as Ciências da Informação e Comunicação como ciência social que trabalha com o fenômeno info-comunicacional centra no comportamento o resultado perceptível e cognoscível da competência essencial desta ciência. Segundo o autor, dominar informática e internet não garante o acesso, assimilação e uso da informação com enriquecimento próprio. A literacia informacional vai além da leitura,

[2] Cesar Coll, Doutor em Psicologia e Catedrático de Psicologia Evolutiva e Educativa da Universidade de Barcelona, atuou como consultor do MEC- Ministério da Educação e Cultura do Brasil na formatação dos "Parâmetros Curriculares Nacionais "para adequações aos ditames da UNESCO sobre Educação para o século XXI.

escrita e oralidade ampliando para a necessidade de novas habilidades cognitivas para compreensão e uso eficaz das ferramentas computacionais.

A matriz dessa compreensão parte da caracterização dada por Castells (2002) estudada em (Coutinho e Lisboa ,2011) sobre a Sociedade da Informação em que:

- Informação é a matéria-prima – Existe uma relação simbiótica entre a tecnologia e a informação. São auto complementares.
- Poder de influência e penetração dos meios tecnológicos na vida social, econômica e política da sociedade
- Lógica de redes – característica predominante deste novo modelo de sociedade, que facilita a interação entre as pessoas ampliada para todos os tipos de processos e organizações, graças as recentes tecnologias da informação.
- Flexibilidade – poder de reconfigurar, alterar e reorganizar as informações.
- Convergência de tecnologias específicas para um sistema altamente integrado – O *contínuo processo de convergência* entre os diferentes campos tecnológicos resulta da sua *lógica comum* de produção da informação, onde todos os utilizadores podem *contribuir*, (grifos nossos) exercendo um papel ativo na produção deste conhecimento.

Para Coutinho e Lisboa (2011), essas características, das quais destacamos a convergência e integração, se ligam à nova democracia do saber, aos espaços de compartilhamento que permitem trânsito sem barreiras comerciais e de comunicação. Afirmam ainda que o "importante nesta sociedade não está na tecnologia em si, mas nas possibilidades de interações que elas proporcionam através da cultura digital" (Coutinho e Lisboa, 2011, p.8)

Essa compreensão, se faz com a integração das características descritas nesse cenário, impulsionadas pela internet e tecnologias digitais definindo para as autoras o novo paradigma social constituído pela Sociedade da Informação. Não há um tempo e depois outro. Na vida em rede os ecossistemas se interligam, as características, possibilidades presentes e futuras se misturam e representam grandes desafios para a convivência planetária e notadamente para o encontro da mais adequada pedagogia educacional para o desenvolvimento das Humanidades.

A produção do conhecimento, base para a geração das riquezas e recursos econômicos de sustentabilidade dependem, segundo Hargreaves, da capacidade de adaptação às mudanças contínuas e ao mesmo tempo da capacidade de aprender de forma autônoma e coletiva. "A Sociedade do conhecimento é uma sociedade da aprendizagem" (Hargreaves, 2003, p.37)

4. O Comportamento Informacional e Cognição Integrada de Novas Competências

As pesquisas realizadas pela OCDE a cada dois anos vêm mostrando uma posição de ranking entre os países privilegiando a pedagogia aplicada com uso tecnológico além de comentários de estímulo às experiências bem-sucedidas. No relatório de 2000 já se apresentava a visão que na sociedade aprendente, processar e administrar a informação, bem como adaptação às mudanças representam sucesso na melhoria das condições dos países (OCDE, 2000). No contexto dessas avaliações e opiniões plurais como, Fisher

(2000), Fabela (2005), Veen e Jacobs (2005) e Veen e Vrakking, (2009), Coutinho e Lisboa (2011) traçaram um perfil de desafios, premissas e competências necessárias para a Educação de hoje e do futuro, conforme resumido abaixo:

1. O aprender ao longo da vida é fundamental para a construção de uma nova sociedade. A sociedade da informação e do conhecimento depende dessa continuidade de aprendizagem (Fisher, 2000).
2. Uma cultura aprendente é aquela que constrói conhecimento de forma coletiva, contínua, horizontal dentro de um processo em que todos os sujeitos individualmente e coletivamente tem voz atuando num sistema de auto formação, (Fabela, 2005).

Essas premissas sugerem alguns elementos condicionantes como: confiança no aprendizado do aluno e relevância percebida, fazendo a ligação com conhecimentos prévios caracterizando aprendizagem significativa; priorização de talentos estimulando aprendizagem e tirando o foco na descoberta de pontos fracos; desafios que ajudem o aluno a estabelecer uma relação com o mundo em que vive e preferir a imersão ao invés dos conteúdos passo a passo; despertar a paixão como chave da motivação; auto regulação de professores e alunos na corresponsabilidade pela aprendizagem com respeito e solidariedade como valores e princípios essenciais de cidadania plena e responsável; espírito empreendedor e criatividade além da transformação do currículo.

Em Silva (2008), observa-se também um importante processo oriundo da evolução tecnológica digital em que "o emergente paradigma pós custodial, informacional e científico" da informação e não mais de sua simples guarda documental provoca a renovação processual da "origem, coleta, organização e armazenamento, recuperação, transmissão e a transformação da informação "Silva (2008, p.27).

Uma nova perspectiva é adicionada ao comportamento informacional mais abrangente cognitivamente e voltado não só aos instrumentos de busca do utilizador e seu atendimento como também à pesquisa científica, requerendo constantes aperfeiçoamentos em termos de processos e competências. Novas formas de inovar, fazer, pesquisar e ensinar, aliando pedagogicamente competências de conteúdo do objeto de estudo com competências tecnológicas, além de significativo avanço da aplicação pedagógica nas Humanidades Digitais (Schulman 86, Freire 2003, Khoeler & Mishra, 2007 e Hirsch 2012).

4.1. O Contexto das Competências Cognitivas e Motivacionais

Observa-se, portanto, na literatura acadêmica referenciada, grande variedade de competências necessárias para a vida contemporânea já no limiar de sua terceira década, o que torna mais difícil a tarefa de superar o atraso cultural presente no globo comparado à velocidade do processo inovativo e de uso das tecnologias digitais.

O cenário que envolve o Ensino e Aprendizagem é cada vez mais complexo puxado pelas tecnologias digitais provocando transformações comunicacionais em nossa vida contemporânea sendo tecida pelo imbricamento dos ecossistemas reticulares, com novas formas de relacionamento que por sua vez produzem transformações em nossa forma de viver e se relacionar. Gradativamente estamos sendo levados a uma nova forma de habitar a biosfera.

O processo de humanização precisa ser renovado e sair das estatísticas de atraso educacional com mais velocidade e eficácia. Com base nas ideias pedagógicas levantadas, esse estudo propõe uma alternativa de adoção integrada de competências cognitivas e motivacionais para uso de tecnologia nos processos educacionais de forma eficaz. Essa alternativa é baseada na reconceituação do modelo de conhecimentos integrados de tecnologia, pedagogia e conteúdos disciplinares – TPACK, technological pedagogical content knowledge de Kholer & Mishra (2007)[3] atualizando-o para o conceito de competências cognitivas integradas acrescidas da componente motivacional e de auto eficácia docente de Bandura (1977), mudanças que seguramente robustecem a construção anterior.

O chamado DCCM -Contexto das Competências Cognitivas e Motivacionais é um construto que representa o design de uma ideia alternativa sobre uma forma de se entender e aplicar de maneira pedagogicamente concebida como mais eficaz em processos de ensino e aprendizagem, ilustrado na figura 1 a seguir:

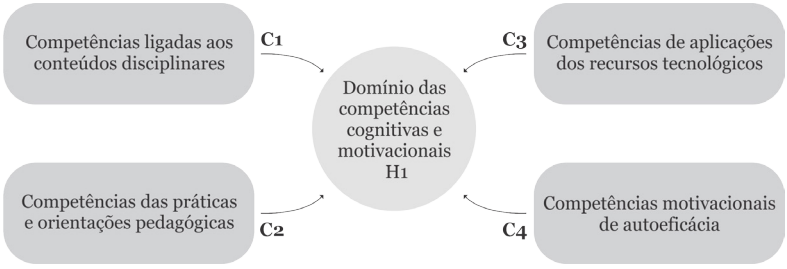


Figura 1 – DCCM: Domínio das Competências Cognitivas Integradas e Motivacionais

É importante para o entendimento do conteúdo pedagógico transformador desse estudo que o conceito de competências nascido no século do conhecimento surge como expressão de autonomia invertendo, no mundo do trabalho os valores pregados pela tradição taylorista (Zarifian, 2003; Coll, 2007).

A chave da compreensão do valor prático do conceito no contexto de seu surgimento é o entendimento de competência como um *saber agir* visto como um comportamento que vai além do conhecimento, pois refere-se ao saber e querer fazer, transformando o conhecimento em atitude e ação explicitada na prática autônoma (Zarifian, 2003; Fleury, 2001).

O estudo resultou na concepçãp metodológica de um design conclusivo sobre as relações de causa e efeito verificadas entre Competências Cognitivas e Motivacionais responsáveis pelo domínio atribuído ao melhor desepenho docente esperado nas atividades de ensino e aprendizagem com o uso da tecnologia de forma pedagogicamente eficaz, na figura acima representadadas por C1,C2,C3 e C4.

[3] Modelo apresentado como extensão ao TPACK framework de Khoeler e Mishra (2007), na CISTI 2016 e agora acrescido da componente motivacional docente de auto eficácia de Bandura (1977), com maior enfoque à Pedagogia na formação do construto

A perspectiva de agência humana vista em Bandura (2008, p.15) apresenta dois aspectos fundamentais para a consideração da possibilidade de autonomia e criação de automotivação na compreensão das formas de desenvolvimento de competências cognitivas e superação de obstáculos de aprendizagem: a intencionalidade em que "as pessoas formam intenções que incluem planos e estratégias de ação para realiza-las." O segundo aspecto está ligado à capacidade de antecipação de situações criando objetivos para si e até mesmo prevendo prováveis resultados, conceito fundamental no construto por caracterizar a relação afetiva do querer adicionado ao conhecimento e habilidades.

Para a explicitação, discussão e validação do DCCM foi utilizada a técnica qualitativa do DSR – Design Science Research, tendo em vista sua grande aplicabilidade e sucesso em Sistemas de Informação para decisão verificado em Hevner et al (2004) em amplo estudo metodológico voltado à aplicação do DSR no âmbito das questões de pesquisa nessa área. Verifica-se similaridades conceituais de grande apoio à aplicabilidade na presente pesquisa tendo em vista que a decisão final produzida dependerá de desempenhos e motivações de pessoas submetidas a um determinado modelo organizacional e institucional tomando decisões e fazendo escolhas alternativas de ideias de forma democrática e autônoma como em projetos de SI.

O Domínio proposto somente é válidado com a força desses indicadores de medida – as 4 competências formadoras do domínio - que o refletem guardando uma correlação entre si e apontados como fator causal para explicar e confirmar essa hipótese.

A discussão e validação de DCCM se deu pela aplicação da técnica qualitativa com 2 Grupos Focais totalizando 19 especialistas efetivamente atuantes em docência bem como pesquisas e gestão educacional. A opção pela técnica de Grupos Focais e pela escolha de especialistas objetivou criticidade e rigor na interpretação das condições docentes para o ensino e aprendizagem dentro de um contexto novo e desafiante em termos de utilização tecnológica em atividades de conteúdos diferenciados tendo ainda que incluir em suas bagagens o domínio dessas novas competências.

4.2. Processo de avaliação

O trabalho com os dois grupos focais foi organizado em 3 etapas conforme descritas a seguir:

Etapa 1: Preparação prévia dos participantes por meio da leitura de um documento contendo objetivos, forma e critérios de discussão, bem como um resumo das abordagens do referencial teórico relativo às competências constantes do construto DCCM apresentado na figura 1.

Etapa 2: Discussão nos grupos focais das 4 questões orientativas abertas relativas às variáveis c1, c2, c3 e c4 do construto DCCM visando avaliar o grau de aderência e significância das mesmas na formação do modelo proposto:

C 1: Competências ligadas aos conteúdos disciplinares

Questão orientativa discutida : *Qual é o entendimento sobre a importância e significado das competências referentes aos conteúdos disciplinares para a*

função docente como parte integrante do Domínio das competências cognitivas e motivacionais?

Contribuições dos grupos focais: O Professor deve conhecer adequadamente os conteúdos de sua disciplina mantendo-se constantemente atualizado a fim de poder dar conta da grade curricular estabelecida na disciplina, orientar pesquisas de novas informações e dúvidas de seus alunos, bem como saber estabelecer relações com as práticas relacionadas com outras disciplinas do currículo. Conhecimento profundo e atualizado da disciplina no seu contexto é fator chave para a docência.

C2: Competências de aplicações dos recursos tecnológicos

Questão orientativa discutida: Baseados em suas práticas e experiências em curso, *qual é o impacto da tecnologia na função docente e que sugestões podem ser aplicadas para sua adoção?*

Contribuições dos grupos focais: Os benefícios são inquestionáveis principalmente quando os docentes percebem os efeitos em seu trabalho. Entretanto, As TIC mudam constantemente e fica muito difícil acompanhá-las, principalmente para os das gerações mais velhas. As competências tecnológicas devem fazer parte do curríiculo formativo dos docentes de forma associada ao aprendizado pedagógico, e dessa forma poderão ser incorporadas à prática do dia a dia. Mesmo assim devem ser suportadas por especialistas no interior das unidades escolares até que a assimilação das habilidades instrucionais e tecnológicas atinjam nível de proficiência na cultura digital escolar.

O aprendizado das aplicações tecnológicas de forma integrada deve ser suportado por um grande esforço inicial do MEC- Ministério da Educação e Cultura pela oferta de recursos instrucionais com aplicações tecnológicas à distância cobrindo de forma planejada todos os conteúdos curriculares e fazendo uso de todas as boas iniciativas estabelecidas no mercado externo já disponível, bem como nos laboratórios ou centros de inovação desenvolvidos no interior das escolas

C3: Competências das práticas e orientações Pedagógicas

Questão orientativa discutida: *Como se avalia o conhecimento e notadamente a prática pedagógica em sala de aula e sua importância na integração com aplicação tecnológica?*

Contribuições dos grupos focais: De fato, há uma visão que o que interessa é saber e dominar os conteúdos da disciplina lecionada. Na maioria das vezes as práticas são desintegradas até mesmo dentro da própria grade curricular. Não se busca as melhores formas para representação de ideias, com melhores analogias, vivências e formulações do objeto como se deveria, porém, por falta da necessária competência que precisa ser mais bem desenvolvida. Muitos professores apenas transportam o uso social das TIC para dentro da sala de aula e a maioria reconhece que os alunos têm maior domínio delas. O avaliadores reforçam a premissa que o conhecimento pedagógico refere-se à forma particular de trasnmitir os conteúdos e que se constitui concretamente na competência de ensinar. As novas tecnologias de informação oferecem grande variedade de possibilidades de representações, porém, exigem sintonia com a correta aplicação

pedagógica, e isto precisa ser bem desenvolvido pois exige habilidades além dos conhecimentos. Não basta levar as TIC da vida social para a escola sem essa devida sintonia.

C4: Competências motivacionais de autoeficácia

Questão orientativa discutida: *Como equilibrar motivação para a docência em meio a tantas dificuldades e ainda mais tendo a tecnologia como mais um fator de desafio?*

Contribuições dos grupos focais: Todos reconhecem que o trabalho docente é, de fato, um composto de diferentes histórias de vida de pessoas que, além dos conteúdos disciplinares necessitam do autocontrole para conhecer e ajudar o aluno, grande objeto de suas missões. Acreditam na autoeficácia como essa capacidade humana de motivação e superação de obstáculos e enaltecem a necessidade de maior reconhecimento social tendo em vista o entorno e desafios que o professor enfrenta em seu trabalho. Essa motivação é também uma resultante das condições pessoais do professor, do aluno, da família e da escola, podendo contribuir, facilitar ou não a consecução dos objetivos desejados.

Etapa 3: Embora a avaliação, conforme escopo do artigo, teve foco na tecnica qualitativa com Grupos Focais, na terceira etapa da reunião optou-se em atribuir uma nota (de 1 a 4), menor para maior, para os indicadores C1, C2, C3 e C4 do construto DCCM – Domínio das competências cognitivas e motivacionais, expresso na figura 1, como forma de confirmação da força contributiva na explicação e justificativas das variáveis de medida sobre o domínio avaliado.

O resultado desse exercício mostrou que os grupos concordaram que cada variável de medida, ou indicador formativo fosse então considerada um domínio de competência recebendo uma redação definitiva na formação do Construto denominado Domínio das Competências Cognitivas Integradas e Motivacionais, ilustrado na figura 2 abaixo

Variável Latente	Indicador	Definições dos Indicadores
H1 - DCCM	C1	Domínio dos conteúdos é essencial para aprendizagem integrada e escolha de recursos pedagógicos
	C2	Domínio conceitual e prático sobre função pedagógica na aprendizagem
	C3	Domínio dos recursos tecnológicos instrucionais e computacionais com sentido pedagógico
	C4	Competências relacionadas ao self e ao contexto das pessoas como agências e motivadas pelo auto eficácia

Figura 2 – Indicadores formativos do construto DCCM

Além de fortemente concordados pelos avaliadores nas discussões dos 2 grupos focais, o exercício quantitativo confirmou a relação causal e formativa dos domínios Ci, C2, C3

e C4 na composição do Construto DCCM, tendo em vista que a média das variáveis é superior a nota 3, sendo que a máxima seria igual a 4. Dessa maneira, considerando que o desvio padrão amostral ao redor da média é menor do que 1 para todas as variáveis, tem-se que a amplitude (diferença entre o limite superior e o limite inferior do intervalo de confiança para a média amostral) do intervalo com 95% de confiança é inferior a 1. Além disso, o fator de carga é superior a 0,7, sendo, 0,88 de média, e, portanto, demonstrando a grande importância dada pelos especialistas a esses indicadores na construção do construto.

Variável	Estatística		
	Média	Desv. Padrão	Load
DCCM	3,75	0,4136	0,9375
c1	3,85	0,3664	0,9625
c2	3,75	0,4443	0,9375
c3	3,05	0,7592	0,7625
c4	3,45	0,7592	0,8625

Tabela 1 – H1/DCCM -Domínio das Competências Cognitivas Integradas e Motivacionais

A tabela 1 acima apenas ilustra essa análie estatistica cujos dados e cálculos foram retidos pelos autores em função das limitações de espaços e da relevância secundária no contexto da técnica qualitativa de grupos focais com especialistas.

4.3. Discussões e contribuições finais dos grupos

Os participantes expressaram seus comentários finais e contribuições para incorporação nas conclusões de acordo com a síntese abaixo. Foram grifados alguns pontos de maior destaque nas discussões.

1. O conhecimeto pedagógico, reconhecido pelos avaliadores como forma particular de conhecimento do conteúdo direcionado concretamente à competência de ensinar apresenta grande potencialidade para atualização e desenvolvimento de modelos e recursos instrucionais a fim de configurar uma pedagogia em linha com as novas formas de aprender e que não se limite aos espaços presenciais em sala de aula, ou à dependência excessiva do direcionamento de professores e livros, sem que sejam tampouco *dependentes unicamente das TIC.*
2. A tecnologia é reconhecidamente necessária na composição da nova pedagogia que deve apoiar o ensino e aprendizagem no contexto contemporâneo e para além de nosso tempo. Porém, apesar dos benefícios, sua assimilação vem como *mais um grande desafio* na cartografia docente tendo em vista sua variedade de aplicabilidades, velocidade de inovação e diferentes graus de eficácia em função da adequada utilização bem como de qualidade da infraestrutura oferecida.
3. A Educação como *ação transformadora* não transfere conhecimentos, mas deve se ocupar da *mediação docente* com o mundo fora da escola buscando

a integração e assimilação tecnológica desenvolvendo o *espírito inovativo* em ambiente de abertura e participação democrática e responsável de todos os envolvidos. Esse pensamento norteou definitivamente a visão dos grupos avaliativos sobre a atividade docente em seu contexto de uma Pedagogia crítica transformadora como valor protagonista do processo educacional.

5. Conclusões

O processo metodológico em si já se constituiu em aprendizagem sobre o modo de construir essas ideias por meio da força do compartilhamento de competências.

Observou-se no desenvolvimento das discussões com avaliadores que propor uma pedagogia é, de fato propor uma visão política e que nessa perspectiva, ao se tratar das práticas de ensinar é preciso falar em política em seu significado social envolvendo a vida dos professores em seu entorno, e o atendimento dos anseios educacionais objetivados democraticamente pelo conjunto da sociedade.

Verificou-se clareza de entendimento sobre adoção tecnológica nos processos de ensino e aprendizagem de forma pedagogicamente eficaz, o que reforça a necessidade de domínio e desenvolvimento constante dessa competência. Diante disso torna-se muito relevante a observância dos "cânones" preconizados por Schulman (1986) enumerando uma série de considerações que cercam o sentido do domínio pedagógico, a saber [tradução nossa]:

- A perspectiva do Conhecimento do professor. Domínios e categorias de conhecimento de conteúdos e suas relações com: (a) conhecimento do conteúdo do assunto em relação ao sujeito, (b) conhecimento do conteúdo pedagógico e (c) conhecimento curricular.
- A perspectiva do Conhecimento do conteúdo pedagógico: forma particular de conhecimento de conteúdo direcionado concretamente à competência de ensinar incluindo as mais usuais formas de representação das ideias, as mais poderosas analogias, ilustrações, exemplos, explanações, demonstrações, enfim, caminhos de representação e formulações do objeto que o tornem compreensível aos alunos.
- A perspectiva do Conhecimento curricular. Espera-se que o professor esteja familiarizado com os materiais curriculares estudados ao mesmo tempo por seus alunos em outras disciplinas relacionando os temas e as discussões.

De fato, como observam Coutinho e Lisboa o docente necessita de uma reconstrução de sua identidade, notadamente os imigrantes digitais que por conta das tecnologias precisam estar preparados para esse novo ambiente," sendo capazes de ressignificar a aprendizagem e estimular a produção coletiva de forma autônoma e organizada através das redes digitais" (2011, p.15)

A incorporação do conceito de A. Bandura sobre motivação e autoeficácia docente contribui ao *adicionar valor* ao modelo de pesquisa e dar *mais voz ao elemento subjetivo* avaliado ao considerar o docente em seu entorno.

Em trabalho em andamento, está sendo estudado um modelo *estrutural alternativo* objetivando incorporação de tecnologia ao processo educacional brasileiro em que se aprofunda a análise dos aspectos institucionais, infraestrutura tecnológica, e

comportamentos esperados de facilitação da incorporação tecnológica por parte de docentes e gestores de forma eficaz e planejada tomando como referência a realidade vivenciada no Brasil.

Além disso, ainda há espaços para realinhamentos, entendimentos e discussões sobre protocolos e normas internacionais, não para frear, mas sim para recuperar a responsabilidade ética que pertence ao criador dos algoritmos e não aos actantes executores. De fato, na vida digital e com as explosões tecnológicas, conforme já previa Negroponte, há quase 25 anos, estamos vivenciando 'o lado obscuro desse progresso tecnológico como abusos de propriedade intelectual, invasão de privacidade, vandalismo digital, pirataria de software e roubo de dados" (Negroponte, 1995, p.215). Não podemos deixar de colocar nessa lista o sucesso negativo das *fakes*.

Diante disso, é preciso colocar com maior veemência e otimismo a necessidade de incrementar o processo de humanização com a apropriação docente das competências cognitivas para desenvolvimento moral e, sobretudo, paradoxalmente, como afirmam Puig e Garcia (2010) atuando fortalecidos com a "própria condição de cidadania e motivação pessoal inerentes à missão docente pelo simples fato de serem pessoas e cidadãos. Não existem especialistas nem saberes especiais." (p. 9) e sim a mediação da escola possibilitando a reflexão humana sobre todo conhecimento disponível

Referências

Bandura, A. (1977). Self-efficacy: towards a unifying theory of behavioral change. *Psychological Review***,** v. 84, n. 2, p. 191–215.

Bandura, A., Azzi R.G., Polydoro, S. (2008). *Teoria Social Cognitiva: conceitos básicos*. Porto Alegre: Artmed.

Castells, M. (2000). *A Sociedade em Rede: A era da informação, economia, sociedade e cultura*. São Paulo: Paz e Terra.

Castells, M. (2002). *A Era da Informação: economia, sociedade e cultura*, Lisboa: fundação calouste gulbenkian.

Coll, C. (2007). Las competencias en la educación escolar: algo más que una moda y mucho menos que un remedio. *Aula de innovación educativa*, *161*, 34–39.

Dacos, M. (2010). Manifesto das digital humanities – THATCamp Paris. Disponível em https://tcp.hypotheses.org/497. Acesso em 16/06/2018

Fabela, S. (2005). A vida toda para aprender. *Portal dos psicólogos*. Disponível em: http://www.psicologia.com.pt/artigos/textos/A0321.pdf. Acessado em: 10/07/2018

Fischer, G (2000). Lifelong Learning – More than training. *Journal of Interactive Learning Research,* Vol. 11 Issue (3/4), pp. 285–29

Fleury, M.T.I. (2001) Construindo o conceito de competência, *Revista de Administração Contemporânea, vol.* 5 (SPE), pp. 183–196

Freire, P. (1982) *Educação e Mudança*. Ed. Paz & Terra. São Paulo.

Freire, P. (2003). *Educação como prática de liberdade*. Rio de Janeiro: Paz e Terra,

Freire, P. (2003). *Pedagogia da Autonomia - saberes necessários à prática educativa*. São Paulo: Paz e Terra

Giroux, H.A., (1994). "Rethinking the Boundaries of Educational Discourse": Modernism, Postmodernism, and Feminism, in Margins in the Classroom: Teaching Literature, ed. Kostas Myrsiades and Linda S. Myrsiades (Minneapolis: University of Minnesota Press), 45. 8

Giroux, H.A. (1999). *Cruzando as fronteiras do discurso educacional: novas políticas em educação*. Porto Alegre, RS; Artes Médicas Sul.

Hargreaves, A. (2003). *O Ensino na Sociedade do Conhecimento: a educação na era da insegurança*. Colecção Currículo, Políticas e Práticas. Porto: Porto Editora

Hirsch, B. D. (Ed.). (2012). *Digital Humanities Pedagogy: Practices, Principles and Politics* (Vol. 3). Open Book Publishers.

Hevner, A.R., March., S.T., Park, J.& Ram, S. (2004). Design Science in information systems research. *MIS Quarterly, 28*(1), 75-106.

Koehler, M. J., & Mishra, P. (2009). What is technological pedagogical content knowledge? *Contemporary Issues in Technology and Teacher Education*, 9(1), 60–70

Levy, P. (2007). *A Inteligência Coletiva: por uma antropologia do ciberespaço* (5a ed.). São Paulo: Loyola

Manovich, L. (2009). Para Lev Manovich falar em "cibercultura" é negar a realidade. Disponível em: https://link.estadao.com.br/noticias/geral,para-lev-manovich-falar-em-cibercultura-e-negar-a-realidade,10000046608. Acesso em: 13 jun.

Mcluhan, M. (2007). *Os meios de comunicação como extensão do homem*. São Paulo: Cultrix,

Negroponte, N. (1995). *A vida digital*. São Paulo: Companhia das Letras.

Organização para a Cooperação e Desenvolvimento Econômico (2000). *Knowledge Management in the Learning Society*. Paris: OCDE. Recuperado de http://ocw.metu.edu.tr/file.php/118/Week11/oecd1.pd.

Pozo, J. I. (2004). A sociedade da aprendizagem e o desafio de converter informação em conhecimento. In: Revista Pátio. Ano VIII – Nº 31- Educação ao Longo da Vida - agosto a outubro de 2004. Disponível em: http://www.revistapatio.com.br/sumario_conteudo.aspx?id=386

Puig, J. M., & García, M. X. (2010). *As sete competências básicas para se educar em valores*. São Paulo: Summus

Saviani, D. (2005). *Pedagogia Histórico Crítica***.** Campinas, SP: Autores Associados.

Saviani, D. (2007). *História das Ideias Pedagógicas no Brasil*. Campinas, SP: Autores Associados.

Silva, A.M. (2008). *Inclusão digital e literacia informacional em Ciência da Informação*. Prisma.com – ojs.letras.up.pt

Shulman, L (1986). Those Who Understand: Knowledge Growth in Teaching. Educational researcher, Vol. 15(2) (Feb 1986), pp. 4–14

Veen, W.; Vrakking, B. (2009). *Homo Sapiens: Educando na era digital*. Porto Alegre: Artmed.

Zarifian, P. (2003). *O modelo da competência: trajetória histórica, desafios atuais e propostas*. São Paulo: Editora Senac São Paulo.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 01/09/2018
Aceitação/Acceptance: 19/11/2018

Explicando el uso de Instagram en Chile: Una aplicación de la técnica de árboles de decisión

Giselle Painén-Aravena[1], Patricio Ramírez-Correa[1], Jorge Alfaro-Peréz[1], Ari Melo Mariano[2], Felipe Machorro-Ramos[3]

giselle.painen@alumnos.ucn.cl, patricio.ramirez@ucn.cl, jalfaro@ucn.cl, arimariano@unb.br, felipe.machorro@udlap.mx

[1] Universidad Católica del Norte, Larrondo 1281, 1781421, Coquimbo, Chile.

[2] Universidad de Brasilia, Campus Darcy Ribeiro, 70904970, Brasilia, Brasil.

[3] Universidad de las Américas Puebla, Ex Hacienda Sta. Catarina Mártir S/N, 72810, San Andrés Cholula, México.

Pages: 626–635

Resumen: La red social Instagram es un fenómeno de interés en la adopción de tecnologías de información. El objetivo de este trabajo es explicar el uso de Instagram en Chile. A partir de datos recopilados en 3.600 hogares sobre usuarios de Internet, se utiliza la técnica de árboles de decisión para predecir la adopción de esta red social. Los resultados informan que la edad del usuario y la adopción de otras aplicaciones predicen el uso de Instagram con una precisión del 86,48%. La principal conclusión del estudio señala que Instagram se posiciona como la red social de gente joven en Chile.

Palabras-clave: Redes sociales; Aceptación de tecnología; Árboles de decisiones; Instagram; Chile.

Explaining the use of Instagram in Chile: An application of the decision tree technique

Abstract: The social network Instagram is a phenomenon of interest in the adoption of information technologies. The objective of this work is to explain the use of Instagram in Chile. Based on data collected in 3,600 households on Internet users, decision tree technique is used to predict the adoption of this social network. The results inform that the age of the user, and the adoption of other applications predict the use of Instagram with an accuracy of 86.48%. The main conclusion of the study indicates that Instagram is positioned as the social network of young people in Chile.

Keywords: Social networks; Technology acceptance; Decision trees; Instagram; Chile.

1. Introducción

Desde su aparición los servicios de redes sociales (SRS) se han expandido a una velocidad impresionante (Meng, Martinez, Holmstrom, Chung, & Cox, 2017), estableciéndose

como un fenómeno extendido y permanente en nuestra sociedad (Waxer, Ninan, Ma, & Dominguez, 2013). Si bien inicialmente los SRS afectaron ámbitos de la vida personal de sus usuarios (Ellison, Steinfield, & Lampe, 2007), la actual influencia de estos sistemas en el ámbito global se ejemplifica tanto en procesos de comunicación política (Kreiss & Mcgregor, 2018) como en la distribución de información financiera de grandes corporaciones (Siikanen et al., 2018). Desde una mirada clásica, Boyd y Ellison (2007) identifican los SRS como servicios de Internet que permiten a los individuos crear un perfil público o semi-público dentro de un sistema cerrado, con el propósito de gestionar una lista de usuarios con los que comparten una conexión a través de crear y consumir contenido. Con una mirada más global, el concepto de SRS se asocia a servicios web en los que los usuarios establecen algún tipo de relación social a través de Internet, lo cual permite la emergencia de cierto tipo de estructura social (Martínez & de Frutos, 2018). De hecho, una característica propia de los SRS es que la decisión de usarlos depende de las interacciones entre los usuarios, es decir, solo tiene sentido cuando un grupo de individuos está dispuesto a utilizar esta tecnología en conjunto (Cheung & Lee, 2010).

Desde la perspectiva de la investigación en aceptación de tecnologías de información, el hecho que los SRS estén basados en tecnología (Qin, Kim, Hsu, & Tan, 2011), su alta popularidad (Choi & Chung, 2013) y sus características únicas (Qin, Kim, & Tan, 2018), han impulsado que su estudio sea considera como un tópico de interés dentro del área. Y si bien diversas investigaciones han tratado frecuentemente a los distintos SRS como un único fenómeno, se ha propuesto el estudio de las distintas plataformas de SRS en forma separada, pues los usuarios de ellas varían tanto en sus características demográficas como sus prácticas de intercambio de información (Choi & Chung, 2013; Meng et al., 2017).

Dentro de los SRS se destaca Instagram (IG), tanto por su rápido crecimiento global (Sheldon & Bryant, 2016) como por sus propiedades estructurales diferenciadas en relación a otros SRS (Manikonda, Hu, & Kambhampati, 2014). IG se lanzó al mercado en el 2010 y tiene por propósito habilitar la comunicación de sus usuarios a través de contenido visual. En el 2018, un total de 1.000 millones de cuentas de IG estaban activas[1]. Desde una mirada académica, la adopción de IG se asocia principalmente al entretenimiento de sus participantes (Alhabash & Ma, 2017), y por tanto, su uso es un fin en sí mismo. Pero más allá de este uso por motivos hedónicos, en el ámbito empresarial IG es considerado un importante factor en el futuro del comercio social, pues a través de este SRS las empresas pueden adquirir nuevos clientes, entregar publicidad e informar de sus promociones (Turban et al., 2017). De hecho, una explicación del éxito de IG señala la unión virtuosa entre la cultura y el comercio (Alang, 2018). No obstante a lo anterior, existe carencia en el estudio académico sobre quiénes son los usuarios de IG (Sheldon & Bryant, 2016).

En este contexto, este trabajo desea aportar con una mirada general de las características de los individuos que optan por incorporarse a este SRS. En específico, este estudio tiene por objetivo explicar el uso de IG en Chile. Para este propósito se utiliza la clasificación basada en árboles de decisiones, una técnica de minería de datos. La importancia científica de este estudio se relaciona con entregar nueva información sobre los usuarios de un SRS de contenido visual, una particularidad distinta a los SRS estudiados con mayor frecuencia en la literatura. Es viable que este tipo de SRS permita la emergencia

[1] https://business.instagram.com.

de un tipo de estructura social distinta, y creemos que las características de sus usuarios permiten escrutar inicialmente esta estructura.

El restante texto se ha organizado de la siguiente forma. Primero realiza una revisión de la literatura sobre IG, y luego se expone la metodología de análisis. Seguido, son presentados los resultados con una breve discusión, y por último se entregan las conclusiones.

2. Revisión de la Literatura sobre IG

Si consideramos que la función básica de IG es el entretenimiento (Alhabash & Ma, 2017), este SRS es un tipo de sistema de información hedónico, es decir, un sistema que tiene como propósito entregar un valor auto realizable a sus usuarios (Ramírez-Correa, Mello, & Mariano, 2018)”type”:”article-journal”,”volume”:”14”},”uris”:[“http://www.mendeley.com/documents/?uuid=96e1d3d5-03f4-4428-89fd-7fbf0d392f58”]}],”mendeley”:{“formattedCitation”:”(Ramírez-Correa, Mello, & Mariano, 2018. Se ha establecido que la intención de utilizar los sistemas hedónicos está determinada por tres antecedentes: la búsqueda de disfrute, la utilidad percibida y la facilidad de uso, siendo también este último antecedente una explicación de los dos anteriores (Van der Heijden, 2004). Por tanto, la diferencia entre la adopción de IG y la adopción de tecnología de funciones utilitaristas está en su antecedente hedónico. En ese sentido, la teoría de uso y gratificación (Katz, Blumler, & Gurevitch, 1974), que señala que las diferencias individuales influyen en las motivaciones para interactuar con los medios y los estilos de vida son más informativos que los datos demográficos para explicar estos deseos, es útil para entender los antecedentes hedónicos del uso de los SRS (Arenas-Gaitán, Rondan-Cataluña, & Ramírez-Correa, 2018)based on social identification theory and uses and gratification theory. This general objective can be divided into two sub-objectives. First, to measure the relationship between social identity, altruism and perceived encouragement as antecedents of WOM in SNS. Second, to study the existence of SNS-user segments which have differentiated behaviors according to the proposed model. Design/methodology/approach – Partial least squares (PLS, y en particular, del uso de IG (Hu, Manikonda, & Kambhampati, 2014).

IG es una plataforma de conciencia social asimétrica, es decir, es muy bajo el valor de reciprocidad entre los usuarios, pues hay muchos seguidores y pocos seguidos (Manikonda et al., 2014). Considerando lo anterior, es muy posible que los motivadores hedónicos del uso de IG sean distintos según se crea contenido, por un usuario seguido, o se consume contenido, por los usuarios seguidores. Asociado a lo anterior, se ha detectado entre los usuarios de IG que, por una parte, una mayor insatisfacción con la vida aumenta el consumo de contenido, y por otra, un mayor nivel de narcisismo aumenta la creación de contenido (Sheldon & Bryant, 2016).

Adicionalmente, los motivadores hedónicos de IG pueden variar según la cultura y género de los individuos. En relación a la variación cultural, existen reportes que muestran es hecho: en estudiantes estadounidenses, un alto nivel de actividad social disminuye el uso de IG (Sheldon & Bryant, 2016), y en cambio, en estudiantes croatas, un objetivo del uso de IG es la interacción social (Sheldon, Rauschnabel, Antony, & Car, 2017)scholars are yet to examine cultural differences among users’ behavior. The current study compares motives for Instagram use between participants from two countries: Croatia, a highly

collectivistic culture, and the United States, a typically individualist culture. Specifically, it examines the relationship between motives and behavioral outcomes of use (time spent on Instagram, the frequency of hashtagging, and the number of Instagram followers. Por otra parte, las diferencias de género han sido ampliamente estudiadas en adopción de tecnologías de información (Ramirez-Correa, Arenas-Gaitan, & Rondan-Cataluna, 2015), y en particular, se ha propuesto esta diferenciación en el caso de los sistemas hedónicos (Wang & Wang, 2008)understanding users' intention to play online games has become a new issue for academics and practitioners. Prior studies have investigated the factors affecting behavioural intention to play online games. However, little research has been conducted to investigate the gender differences in the acceptance of online games. Thus, this study is to investigate the effects of perceived playfulness and its potential antecedents (ie, computer self-efficacy, computer anxiety, challenge, speed and feedback. En el contexto de SRS, se ha demostrado que el género es un moderador de los motivadores psicológicos del uso (Mouakket, 2018), y que en comparación a los hombres, las mujeres crean más contenido con fotos de sí mismas (Dhir, Pallesen, Torsheim, & Andreassen, 2016). En el caso de IG, estudios en Estados Unidos señalan, por una parte, una mayor proporción de mujeres con perfiles activos (Pew Research Center, 2018), y por otra, que ser mujer disminuye la probabilidad de no usar IG (Sheldon & Bryant, 2016).

3. Metodología

Para realizar este estudio se obtuvieron datos desde la IX Encuesta de Acceso y Usos de Internet de Chile (SUBTEL, 2017). Estos datos corresponden a información del año 2017 de 3.600 hogares distribuidos en todo de Chile. En específico, la encuesta fue aplicada con un diseño muestral probabilístico, estratificado por región y zona, y con selección polietápica.

Los datos de la encuesta fueron procesados para obtener a los usuarios de IG y las posibles variables motivadoras de este uso. Además se utilizó la clasificación basada en árboles de decisión (CBAD) con el propósito de explicar el uso de IG. CBAD es una técnica no paramétrica para segmentar un conjunto de datos a partir de los valores de sus atributos, posteriormente esta segmentación es aplicada para predecir una variable específica. El algoritmo de CBAD divide los datos en diferentes segmentos en forma iterativa, considerando que cada uno de los segmentos debe representar lo mejor posible una de las posibles clases de la clasificación analizada, y luego de concluir la creación del modelo de clasificación, cada nuevo caso es evaluado utilizando esta base para predecir una variable objetivo. Dado que gestiona a conjuntos de datos complejos en forma eficiente y genera resultados de apreciación intuitiva, CBAD es utilizado frecuentemente en minería de datos (Ramírez & Grandón, 2018). En concreto, se utilizó SPSS 20.0 para el análisis exploratorio de datos y RapidMiner Studio 9.0 para el análisis de CBAD, RapidMiner utiliza el algoritmo C4.5 para aprendizaje automático en CBAD.

4. Resultados

Se realizó un análisis de las correlaciones de Spearman entre los datos de la encuesta y el uso de IG, con el propósito de seleccionar los atributos para el análisis de la CBAD. A partir de este análisis se eliminaron aquellos atributos con correlación menor a 0,2.

Los atributos seleccionados tienen relación con 1) antecedentes demográficos, 2) uso de Internet y 3) propósitos del uso de Internet. Las Tablas 1 y 2 describen los atributos utilizados para el análisis.

Atributo	Tipo	Media	Desv. Est.
Edad de Acceso	Numérico	44,7	16,5
Edad de Uso	Numérico	43,0	16,8
Ultima vez que uso el internet	Nominal	N	%
Hace 3 meses o menos		3.012	83,7
Entre 3 meses y un año		38	1,1
Hace más de un año		54	1,5
Nunca he utilizado internet		496	13,8
Compras por internet de últimos 12 meses	Nominal	N	%
Al menos una vez al día		385	10,7
Al menos una vez a la semana		85	2,4
Al menos una vez al mes		273	7,6
Menos de una vez al mes		542	15,1
Género	Nominal	N	%
Femenino		1.830	50,8
Masculino		1.770	49,2
Total		3.600	100,0

Tabla 1 – Atributos no dicotómicos para el análisis de la CBAD

Para determinar los parámetros asociados a los criterios a usar en la CBAD se ejecutó un proceso de optimización en RapidMiner. La Tabla 3 muestra los resultados de este proceso para cada parámetro analizado.

Atributos
Conoce Cdf Estadio; Conoce Claro Video; Conoce Google; Conoce Juegos o aplicaciones de smartphone; Conoce Linked-In; Conoce Netflix; Conoce Pinterest; Conoce Skype; Conoce Snapchat; Conoce Spotify; Conoce Twitter; Conoce Vevo; Conoce Youtube; En los últimos 12 meses ha utilizado internet en el hogar de otra persona (amigos, conocidos o parientes); En los últimos 12 meses ha utilizado Internet en el lugar de educación/capacitación; En los últimos 3 meses ha cambiado los ajustes de privacidad en una red social (ejemplo, facebook, twitter, instagram, linkedin, etc); En los últimos 3 meses ha realizado en términos de seguridad algún tipo de herramienta de protección o software (antivirus, anti spam, firewall, etc.) para proteger su computador; En los últimos 3 meses ha usado internet para adquirir información para realizar tareas/información de algún curso, materia o capacitación; En los últimos 3 meses ha usado internet para chatear por whatsapp; En los últimos 3 meses ha usado internet para compartir contenidos (noticias, memes, links) a través de las redes sociales pensando en su uso personal; En los últimos 3 meses ha usado internet para compartir estados de ánimo y reflexiones a través de las redes sociales ; En los últimos 3 meses ha usado internet para crear y subir contenidos propios en redes sociales como facebook, twitter, youtube (fotos, información, videos) pensando en su uso personal; En los últimos 3 meses ha usado internet para crear y subir contenidos a la web; En los últimos 3 meses ha usado internet para descargar, o escuchar en línea: música, películas, cortometrajes o imágenes, estaciones de radio, canales de televisión.; En los últimos 3 meses ha usado internet para enviar y/o recibir correos electrónicos; En los últimos 3 meses ha usado internetpara realizar llamadas telefónicas/videoconferencia por internet; En los últimos 3 meses ha usado internet para utilizar redes sociales (facebook, twitter, linkedin etc.); Paga un monto extra por el uso de Netflix; Usa el computador para realizar tareas como encontrar e instalar programas (software) desde internet; Usa

el computador para realizar tareas como postear/ publicar mensajes en plataformas especializadas (foros online, newsgroups, etc.); Usa el computador para realizar tareas como transferir archivos entre artículos electrónicos (cámaras digitales, celulares, reproductores de música); Usa el computador para realizar tareas como el uso avanzado de planillas de cálculo: funciones de filtrado de datos, organización de datos, análisis estadísticos, gráficos; Usa el computador para realizar tareas como el uso de procesador de texto; Usa el computador para realizar tareas como el uso de software de presentación; Usa el computador para realizar tareas como fórmulas aritméticas sencillas en planillas de cálculo ; Usa el computador para realizar tareas como modificar configuración es de seguridad de navegadores de Internet (Chrome, Explorer, Firefox, etc.); Usa el computador para realizar tareas como modificar o verificar configuración de programas y aplicaciones; Usa Facebook habitualmente; Usa google cuando quiere buscar información de productos o servicios; Usa google cuando quiere buscar información para tareas, actividades, trabajos; Usa Google habitualmente; Usa habitualmente juegos o aplicaciones de smartphone; Usa Netflix habitualmente; Usa Pinterest habitualmente; Usa Snapchat habitualmente; Usa Spotify habitualmente; Usa Twitter habitualmente; Usa twitter para entretenerse; Usa whatsapp habitualmente; Usa Youtube habitualmente.

Tabla 2 – Atributos dicotómicos para el análisis de la CBAD

Parámetro	**Valor**
Relación de separación	0,85
Criterio para seleccionar atributos para dividir	Índice Gini
Máxima profundidad del árbol	5
Se aplica poda	Falso
Se aplica pre poda	Falso

Tabla 3 – Valor de parámetros de CBAD

Considerando los parámetros determinados previamente se ejecutó el algoritmo de la CBAD. La Figura 1 muestra el resultado.

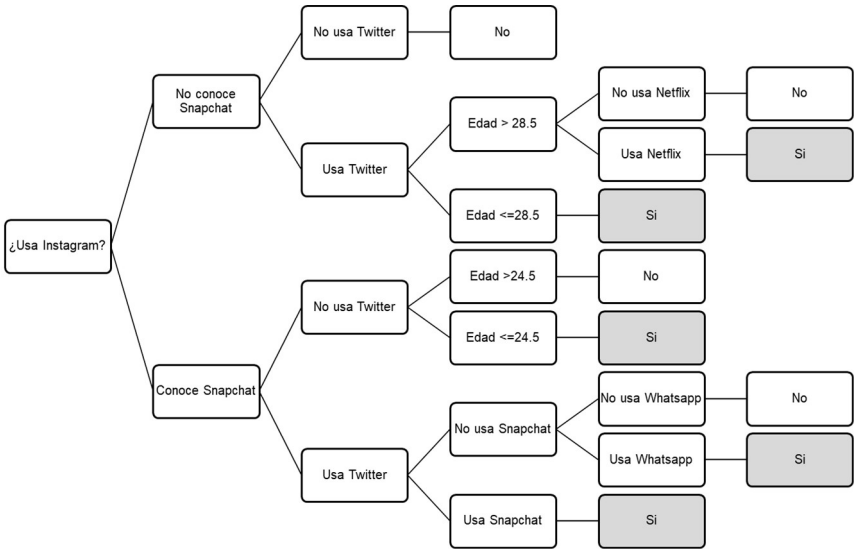


Figura 1 – Modelo de predicción del uso de IG

Finalmente, para realizar el análisis predictivo se dividieron los datos en dos grupos, el primero con 85% de los datos y el segundo con el 15% restante. El primer grupo de datos se utilizó para generar el modelo, y el segundo grupo para evaluar su capacidad predictiva. La Tabla 4 presenta la matriz de confusión para la predicción de uso de IG, a partir de su análisis se determinó que el modelo de predicción tiene un ratio de precisión del 86,48% (56/110*110/540+411/430*430/540).

		Predicción de Uso de IG		
		Sí	**No**	**Total**
Uso Real de IG	*Sí*	56	19	75
	No	54	411	465
Total		110	430	540

Tabla 4 – Matriz de confusión para la predicción de uso de IG

En resumen, la aplicación de la CBAD muestra que se puede explicar el uso de IG en base a la edad del usuario, el uso de otros SRS y de otros sistemas hedónicos.

4. Discusión

A partir de los resultados podemos discutir algunos elementos relevantes. Primero, y asociado a la teoría de uso y gratificación, encontramos en el conocimiento y uso de Snapchat son claves del estilo de vida de los usuarios de IG. Snapchat es un SRS que apunta la mensajería efímera, videos cortísimos e imágenes que duran solo horas en la web. Los usuarios prefieren este SRS para compartir experiencias espontáneas con los más cercanos (Bayer, Ellison, Schoenebeck, & Falk, 2016), y lo usan más para fortalecer capital social que para incrementarlo con nuevas relaciones (Piwek & Joinson, 2016), es decir, en Snapchat no se buscan extraños (Vaterlaus, Barnett, Roche, & Young, 2016). Esto puede relevar un aspecto importante de los usuarios chilenos de IG, pues probablemente busquen en IG los mismos resultados que en Snapchat, fortalecer una la estructura social existente fuera de Internet.

Segundo, el género de los usuarios chilenos no aparece como un antecedente de utilizar IG, si bien este hallazgo es contradictorio con estudios anteriores, se puede entender si consideramos que en un contexto de búsqueda de gratificaciones de auto mejoramiento, los hombres realizan un comportamiento más de consumo de contenido, y las mujeres más de generación de contenido (Krasnova, Veltri, Eling, & Buxmann, 2017). Esto último puede explicar la mayor implicación de mujeres en IG reportada en trabajos previos. Es decir, el género no sería un predictor si consideramos solo el uso de IG, pero esto podría variar si separamos el uso en consumo y generación de contenido.

Finalmente, se debe destacar que el análisis indica que ser joven determina en forma importante el uso de IG. Si bien este resultado es predecible, dado la orientación de IG, sus implicaciones son interesantes. La mayoría de los usuarios de IG en Chile son también consumidores iniciales, dado que aún están en la etapa de formación académica o iniciando su vida laboral, y en medida que aumente su edad promedio, su nivel de consumo aumentará, y por tanto, la importancia de IG para las empresas como medio para llegar a este mercado.

5. Conclusiones

Este trabajo empleó información secundaria para explicar la utilización de IG en Chile. A partir de estos datos fue posible determinar un modelo de CBAD con una precisión en la predicción superior al 86%. Dado este análisis es posible predecir el uso de IG a partir variables demográficas y comportamentales. En específico, la juventud de los usuarios como el uso de las aplicaciones Snapchat, Whatsapp, Twitter y Netflix son fuertes antecedentes de aceptar IG. La principal conclusión del estudio es que IG se posiciona como el SRS de los jóvenes en Chile.

Una de las principales limitaciones de este estudio es la carencia de datos sobre el propósito del uso de IG. Y dada la evidencia en la literatura, creemos que esto podría ser un factor que modere los resultados, y por tanto, en futuros estudios se debería examinar si el uso de IG es para la creación o el consumo de contenido.

Referencias

Alang, N. (2018, June 29). Why Instagram conquered social media. *The Week*. Retrieved October 25, 2018, from https://theweek.com/articles/781459/why-instagram-conquered-social-media

Alhabash, S., & Ma, M. (2017). A Tale of Four Platforms: Motivations and Uses of Facebook, Twitter, Instagram, and Snapchat Among College Students? *Social Media + Society*, *3*(1), 205630511769154. Retrieved from http://journals.sagepub.com/doi/10.1177/2056305117691544

Arenas-Gaitán, J., Rondan-Cataluña, F. J., & Ramírez-Correa, P. E. (2018). Antecedents of WOM: SNS-user segmentation. *Journal of Research in Interactive Marketing*.

Bayer, J. B., Ellison, N. B., Schoenebeck, S. Y., & Falk, E. B. (2016). Sharing the small moments: ephemeral social interaction on Snapchat. *Information Communication and Society*, *19*(7), 956–977.

Boyd, D. M., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, *13*(1).

Cheung, C. M. K., & Lee, M. K. O. (2010). A theoretical model of intentional social action in online social networks. *Decision Support Systems*, *49*(1), 24–30.

Choi, G., & Chung, H. (2013). Applying the Technology Acceptance Model to Social Networking Sites (SNS): Impact of Subjective Norm and Social Capital on the Acceptance of SNS. *International Journal of Human-Computer Interaction*, *29*(10), 619–628. Taylor & Francis .

Dhir, A., Pallesen, S., Torsheim, T., & Andreassen, C. S. (2016). Do age and gender differences exist in selfie-related behaviours? *Computers in Human Behavior*, *63*, 549–555.

Ellison, N. B., Steinfield, C., & Lampe, C. (2007). The benefits of facebook "friends:" Social capital and college students' use of online social network sites. *Journal of Computer-Mediated Communication*.

Van der Heijden, H. (2004). User acceptance of hedonic information systems. *MIS quarterly*, 695–704. JSTOR.

Hu, Y., Manikonda, L., & Kambhampati, S. (2014). What we Instagram : a first analysis of Instagram photo content and user types. *Proceedings of the Eight International AAAI Conference on Weblogs and Social Media* (pp. 595–598).

Katz, E., Blumler, J. G., & Gurevitch, M. (1974). *The uses of mass communication:Current Perspectives on Gratifications Research. Sociology The Journal Of The British Sociological Association*.

Krasnova, H., Veltri, N. F., Eling, N., & Buxmann, P. (2017). Why men and women continue to use social networking sites: The role of gender differences. *The Journal of Strategic Information Systems*. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0963868717300331

Kreiss, D., & Mcgregor, S. C. (2018). Technology Firms Shape Political Communication: The Work of Microsoft, Facebook, Twitter, and Google With Campaigns During the 2016 U.S. Presidential Cycle. *Political Communication*.

Manikonda, L., Hu, Y., & Kambhampati, S. (2014). Analyzing user activities, demographics, social network structure and user-generated content on Instagram. *arXiv preprint arXiv:1410.8099*.

Martínez, J. A. D., & de Frutos, T. H. (2018). Connectivism in the Network Society. The Coming of Social Capital Knowledge. *Tendencias Sociales. Revista de Sociología*, (1), 21–37.

Meng, J., Martinez, L., Holmstrom, A., Chung, M., & Cox, J. (2017). Research on Social Networking Sites and Social Support from 2004 to 2015: A Narrative Review and Directions for Future Research. *Cyberpsychology, Behavior, and Social Networking*.

Mouakket, S. (2018). The role of personality traits in motivating users' continuance intention towards Facebook: Gender differences. *Journal of High Technology Management Research*, *29*(1), 124–140.

Pew Research Center. (2018). *Social Media Use 2018. Pew Research Center*. Retrieved from http://www.pewinternet.org/2018/03/01/social-media-use-in-2018/

Piwek, L., & Joinson, A. (2016). "What do they snapchat about?" Patterns of use in time-limited instant messaging service. *Computers in Human Behavior*, *54*, 358–367.

Qin, L., Kim, Y., Hsu, J., & Tan, X. (2011). The effects of social influence on user acceptance of online social networks. *International Journal of Human-Computer Interaction*.

Qin, L., Kim, Y., & Tan, X. (2018). Understanding the Intention of Using Mobile Social Networking Apps Across Cultures. *International Journal of Human-Computer Interaction*.

Ramirez-Correa, P. E., Arenas-Gaitan, J., & Rondan-Cataluna, F. J. (2015). Gender and Acceptance of E-Learning: A Multi-Group Analysis Based on a Structural Equation Model among College Students in Chile and Spain. *Plos One*, *10*(10).

Ramírez-Correa, P. E., Mello, T. M., & Mariano, A. M. (2018). A aceitação da Netflix: um estudo utilizando equações estruturais. *Revista Tecnologia e Sociedade*, *14*(30).

Ramírez, P. E., & Grandón, E. E. (2018). Predicción de la Deserción Académica en una Universidad Pública Chilena a través de la Clasificación basada en Árboles de Decisión con Parámetros Optimizados. *Formación universitaria*, *11*(3), 3–10. Centro de Información Tecnológica.

Sheldon, P., & Bryant, K. (2016). Instagram: Motives for its use and relationship to narcissism and contextual age. *Computers in Human Behavior*, *58*, 89–97.

Sheldon, P., Rauschnabel, P. A., Antony, M. G., & Car, S. (2017). A cross-cultural comparison of Croatian and American social network sites: Exploring cultural differences in motives for Instagram use. *Computers in Human Behavior*, *75*, 643–651.

Siikanen, M., Baltakys, K., Kanniainen, J., Vatrapu, R., Mukkamala, R., & Hussain, A. (2018). Facebook drives behavior of passive households in stock markets. *Finance Research Letters*.

SUBTEL. (2017). *IX Encuesta Accesos y Usos de Internet*. Retrieved from www.subtel.gob.cl/wp-content/uploads/2018/05/ppt_usos_may2018.pdf

Turban, E., Outland, J., King, D., Lee, J. K., Liang, T.-P., & Turban, D. C. (2017). *Electronic commerce 2018: a managerial and social networks perspective*. Springer.

Vaterlaus, J. M., Barnett, K., Roche, C., & Young, J. A. (2016). "snapchat is more personal": An exploratory study on Snapchat behaviors and young adult interpersonal relationships. *Computers in Human Behavior*, *62*, 594–601.

Wang, H. Y., & Wang, Y. S. (2008). Gender differences in the perception and acceptance of online games. *British Journal of Educational Technology*, *39*(5), 787–806.

Waxer, N., Ninan, D., Ma, A., & Dominguez, N. (2013). How Cloud Computing and Social Media Are Changing the Face of Health Care. *Physician Executive*, *39*(2), 58–62.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 17/09/2018
Aceitação/Acceptance: 06/11/2018

Metodología para hallar patrones de movimiento en trayectorias sobre DBMS

Diego Fernando Rodriguez Lamus[1], Álvaro Enrique Ortiz Dávila[1]

dfrodriguezl@correo.udistrital.edu.co, aeortizd@udistrital.edu.co

[1] Universidad Distrital Francisco José de Caldas, Facultad de Ingeniería, Bogotá, Colombia.

Pages: 636–647

Resumen: Con los dispositivos móviles actuales es posible capturar una gran cantidad de información proveniente del movimiento de los objetos como automóviles, bicicletas, etc., donde se tienen datos de ubicación y tiempo que permiten hacer un rastreo de la posición de un objeto en el tiempo. Como la información capturada aumenta día tras día, la cantidad ha sobrepasado la capacidad de análisis por parte de los analistas, por lo que se recurre a técnicas que permiten descubrir información implícita en los datos a través de técnicas computacionales, una de ellas es el descubrimiento de patrones sobre trayectorias ya que brinda una aproximación de un ciclo de movimiento de un objeto. Para detectar estos patrones en trayectorias se hace uso de dos propuestas encontradas en la literatura y se implementan en un motor de base de datos utilizando el marco de las bases de datos de objetos en movimiento.

Palabras-clave: trayectorias; DBMS; patrones; bases de datos de objetos en movimiento; movimiento.

Methodology to find patterns of movement in trajectories on DBMS

Abstract: In current mobile devices it is possible to capture a large amount of information about movement of objects such as cars, bicycles, etc., where you can obtain location and time data that allow you to track position of an object over time. As information captured increases day after day, amount has exceeded analysis capacity of the analysts, so they resort to techniques that allow to discover implicit information in data through computational techniques, one of them is discovery of trajectory patterns offered by an approximation of a movement cycle of an object. To detect these patterns in trajectories, two proposals found in literature are used and they are implemented in a database engine using the framework Moving Objects Database – MOD.

Keywords*:* Trajectories; DBMS; patterns; moving objects databases; movement.

1. Introducción

Con el auge de los dispositivos móviles y por ende de los sensores incorporados a ellos, como los que recepcionan la señal de los satélites de posicionamiento global – GNSS, que

indican la posición en centésimas de segundo, se ha abierto un abanico de posibilidades para extraer comportamientos de objetos en base a su movimiento. Una de las opciones es la construcción y análisis de trayectorias a partir de los datos de posición arrojados cada determinado intervalo de tiempo por algún dispositivo receptor, lo cual permite extraer conocimiento de dichos datos como el descubrimiento de patrones, detección de anomalías, agrupación y clustering, etc., donde la detección de anomalías juega un papel muy importante ya que es la que permite descubrir comportamientos anormales en un conjunto de datos con respecto al resto del conjunto o a otro que se tenga como modelo.

Una de las tareas más importantes en la minería de datos es la detección de patrones que consiste en el descubrimiento de comportamientos regulares en los datos. Para hallar esos patrones en trayectorias se parte de la premisa que el movimiento se representa en 3 dimensiones (ubicación XY y tiempo T), lo cual le agrega algo de complejidad al diseño y ejecución de ciertos algoritmos que deben ser adaptados al contexto de los objetos en movimiento y sus trayectorias.

En la actualidad existen diversos algoritmos que proponen formas de detectar patrones, en el presente artículo se aborda una metodología mezclada entre (Ying, Xu, & Guo Yin, 2009) y (Jae-Gil, Han, & Whang, 2007), donde se trata de reducir la trayectoria a su mínima expresión a través de puntos característicos, para luego aplicar una variante del algoritmo DB-SCAN orientado hacia trayectorias, utilizando conceptos como el volumen de los mínimos cubos envolventes de las sub trayectorias. Existe una parte implementada en base de datos que es la de los puntos característicos, mientras que lo demás existe en propuesta lo cual se implementó en Oracle bajo las extensiones Hermes (Pelekis N. , Frentzos, Giatrakos, & Theodoridis, 2010) y Oracle Spatial para que pueda ser utilizada en casos reales de aplicación.

2. Marco Teórico

2.1. Trayectorias y su modelamiento en DBMS.

2.1.1. Trayectorias

Las trayectorias son aquellas que representan el movimiento de un objeto desde el inicio hasta el fin del mismo. Como primera definición para entender que es una trayectoria, es necesario comprender que es el movimiento, el cual consiste en el "cambio en el tiempo de la posición u orientación de un cuerpo" (Encyclopædia Britannica, 2017), como se deduce de la definición planteada, el movimiento de un objeto depende esencialmente de dos componentes: la posición y el tiempo, lo cual implica que para modelar el mismo es necesario tener en cuenta los componentes mencionados (Zhou, Zheng, Jueng, Xu, & Sadiq, 2015).

La construcción derivada de la anterior definición asume que el movimiento es de naturaleza continua, lo que representa un reto para su representación en los ordenadores, siendo necesario un proceso de discretización a partir de una serie de puntos ordenados compuestos de coordenadas en un plano (p_i) y el tiempo t de toma de la ubicación, estos puntos deben ser representativos dentro del trazado de la trayectoria para trabajar con datos aproximados a la realidad, como se observa:

$$T = \{< p_1, t_1 >, < p_2, t_2 >, ..., < p_n, t_n >\}$$

Donde $p_i \in \mathbb{R}^2, t_i \in \mathbb{R}, 1 \leq i \leq n$ y $t_1 < t_2 < ... < t_n$

Existe una pregunta acerca del movimiento que hace concluir que nuestro conocimiento acerca del mismo es de naturaleza continua, y la pregunta es ¿Dónde estuvo localizado un objeto en un tiempo entre dos puntos tomados?, la respuesta incluye cierta incertidumbre que depende de la tasa de muestreo y los parámetros del movimiento. Se utiliza técnicas de interpolación, siendo la más popular la interpolación lineal debido a su fácil implementación.

En esta técnica de implementación se asume que la velocidad y dirección permanecen constantes durante el intervalo de tiempo definido. Este modelo se conoce como "representación cortada", que consiste en la descomposición del desarrollo temporal del movimiento en fragmentos llamados "rebanadas" como se presenta en la Figura 1, donde entre cada trozo o punto muestreado existe una función simple que describe el movimiento entre ellos que puede ser de tipo lineal o de arco (Pelekis & Theodoridis, 2014).

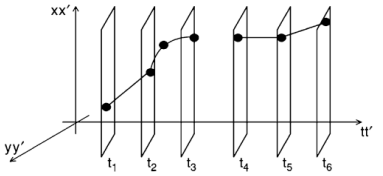


Figura 1 – Representación por trozos para la definición de objetos en movimiento (Pelekis & Theodoridis, 2014).

2.2. Descubrimiento de patrones en trayectorias.

Patrón

Un patrón se define como un comportamiento regular hallado en los datos, es decir, cuando de alguna manera estos tienden a repetirse de una manera sistemática o predecible. Específicamente para propósitos de este trabajo, un patrón espacio-temporal sobre un grupo de trayectorias se refiere a la existencia de objetos ocupando y moviéndose en un mismo espacio geográfico durante un mismo intervalo de tiempo que debe o puede ser representado por caminos o trayectorias representativas del movimiento de los objetos que son similares (Han, Kamber, & Pei, 2012) (Yu, 2015).

2.2.1. Clúster

El clustering es el proceso en el cual se agrupan un conjunto de datos en múltiples grupos o clusters, de lo anterior se deduce que los objetos dentro de un mismo grupo tienen alta similitud, mientras que son muy diferentes de los objetos en otros grupos. Las medidas de similitud y no similitud se calculan a partir de los valores de los atributos que describen los objetos que con frecuencia involucran medidas de distancia (Han, Kamber, & Pei, 2012) (Panagiotakis, Pelekis, Kopanakis, Ramasso, & Theodoridis, 2012).

Para definir los cluster es necesario definir cuál es la medida de similitud o no similitud entre los grupos encontrados, así como la determinación de los atributos involucrados en la misma (Li, y otros, 2010).

En el caso del presente trabajo, los criterios que se tienen en cuenta para realizar el hallazgo de los patrones son:

- La dirección del movimiento debe ser la misma (dimensión espacial)
- El intervalo de tiempo entre dos puntos debe ser similar (dimensión temporal)

Para encontrar los patrones del movimiento, es requerido, realizar las segmentaciones de las trayectorias en sub-trayectorias, donde se utiliza el concepto de puntos característicos para realizar la partición en zonas donde el comportamiento del movimiento cambia rápidamente (Jae-Gil, Han, & Whang, 2007).

2.2.2. Puntos Característicos

Como se mencionó los puntos característicos, son los puntos donde el comportamiento de una trayectoria cambia rápidamente. Se tiene una trayectoria:

$$TR_i = p_1 p_2 p_3 \ldots p_j \ldots p_{len_i}$$

La cual tiene un conjunto de puntos característicos

$$\{p_{c_1}, p_{c_2}, p_{c_3}, \ldots, p_{c_{par_i}}\}$$

Donde $(c_1 < c_2 < c_3 < \ldots < c_{par_i})$

Luego la trayectoria TR_i es particionada en cada punto característico, y cada segmento de la misma es representado por un segmento de línea entre dos puntos característicos sucesivos. La trayectoria es segmentada en un conjunto de $(par_i - 1)$ segmentos $\{p_{c_1}, p_{c_2}, p_{c_3}, \ldots, p_{c_{par_i-1}} p_{c_{par_i}}\}$

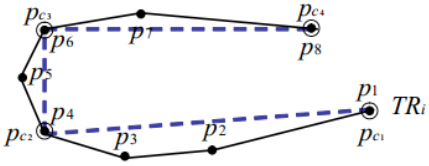


Figura 2 – Ejemplo de una trayectoria y sus particiones (Jae-Gil, Han, & Whang, 2007)

En la Figura 2, se observa en línea de color negro la trayectoria y en línea azul punteada se observan las particiones de la trayectoria o sub-trayectorias que se obtienen entre los puntos característicos simbolizados con una circunferencia alrededor del punto. La partición óptima debe poseer dos características: precisión y concisión (Jae-Gil, Han, & Whang, 2007).

- **Precisión:** Se refiere a que la diferencia entre una trayectoria y un conjunto de las particiones de la misma, debe ser lo más pequeña posible.
- **Concisión:** Significa que el número de particiones de la trayectoria debe ser lo más pequeño posible.

Para encontrar el punto de equilibrio entre la concisión y la precisión, se adopta el método de Longitud de descripción mínima (Minimum Description Length - MDL), como principio general este consiste en que cualquier regularidad en los datos puede ser usada para comprimirlos, es decir, describirlos usando la mínima cantidad de símbolos que el número de símbolos necesarios para describir los datos literalmente, entre más regularidad exista mayor es la tasa de compresión (Grünwald, Myung, & Pitt, 2005).

La definición de MDL incluye dos conceptos claves que son $d_\perp$ y d_θ, equivalentes a la distancia perpendicular y a la distancia angular, respectivamente. Estas distancias forman los componentes de la función de distancia para la segmentación de trayectorias utilizando puntos característicos como se observa en la Figura 3.

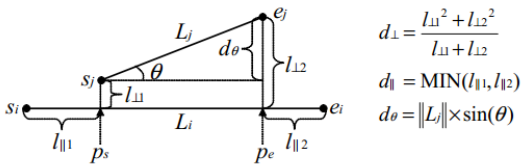


Figura 3 – Componentes de la función de distancia para la segmentación de trayectorias a través de puntos característicos (Jae-Gil, Han, & Whang, 2007)

2.2.3. Mínimo cubo envolvente (MBB)

El mínimo cubo envolvente, se refiere al mínimo cubo que es capaz de envolver la totalidad del segmento de trayectoria hallada mediante puntos característicos, es decir, está dado por las coordenadas de los dos puntos característicos que conforman el segmento como se presenta en la Figura 4:

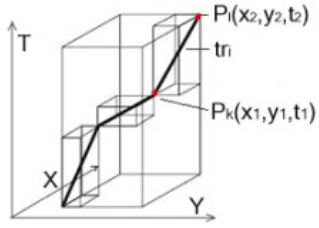


Figura 4 – Mínimo cubo envolvente (Ying, Xu, & Guo Yin, 2009)

Se parte de la representación tridimensional del movimiento a partir de trayectorias, donde se cuenta con los dos puntos característicos P_i y P_k, compuestos por coordenadas geográficas (x-y) y el tiempo, lo cual define un sistema de coordenadas en tres dimensiones

como se ve en la Figura 4. Esta forma de representar la información permite resumir la información en estos cubos lo cual acelera la velocidad del procesamiento que se requiere para procedimientos posteriores.

Una vez definido el concepto de mínimo cubo envolvente, se puede definir el cálculo de distancias de similitud entre dos segmentos de trayectorias, a partir del volumen y el traslape:

$$Volume(MBBtr_i) = (x_2 - x_1) * (y_2 - y_1) * (t_2 - t_1)$$

La fórmula es la misma del cálculo del volumen tradicional de un cuerpo donde según la Figura 4, el alto es la diferencia entre t_2 y t_1, el ancho la diferencia entre coordenadas x´s y el largo la diferencia entre las coordenadas y's.

La otra medida importante es la métrica de Traslape, que representa el volumen de intersección entre dos cubos, donde si el volumen de intersección es igual al volumen de un cubo, se deduce que el movimiento tiene comportamiento similar, lo cual quiere decir que son vecinos en las dimensiones espacio temporales. En cuanto a esta medida pueden existir dos tipos de relaciones espaciales: un traslape y "Disjoint", donde el primero significa que los cubos tienen algún tipo de relación espacial y el segundo que no hay ningún tipo de relación entre los dos cubos.

Con base en las dos medidas anteriores, se estructura la medida de similitud que puede ser calculada de la siguiente manera (Ying, Xu, & Guo Yin, 2009; Ying, Xu, & Guo Yin, 2009):

$$Sim = \frac{OV(MBB_i, MBB_j)}{Volume(MBB_i)}$$

Donde $OV(MBB_i, MBB_j)$, es el área de traslape entre los dos cubos envolventes y $Volume(MBB_i)$ es el volumen del cubo.

Una vez definida la medida de similitud, se puede definir la medida de distancia de similitud entre tr_i y tr_j como (Ying, Xu, & Guo Yin, 2009):

$$SDist = (1 - Sim) * \beta * (len(tr_i) + len(tr_j))$$

Donde $len(tr_i)$ representa la longitud del segmento tr_i, β es un peso espacio-temporal para representar distancia de similitud. El peso está sujeto a un intervalo de

$[0, \frac{\sqrt{2}}{2} * (len(tr_i) + len(tr_j))]$.

Algoritmo DB-SCAN para trayectorias

Para el hallazgo de los patrones se utiliza una de las tareas de minería de datos más famosas, el clustering, utilizando el algoritmo DB-SCAN (Density-Based Clustering Based in Connected Regions with High Density) orientado hacia trayectorias. En su versión

original este algoritmo encuentra objetos núcleo, donde su principal característica es la presencia de vecindarios densos. Se utilizan principalmente dos parámetros (Han, Kamber, & Pei, 2012):

- *E-neighborhood,* que es usado para especificar el radio de un vecindario, el radio de un objeto *o* es el espacio dentro de un radio E centrado en *o*.
- *MinTrs,* que se refiere al límite de densidad de las regiones. Un objeto es un objeto núcleo si y solo si, *E-neighborhood* del objeto contiene por lo menos *MinTrs* objetos.

En el contexto actual de las trayectorias, las modificaciones propuestas para utilizar este algoritmo son las siguientes (Ying, Xu, & Guo Yin, 2009):

- Se realiza clúster de trayectorias en vez de puntos
- Se utiliza la medida de distancia previamente hallada SDist para encontrar sub trayectorias vecinas.

El algoritmo seleccionado para el clustering de las trayectorias funciona de la siguiente manera, inicialmente las sub trayectorias son consideradas como no clasificadas. El algoritmo consiste de dos fases, el primer paso consiste en el cálculo del parámetro Eps-Neighborhood para cada sub trayectoria no clasificada tr. Si tr es una sub trayectoria núcleo, se verifica $\left|N_{Eps}(tr_j)\right| \geq MinTrs$, si se cumple la condición planteada, para cada tr_i que pertenezca o se encuentre dentro del radio de la sub trayectoria núcleo tr, asigne tr_i al clúster especificado.

3. Metodología para hallar patrones de movimiento en trayectorias sobre DBMS

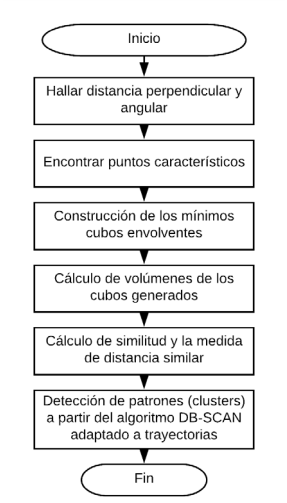


Figura 5 – Diagrama de flujo que indica el procedimiento general de la metodología planteada en el artículo.

El diagrama presentado en la Figura 5, se puede explicar de la siguiente manera: el primer algoritmo descrito es el de "puntos característicos", que fue seleccionado debido a que reduce una trayectoria a su mínima representación posible sin perder la información relevante que describe la trayectoria, lo cual hace que como insumo en los siguientes algoritmos para la detección de patrones se tenga una trayectoria mucho más sencilla que la original que a su vez aumenta la velocidad de ejecución haciendo mucho más eficiente los procedimientos posteriores.

Esta reducción de la información a su mínima expresión en el caso específico de los datos sobre movimiento recurre a la distancia perpendicular y angular entre los segmentos que componen una trayectoria.

Una vez se obtienen los puntos característicos se aborda otro algoritmo denominado "Mínimo Cubo Envolvente", que permite descomponer una trayectoria en cubos que son los de menor tamaño que pueden encerrar a cada uno de los segmentos que la conforman, lo cual permite describir cada uno de esos trozos a través de dos magnitudes denominadas "volumen" y "traslape", esto permite describir las trayectorias mediante cifras para hallar la distancia utilizada en el algoritmo de detección de clústers, reemplazando los datos crudos que representan un alto costo computacional.

A partir de las dos magnitudes que describen los mínimos cubos envolventes (volumen y traslape), se calcula la medida de similitud entre dos segmentos, donde esta última se utiliza para calcular a su vez la distancia de similitud que es finalmente la utilizada para determinar a qué grupo o clúster pertenece determinado segmento.

Por último, el algoritmo que es el núcleo de esta fase del trabajo, es una versión adaptada del método DB-SCAN al tema de las trayectorias, donde se parte de dos conceptos clave: el radio de los clúster y los objetos núcleo (mínima cantidad de trayectorias en un clúster).

Sintetizando, en una primera etapa se reducen las trayectorias a su mínima descripción posible (generalización) conservando la mayor cantidad de información relevante a través del método denominado "Minimum Description Length" para hallar los puntos característicos que representan una trayectoria, para luego con estos, extraer dos magnitudes (volumen y traslape) a partir de los mínimos cubos envolventes que permiten reducir de manera significativa el procesamiento de los datos ya que no se utilizan los datos crudos para hallar las distancias de similitud si no cifras que representan de una forma eficiente el comportamiento de las mismas. Finalmente, a partir de estas magnitudes calculadas se calcula el radio de cada uno de los grupos hallados mediante el algoritmo DB-SCAN adaptado a trayectorias, con el fin de hallar las trayectorias representativas que describan el comportamiento general de los objetos en movimiento analizados.

3.1. Hallar los puntos característicos de las trayectorias

Para hallar los puntos característicos de las trayectorias analizadas, se utilizó el procedimiento almacenado denominado TRACLUS.approximate_traj_partitioning que ya se encuentra implementado en la extensión HERMES para ORACLE, que funciona de la manera que se presenta en la Figura 6:

Algorithm 1: Partición aproximada de una trayectoria

Input: Una trayectoria $TR_i = p_1p_2p_3...p_j...p_{len_i}$
Output: Un conjunto CP_i de puntos característicos
Añadir p_1 al conjunto Cp_i;
/* El punto de inicio */
$startIndex := 1, length := 1$;
while $startIndex + length <= len_i$ **do**
  $currIndex := startIndex + length$;
  $cost_{par} := MDL_{par}(p_{startIndex}, p_{currIndex})$;
  $cost_{nopar} := MDL_{nopar}(p_{startIndex}, p_{currIndex})$;
  /* Validar si la partición en el punto actual hace que el costo MDL sea más grande que la no partición */
  if $(cost_{par} > costnopar)$ **then**
    /* partición en el punto previo */
    Añadir $p_{currIndex-1}$ al conjunto CP_i;
    $startIndex := currIndex - 1, length := 1$;
  else
    $length := length + 1$;
  end
end
/* El punto final */
Añadir p_{len_i} al conjunto CP_i;

Figura 6 – Pseudocódigo – algoritmo partición aproximada de una trayectoria (Ying, Xu, & Guo Yin, 2009)

3.2. Medidas de similitud

Una vez calculados los dos volúmenes, se puede calcular a partir de estos las medidas de similitud que son las encargadas de medir la similitud entre sub trayectorias con el fin de incluirlas en algún clúster.

La primera medida denominada Sim se calcula a partir del cociente entre el volumen de traslape y el volumen de la MBB correspondiente a la sub trayectoria que se está analizando como se ve en la Figura 7:

Algorithm 4: Cálculo de similitud - *Sim*

Input: El volumen de traslape OV
Input: El volumen del MBB analizado V
Output: La similitud entre dos MBB's - Sub trayectorias *Sim*
/* Calcular similitud a partir del cociente de las dos entradas */
Sim := OV/V;

Figura 7 – Pseudocódigo para calcular similitud – Sim.

A partir de la distancia de similitud – Sim se calcula la métrica que utilizara el algoritmo de agrupamiento denominada SDist (Similar Distance Measure – Medida de distancia similar), donde se requiere un peso β, que según (Ying, Xu, & Guo Yin, 2009) puede aproximarse a un valor de √2/2, con lo cual se tienen todos los términos para la implementación utilizando el esquema planteado en la Figura 8:

Algorithm 5: Cálculo de medida de distancia similar - *SDist*

Input: La similitud entre dos MBB's - Sub trayectorias *Sim*
Input: El punto mínimo del cubo No. 1 p_{min_1} - algún punto característico
Input: El punto máximo del cubo No. 1 p_{max_1} - algún punto característico
Input: El punto mínimo del cubo No. 2 p_{min_2} - algún punto característico
Input: El punto máximo del cubo No. 2 p_{max_2} - algún punto característico
Output: Medida de distancia similar - *SDist*
/* Asignar valor al peso β */
$\beta := \sqrt{2}/2$;
/* Calcular longitud de la sub trayectoria No. 1 */
length1 := distance($p_{min_1}.x, p_{min_1}.y, p_{max_1}.x, p_{max_1}.y$);
/* Calcular longitud de la sub trayectoria No. 2 */
length2 := distance($p_{min_2}.x, p_{min_2}.y, p_{max_2}.x, p_{max_2}.y$);
/* Calcular SDist */
SDist := $(1 - Sim) * \beta * (length1 + length2)$;

Figura 8 – Pseudocódigo para calcular SDist.

3.3. Algoritmo DB-SCAN para trayectorias.

Por último, se aborda el algoritmo DB-SCAN orientado hacia trayectorias el cual necesita una serie de pasos previos para poder ser ejecutado e implementado, entre ellos los siguientes procedimientos: Calcular el vecindario de una sub trayectoria (Calcular Neps), determinar cuándo una sub trayectoria es un elemento núcleo y asignar cluster a una sub trayectoria.

Procedimiento principal DB-SCAN para trayectorias

En este procedimiento se hace un llamado de forma lógica a los procedimientos de partición de trayectorias en puntos característicos, cálculo de vecindario, asignación de cluster, determinar si una sub trayectoria es núcleo, entre otros; este proceso retorna un conjunto de clusters de trayectorias, donde cada uno de ellos representa un patrón encontrado, como se presenta en la Figura 9.

Algorithm 9: DB-SCAN Principal
Input: Arreglo de sub trayectorias D
Input: Limite para el vecindario de la sub trayectoria tr_i - eps
Input: Mínimo número de trayectorias por cluster - $MinTrs$
Output: Arreglo de clusters - O
c_id := 0;
/* Se transforma D en puntos */
pointsflines := Segments2spt_pos(D);
/* Se extraen los puntos característicos de D */
$cp := approximate_traj_partition(pointsflines)$;
/* Se reconstruyen los segmentos basados en los puntos característicos */
$linesfcp := Construct_Segments(cp)$;
/* Se recorre el arreglo de segmentos $linesfcp$ */
$pos1 := linesfcp_1$;
while $pos1$ *is not null* **do**
  /* Se verifica cuales segmentos estan clasificados */
  if $linesfcp_{pos1}.classified = 0$ **then**
    /* Se calcula el vecindario del segmento */
    Ngh := Calculate_Neps(linesfcp,pos1,Eps);
    /* Se verifica si el segmento es núcleo */
    is_core := IsCore(Ngh,MinTrs);
    if *is_core* **then**
      c_id := c_id + 1;
      /* Se asigna el vecindario del segmento al cluster correspondiente */
      c1 := AssignCluster(Ngh,linesfcp,c_id);
      Agregar $c1$ a O;
    end
  end
  $pos1 := linesfcp_{pos1+1}$;
end

Figura 9 – Pseudocódigo procedimiento principal algoritmo DB-SCAN para trayectorias.

4. Implementación

La implementación de la mencionada metodología se realizó en el motor de base de datos Oracle versión 11g utilizando la extensión Hermes, que brinda capacidades de almacenamiento y manipulación de datos de objetos en movimiento (bases de datos de objetos en movimiento) a la base de datos aumentando los tipos de datos abstractos (Moving Point, SpatioTemporal Point, etc.), procedimientos y funciones que además permiten la integración con la extensión espacial de Oracle – Oracle Spatial. Se implementaron un total de 13 funciones bajo un mismo paquete y 2 procedimientos

almacenados para lograr el funcionamiento de la metodología propuesta en el presente trabajo, las cuales son llamadas por una función denominada Main, encargada de realizar llamados a los otros procedimientos. Todos los procesos se implementaron utilizando la extensión Hermes. El resultado es una tabla donde se indica el cluster, coordenadas de inicio y coordenadas finales de cada una de las sub trayectorias procesadas. A continuación, se presenta un ejemplo de la visualización en software SIG (ArcGIS) del resultado del algoritmo:

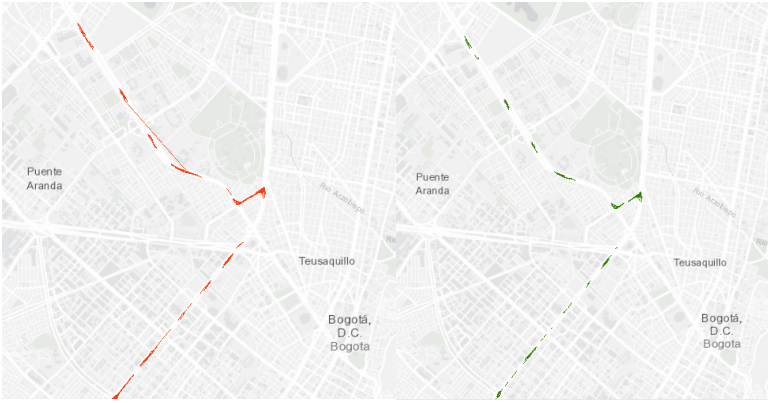


Figura 10 – Muestra de cluster detectados por el algoritmo.

El conjunto de datos sobre el cual se realizaron las pruebas de la implementación corresponde a las trayectorias descritas por una ruta (G43-K43) el día 28 de febrero de 2018 del sistema de buses articulados Transmilenio de la ciudad de Bogotá, el algoritmo hallo 30 clusters de los cuales se muestran 2 en la Figura 10, los cuales corresponden al cluster 2 y 10, que representan lugares donde tanto en espacio como en tiempo existen unos patrones de movimiento durante el trazado de la ruta.

5. Conclusiones

Se evidencia que existen algoritmos que involucran datos de trayectorias o de objetos en movimiento que solo se presentan en forma de planteamiento, el trabajo en el presente artículo presenta la plataforma sobre la cual se pueden implementar para que sean probados y utilizados en situaciones de la vida real utilizando el marco de las bases de datos de objetos en movimiento.

Se aprovechó toda la cantidad de tipos de datos y procedimientos que contiene HERMES para el manejo de datos de objetos en movimiento representados en trayectorias, lo cual brinda un margen de ejecución para un gran número de tareas de analítica y minería de datos, ya que en la actualidad el mercado y la academia van enfocados hacia el descubrimiento de información que no se halla de forma explícita en los datos.

Así como se muestra el pseudocódigo de los distintos algoritmos, también se cuenta con la implementación de los mismos en la base de datos Oracle teniendo como soporte las extensiones de Oracle Spatial y Hermes, lo cual permitirá en un futuro utilizarlos los

procedimientos implementados para la ejecución de tareas de minería de datos sobre trayectorias que ha sido un campo poco explorado en los últimos años.

Referencias

Encyclopædia Britannica. (2017). Motion. Obtenido de http://academic.eb.com.bdigital.udistrital.edu.co:8080/levels/collegiate/article/motion/53959

Grünwald, P. D., Myung, I. J., & Pitt, M. A. (2005). Advances in Minimum Description Length: Theory and Applications. Cambridge. MA, USA: Massachussetts Institute of Technology.

Güting, R. H., & Schneider, M. (2005). Moving Objects Databases. San Francisco, USA: Morgan Kaufmann Publishers.

Han, J., Kamber, M., & Pei, J. (2012). Data Mining Concepts and Techniques. Waltham, MA, USA: Morgan Kaufmann Publishers.

Jae-Gil, L., Han, J., & Whang, K.-Y. (2007). Trajectory Clustering: A Partition-and-Group Framework. SIGMOD'07. Beijing, China.

Li, Z., Ji, M., Lee, J.-G., Tang, L.-A., Yu, Y., Han, J., & Kays, R. (2010). MoveMine: Mining Moving Object Databases. SIGMOD. Indianapolis, Indiana.

Panagiotakis, C., Pelekis, N., Kopanakis, I., Ramasso, E., & Theodoridis, Y. (2012). Segmentation and Sampling of Moving Object Trajectories Based on Representativeness. IEEE Transactions on Knowledge and Data Engineering, 1328–1343.

Pelekis, N., & Theodoridis, Y. (2014). Mobility Data Management and Exploration. New York: Springer.

Pelekis, N., Frentzos, E., Giatrakos, N., & Theodoridis, Y. (2011). HERMES: A Trajectory DB Engine for Mobility-Centric Applications. International Journal of Knowledge. based Organizations.

Pelekis, N., Frentzos, P., Giatrakos, N., & Theodoridis, Y. (2010). Supporting Movement in ORDBMS-the HERMES MOD Engine. Technical Report Series, UNIPI-INFOLAB-TR-2010-01.

Ying, X., Xu, Z., & Guo Yin, W. (2009). Cluster-based Congestion Outlier Detection Method on Trajectory Data. Sixth International Conference on Fuzzy Systems and Knowledge discovery, (págs. 243–247)

Yu, Z. (2015). Trajectory Data Mining: An Overview. ACM Transactions on Intelligent Systems and technology, 6(3).

Zhou, X., Zheng, K., Jueng, H., Xu, J., & Sadiq, S. (2015). Making sense of spatial trajectories. CIKM'15, (págs. 671–672). Melbourne.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 23/10/2018
Aceitação/Acceptance: 04/12/2018

Pruebas de concepto automatizado sobre aplicaciones web utilizando Raspberry Pi

Geraldín Vergara Fajardo[1], Diana Marcela Montaño[2], Siler Amador Donado[3], Katerine Márceles Villalba[4]

geralvefa@unimayor.edu.co[1], diana_montano@unimayor.edu.co[2], samador@unicauca.edu.co[3], kmarceles@unimayor.edu.co[4]

[1, 2, 4] Facultad de Ingeniería, Institución Universitaria Colegio Mayor del Cauca, Carrera 7 N° 2-34 Edificio Bicentenario, 190003, Popayán- Cauca, Colombia.

[3] Facultad de Ingeniería Electrónica y Telecomunicaciones, Universidad del Cauca, Carrera 2 #4N-140, 190003, Popayán- Cauca, Colombia

Pages: 648–659

Resumen: En la actualidad se ha identificado que existen organizaciones que carecen de aplicaciones que le permitan automatizar el proceso de pentesting y que utilicen una metodología de seguridad de la información, como es el OWASP 2017; dado a lo anterior este proyecto cuenta con un plan de trabajo para llegar a la creación de un diseño automatizado de pruebas de concepto web que lleve consigo scripts de seguridad que permitan detectar vulnerabilidades de los sitios web; todo esto bajo la utilización de un Raspberry Pi, para conseguir un beneficio no solo en lo técnico y económico sino también conseguir que todo este proceso se realice de manera portable y automatizada.

Palabras-clave: OWASP, Pentesting, Seguridad, Raspberry Pi, Scripts.

Automated concept testing on web applications using Raspberry Pi

Abstract: At present it has been identified that there are organizations that lack applications that allow them to automate the pentesting process and that use an information security methodology, such as the OWASP 2017; Given this, this project has a work plan to arrive at the creation of an automated design of web concept tests that carries with it security scripts that allow the detection of vulnerabilities of the websites; all this under the use of a Raspberry Pi, to achieve a benefit not only in the technical and economic but also to ensure that all this process is done in a portable and automated.

Keywords*:* OWASP, Pentesting, Security, Raspberry Pi, Scripts.

1. Introducción

En la actualidad unos de los principales inconvenientes presentados en los sitios web ha sido su seguridad, esto se ha venido evidenciando tanto en medios de comunicación

como también en redes sociales que ninguna es 100% libre de vulnerabilidades y ataques, y que aún existiendo herramientas disponibles para disminuir o prever éstas, toma tiempo hacer uso de cada una de dichas herramientas de acuerdo al análisis requerido. Por tanto, una manera de minimizar riesgos estaría dada por la realización de un pentesting, el cual debe cumplir una función de evaluación de seguridad de alto nivel, siguiendo una metodología que deberá ejecutarse durante la realización de cualquier programa de auditoría en seguridad de la información. Con esto se lograría ayudar a las empresas desarrolladoras de software a identificar vulnerabilidades de acuerdo al TOP 10 de OWASP 2017, durante la etapa de desarrollo de las aplicaciones web.

El artículo se encuentra organizado de la siguiente manera: en la sección dos están los trabajos previos como los conceptos claves del trabajo, en la sección tres el desarrollo del proyecto, en la sección cuatro experimentos y resultados, finalmente en la quinta sección las conclusiones y trabajos futuros.

2. Teoria del domínio y trabajos relacionados

En esta sección se citan algunos de los trabajos que fueron tomados como referentes para la realización de este proyecto:

Desarrollo de una propuesta metodológica para determinar la seguridad en una aplicación web, presentada por los autores; Ascencio Mendoza, M. Moreno Patiño, P. Trabajo de investigación que se realizó con el fin de elaborar una herramienta metodológica que le permitió determinar que vulnerabilidades posee una aplicación web, también una herramienta útil que mantuvo a la vanguardia a los ingenieros informáticos de las nuevas exigencias de mercado. En este trabajo se analizaron los conceptos y fundamentos de seguridad en las aplicaciones web, se determinó los tipos de vulnerabilidades en las aplicaciones web, se establecieron que herramientas existen actualmente para la detección de vulnerabilidades en las aplicaciones web, y por último se determinaron las características principales de algunas metodologías existentes para la seguridad en aplicaciones web. (Ascencio Mendoza & Moreno Patiño, 2011)

Penetration Testing with Raspberry Pi, con sus autores Lakhani, A. Muniz, J. Construyeron un libro donde realizaron pruebas de penetración utilizando el sistema operativo Kali Linux y el SBC Raspberry Pi, el enfoque principal de este libro era aprender cómo combinar el poder de Kali Linux con la portabilidad y bajo costo de Raspberry Pi. El resultado de este fue una plataforma de pruebas de penetración extremadamente flexible para proyectos específicos que no requieren aplicaciones con necesidades de alta potencia de procesamiento. Utilizaron un conjunto de herramientas para llevar a cabo el análisis de penetración y las pruebas de vulnerabilidad de las posiciones remotas, se utilizó la portabilidad de Raspberry Pi a prueba de evaluación de seguridad de forma encubierta en distintos lugares, y se configuró éste para ser gestionados de forma remota. El uso de Kali Linux sobre un Raspberry Pi permitió proporcionar una prueba de intrusión con una opción única y rentable para lograr los objetivos de prueba. (Lakhani & Muniz, 2015).

Propuesta de implementación de una metodología de auditoria de seguridad informática, con su autor López Santoyo, R; realizó un trabajo de grado que buscaba

elegir la metodología que tuviera las mejores perspectivas para la ciberseguridad y de esta manera implementarla; en este trabajo se analizó cada una de las metodologías con el objetivo de compararlas evaluando así cuales eran sus ventajas e inconvenientes de cada una de ellas respecto a las demás, de esta manera se eligió la metodología más apropiada para realizar un análisis de seguridad completo, en todos sus posibles ámbitos sobre una infraestructura TI(Tecnología de información) de una gran organización, y por último se implementó la metodología OSSTMM (Sirve para evaluar la seguridad de localizaciones físicas, interacciones humanas y todas las formas de comunicación existentes, esta metodología se basa en los objetivos a analizar, como se realizan las pruebas sobre estos, los controles de seguridad encontrados y que es lo que no se ha probado. El objetivo de esta metodología es la de responder las preguntas del STAR (Security Test Audit Report)), indicando las herramientas que podían utilizarse, donde utilizarlas y donde localizar los recursos necesarios para llevarlas a cabo. Esta guía pretendía facilitar el "qué" debe hacerse atendiendo a la metodología y el "cómo" se debe ejecutar en la práctica. (López Santoyo, 2015)

IDS on Raspberry Pi, Trabajo de diplomado realizado por Aspernas, A. Simonsson, T. En éste se realizó un estudio comparativo de Raspberry Pi 2 modelo B y Raspberry Pi 2 modelo B+ donde estos actuaban como un IDS (Intrusion Detection System) en una red doméstica, implementando Arch Linux ARM, donde el objetivo principal era examinar el efecto sobre el rendimiento de la red con el uso de estos computadores de bajo costo, se observaron las capacidades para manejar el tráfico de red mientras se examinaba el tráfico de intrusión, también se midieron el uso de la CPU y la memoria en los dispositivos para detectar las posibles limitaciones de hardware, de esta manera se logró concluir que Raspberry Pi se puede utilizar como un IDS pero solo bajo ciertas circunstancias. (Aspernas & Simonsson, 2015).

2.1. Aportes de los antecedentes al proyecto

Uno de los principales aportes fue el análisis de las vulnerabilidades web; lo cual permitió determinar para este trabajo de investigación algunos dominios a implementar para la detección de dichas vulnerabilidades; también contribuyó con la interacción del Raspberry Pi con el sistema operativo Kali Linux en la realización de pruebas de penetración y pruebas de vulnerabilidad, en el cual resulto muy rentable el manejo en conjunto de ambos; por otra parte se dio una visión de las capacidades de utilizar sobre el Raspberry pi; como el uso para el manejo de trafico de red y tráfico de intrusión, asimismo se midió el rendimiento del dispositivo en cuanto al uso de CPU y de memoria para determinar ciertas limitaciones de hardware que pudieran afectar el uso de algunas herramientas en este proyecto. Por último se observó la comparación de las metodologías dirigidas a la seguridad de la información para establecer la mejor metodología que permitiera la realización de análisis de penetración a aplicaciones web, en la cual se estableció la metodología OWASP, por lo que está dirigida específicamente ha aplicaciones web.

3. Desarrollo del proyecto

Para el cumplimiento de la primera parte del proyecto del porqué emplear el dispositivo Raspberry pi?, se inició con una recopilación de información necesaria sobre una

cantidad de SBC´s, mediante un proceso de indagación y exploración, en el cual se recolectaron 46 SBC´s, a partir de ello se definieron siete indicadores y se realizó un proceso de normalización, para la selección del SBC que se ajustara más al proyecto, donde resulto que el SBC a escoger seria el Raspberry Pi 3 Model B, los criterios que lo determinaron fueron: Frecuencia, RAM, puertos USB, puerto Ethernet, Wi-Fi, precio del SBC y si el dispositivo soportaba el sistema operativo Kali Linux.

1. Para la frecuencia de CPU, se estableció un 20% del total de los indicadores, debido a que para la ejecución del pentesting se necesitaba velocidad, para reducir tiempo como se justificó en el proyecto.
2. La RAM representó el 3%, dado que la mayoría de los SBC´s cuentan con la RAM suficiente para cargar un sistema operativo.
3. Los puertos USB se le establecieron un 5%, puesto que se utilizaron periféricos como teclado y mouse para el manejo del SBC.
4. El puerto de Ethernet representó el 15%, debido a que era primordial conectarse a internet y realizar el pentesting a sitios web en tiempo real.
5. El WI-FI integrado determino el 5% de los indicadores, porque si no se cuenta con conexión LAN era necesario hacer uso del servicio WI-FI para la conexión a internet para la ejecución del pentesting.
6. El precio del SBC representó el 37%, dado a que se requiere un SBC de un precio bajo, en vista de que es el enfoque primordial del proyecto.
7. El sistema operativo estableció el 15% de los indicadores, debido a que se va a utilizar una distribución basada en Debian GNU/Linux llamada Kali Linux, dado que es una herramienta diseñada principalmente para realizar pruebas avanzadas de penetración y auditoria de seguridad (Kali Linux, 2018), y además es de uso libre lo que resulta en una reducción de costos del proyecto.

Por otro lado para la realización de esta investigación también se analizaron 17 herramientas, que se muestran en la Figura 1, destinadas al pentesting recomendadas por OWASP, donde se observaron algunas características en particular para ver su funcionamiento y se determinó que ninguna de las herramientas efectua un pentesting completo, sino más bien se encargan de analizar algunas vulnerabilidades por separado dando en ocasiones resultados erróneos; y provocando de esta manera la ejecución de todas estas herramientas para tener un pentesting completo, lo que ocasiona que se le dedique más tiempo del esperado, para poder escoger la información valiosa para un proceso de prueba de concepto.

Es de resaltar que a pesar de que la herramienta OWTF, se encargue de realizar una prueba de concepto basado en la guía 4.0 y el top 10 de OWASP, utiliza muchos recursos de la máquina, debido al uso de más de 100 plugins, ejecutandose éstos por partes, es decir tiene sesiones diferentes para cada control, mientras que este proyecto une los controles en una sola sesión, donde solo debes colocar el target una vez, y esperar los resultados, y lo más importante sobre un Raspberry Pi 3 Modelo B, que no solo baja los costos de una prueba de concepto que es equivalente a un pentesting sino también busca la portabilidad.

NO.	NOMBRE	VULNERABILIDADES QUE DETECTA	VELOCIDAD DE DETECCIÓN DE RESULTADOS	LEGIBILIDAD DEL REPORTE FINAL	MULTIPLATAFORMA	OPENSOURCE Y FREEWARE
1	Antisamy	XSS	ALTA	BAJA	SI	SI
2	CAL9000	XSS, http	ALTA	BAJA	SI	SI
3	Code pulse	pruebas caja negra	MEDIA	BAJA	SI	SI
4	Dependency track	uso componente, XSS, inyección SQL	ALTA	BAJA	SI	SI
5	DirBuster	fuerza bruta	ALTA	BAJA	NO	SI
6	Encoding	Técnicas XSS	ALTA	MEDIA	SI	SI
7	JBroFuzz	Inyección SQL, XSS	ALTA	BAJA	SI	SI
8	LAPSE	Inyección SQL, XSS	ALTA	BAJA	SI	SI
9	O-Saft	http, vulnerabilidades SSL	ALTA	BAJA	NO	SI
10	OWTF	sesiones de alojamiento virtual	ALTA	MEDIA	SI	SI
11	Pantera Web Assessment studio	archivos XML para configuración y ataques	ALTA	ALTA	SI	SI
12	SQLiX	Inyección SQL	ALTA	BAJA	SI	SI
13	WebGoat	Inyección SQL, XSS, sesión insegura, comentarios HTML	ALTA	MEDIA	SI	SI
14	WebScarab	http, sesión, cookies, XSS, CRLF	ALTA	BAJA	SI	SI
15	WSFuzzer	http, XSS	MEDIA	BAJA	SI	SI
16	Xenotix XSS	XSS, falsos positivos	ALTA	MEDIA	SI	SI
17	ZAP	Sistemas de autenticación, XSS, fuerza bruta	ALTA	BAJA	SI	SI

Figura 1 – Herramientas sugeridas por Owasp.

Es importante mencionar que hasta el momento se han creado veintidós scripts que corresponden a siete de los ataques del Top 10 de OWASP 2017 y de los cuales se han realizado pruebas a un sitio web que la universidad ha autorizado.

4. Experimentos y Resultados

De la normalización de los 46 SBC´s se seleccionaron los primeros cinco SBC´s, que obtuvieron mayor puntaje, discriminando sus respectivas características, como se referencia en la Figura 2:

NOMBRE	CPU	RAM	USB	ETHERNET	WI-FI	OS	PRECIO	PUNTAJE
Nano Pi M3	1.4GHz	1GB	SI	SI	SI	SI	Aprox. 90.000 COP	0.920
Raspberry Pi 3 Model B	1.2GHZ	1GB	SI	SI	SI	SI	Aprox. 138.000 COP	0.877
Raspberry Pi Zero W	1GHz	0.512GB	SI	SI	SI	SI	Aprox. 60.000 COP	0.869
Odroid XU3 Lite	1.8 GHz	2GB	SI	SI	NO	SI	Aprox. 343.832 COP	0.852
CuBox	0.8GHz	1GB	SI	SI	NO	SI	Aprox. 90.450 COP	0.790

Figura 2 – Clasificación de SBC con mayor puntaje a partir de criterios establecidos.

Para la elección del SBC final se sumaron dos criterios más que fueron de ayuda para tener el SBC adecuado para el proyecto, los cuales son la asequibilidad y la comunidad de dicho dispositivos en la red, donde se determinó que el dispositivo más asequible en la ciudad de Popayán era el Raspberry Pi 3 model B y contaba con una comunidad de soporte bastante significativa en internet; además de ser un dispositivo utilizado en varios proyectos enfocados a la seguridad de la información como se puede notar en los antecedentes ya mencionados.

Al seleccionarse el dispositivo SBC, se determinaron las pruebas a realizarse para la automatización de la prueba de concepto, que se muestran de la Tabla 1 a la Tabla 8, las cuales se eligieron a partir de un proyecto ya realizado (Peña Torres, 2016) donde se obtuvieron las más representativas que suplieran cada ataque del Top 10 de OWASP:

A1- Inyección	
OTG-INPVAL-005	Prueba de Inyección SQL
	Oracle Testing
	Prueba de MySQL
	Prueba de SQL Server
	Probando PostgreSQL
OTG-INPVAL-006	Prueba de Inyección LDAP
OTG-INPVAL-012	Prueba de Inyección de Código
OTG-INPVAL-013	Prueba de Inyección de Comando

Tabla 1 – Riesgo A1

A2- Autenticación rota	
OTG-AUTHN-003	Prueba de mecanismo de bloqueo débil

Tabla 2 – Riesgo A2

A3- Exposición a datos sensibles	
OTG-INFO-005	Revise los comentarios de la página web y los metadatos para la fuga de información
OTG-CRYPST-001	Prueba de cifrado SSL / TSL débil, protección insuficiente de la capa de transporte

Tabla 3 – Riesgo A3

A4- Entidades externas 2017-XML (XXE)	
OTG-INPVAL	Prueba de inyección XML

Tabla 4 – Riesgo A4

A5- Control de acceso roto	
OTG-AUTHZ-002	Prueba para eludir el esquema de autorización

Tabla 5 – Riesgo A5

A6- Mala configuración de seguridad	
OTG-CONFIG-001	Prueba de red / configuración de infraestructura
OTG-CONFIG-002	Configuración de plataforma de aplicación de prueba
OTG-CONFIG-005	Enumerar interfaces de administración de aplicaciones y de infraestructura
OTG-CONFIG-006	Prueba de métodos HTTP
OTG-CONFIG-007	Pruebe la seguridad de transporte estricta de HTTP

Tabla 6 – Riesgo A6

A7- Cross-site scripting	
OTG-INPVAL-001	Prueba de secuencias de comandos cruzadas reflejadas en el sitio

Tabla 7 – Riesgo A7

A9- Uso de componentes con vulnerabilidades conocidas	
OTG-INFO-002	Servidor web de huellas digitales
OTG-INFO-007	Mapa de rutas de ejecución a través de la aplicación
OTG-INFO-008	Marco de aplicación web de huellas dactilares
OTG-INFO-009	Aplicación web de huellas dactilares
OTG-INFO-010	Arquitectura de aplicaciones de mapas

Tabla 8 – Riesgo A9

Para las vulnerabilidades correspondientes al A8 y A10 del Top Ten de OWASP 2017, no se realizaron scripts debido a que a lo largo de la investigación se encontró que estas son de difícil automatización ya que se ve la necesidad de realizar cambios o ajustes específicos en el código de explotación subyacente para cada aplicación web, según sea la estructura de la misma y de esta manera se logre una prueba de concepto satisfactoria, ya que si se realizan scripts generales para todas las aplicaciones web estos podrían arrojar falsos positivos (resultados equívocos) o no generar resultados al realizarse las pruebas, y es aquí donde se necesita de asistencia humana para validar los problemas.

Por último, se han realizado las pruebas mencionadas, donde arroja información en la consola del sistema operativo y también en archivos que al final de este proyecto de investigación serán interpretados, para así generar un reporte técnico y ejecutivo de la prueba de concepto realizada.

La prueba de concepto como todo proyecto en Python tiene un comando de ayuda para poder entender el manejo del mismo, el cual se puede observar en la Figura 3, donde aparecen tres banderas para la realización de la prueba de pentesting.

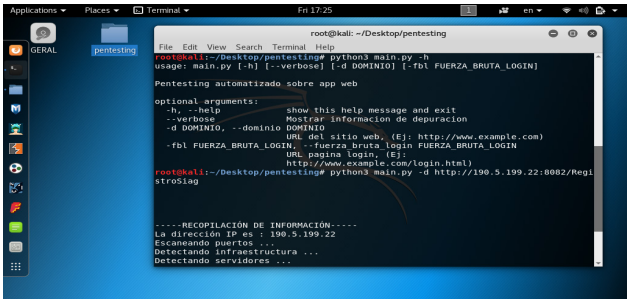


Figura 3 – Ayuda del Pentesting.

Utilizando la bandera *–d* y el sitio web a analizar se realizan las pruebas que no hacen referencia a fuerza bruta, en la primera parte muestra la recopilacion de la informacion y

spidering del sitio web como se observa en la Figura 4, y a partir de estas pruebas se arroja la informacion en archivos como se muestra en la Figura 5, los cuales hacen referencia a las pruebas OTG-INFO, que es practicamente la parte de escaneo del sitio web.

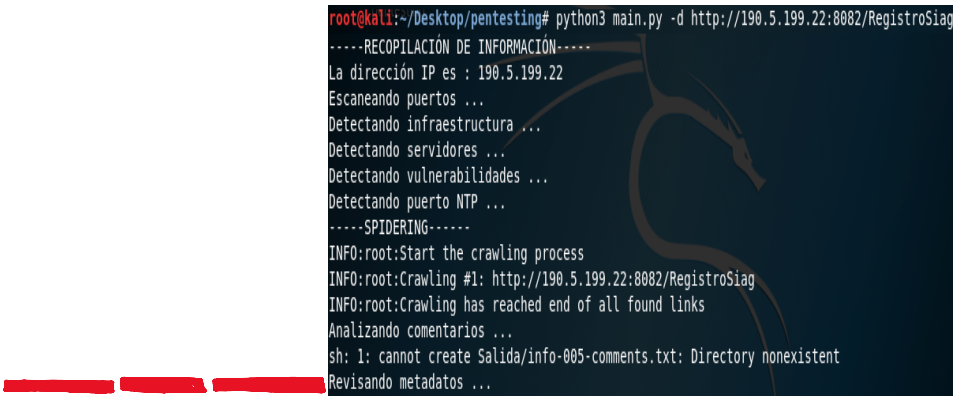


Figura 4 – Recopilacion de Informacion y spidering del sitio asignado (Se protege la dirección web por seguridad).

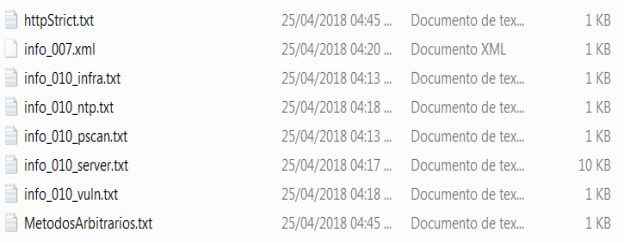


Figura 5 – Archivos de las pruebas OTG-INFO.

En la Figura 6, se ejecuta la prueba OTG-CONFIG-006, la cual tiene como objetivo descubrir si el sitio web soporta los metodos HTTP como: HEAD, GET, POST, PUT, DELETE, TRACE, OPTIONS y CONNECT.

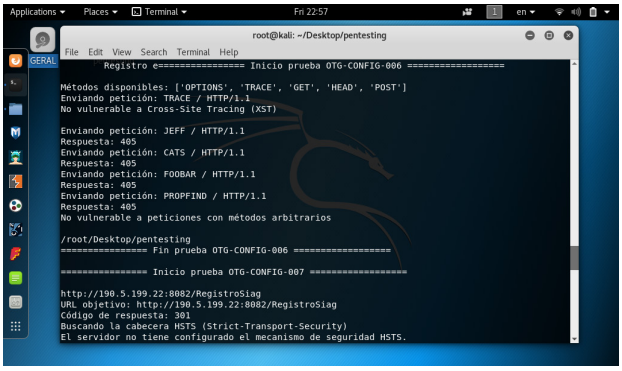


Figura 6 – Ejecucion Prueba OTG-CONFIG-006.

En la Figura 7, se muestra el resultado de la prueba OTG-CONFIG-007 donde se verifica que el sitio web utilice el encabezado HTTP Strict Transport Security, para garantizar que todos los datos viajen cifrados desde el navegador al servidor. Y tambien se muestra la prueba OTG-CRYPST-001, donde el objetivo es comprobar que tipo de canales se utilizan para el envio de información, esta prueba solo se puede realizar si el sitio web tiene el puerto 443 abierto, de lo contrario no es vulnerable.

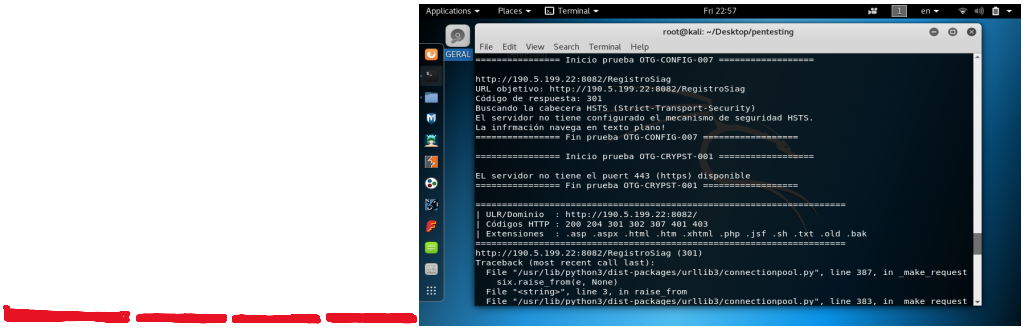


Figura 7 – Ejecucion prueba OTG-CONFIG-007 y Prueba OTG-CRYPST-001.

En la Figura 8, se observa que la prueba busca direcciones enlazadas al dominio principal con el fin de obtener de alguna de ellas un acceso del sistema.

Cada uno de los códigos HTTP tiene un significado, como no autorizado, sin contenido, pero el que resalta aquí es el código 200, que indica una posibilidad de acceso e información en esa dirección.

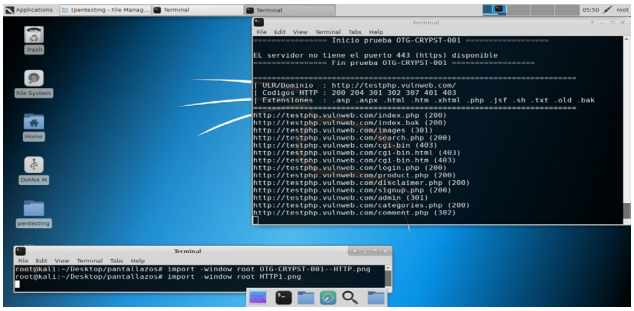


Figura 8 – Ejecución prueba OTG-AUTHZ-002.6

Como lo muestra la Figura 9 se utilizan diccionarios para usuario y contraseñas con el fin de encontrar un posible acceso mediante fuerza bruta, los datos resultantes muestran cada usuario intentando acceder con una contraseña hasta terminar el listado y donde finalmente se evidencia su éxito, resaltando el usuario y contraseña con el cual se puede ingresar al sistema.

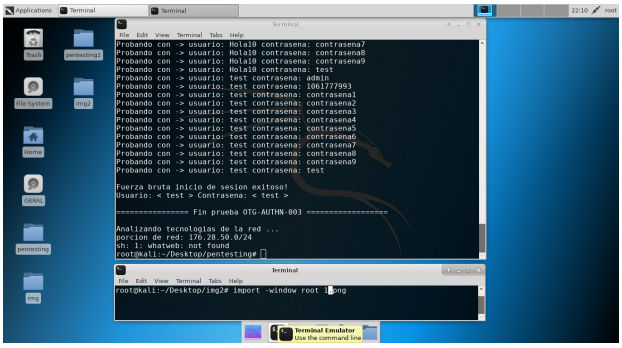


Figura 9 – Ejecución prueba OTG-AUTHN-003.

5. Conclusiones y trabajos futuros

Debido a la inexistencia de una aplicación que automatice el proceso de pentesting y que utilice una metodología de seguridad de la información, en este caso OWASP 2017, el proyecto contó con un plan de trabajo para lograr la creación de un diseño de pruebas de concepto automatizado de vulnerabilidades web que llevará consigo scripts de seguridad que permitieran detectar los ataques informáticos que en los sitios web se presentan, todo esto bajo la utilización de un SBC de bajo costo. A continuación en la figura 10. Se muestra el diseño estructural de la automatización del los scripts.

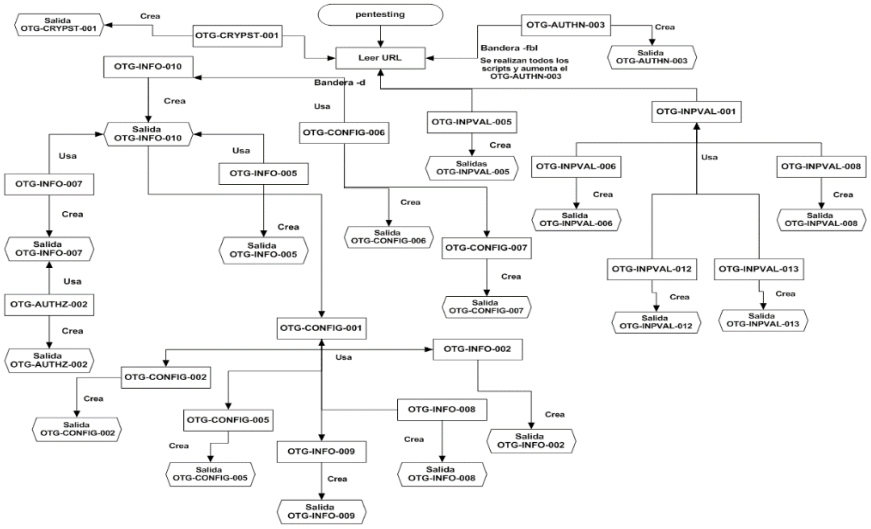


Figura 10 – Diagrama automatización de la prueba de concepto.

Como resultado de este proyecto de investigación se destaca la existencia de los diferentes dispositivos para la realización de pruebas de concepto de manera portable, en este caso

el Raspberry pi 3 Modelo B, donde se pueden realizar pruebas de penetración de manera conjunta por medio de scripts, minimizando así el proceso de búsqueda de herramientas por separado dedicadas a buscar vulnerabilidades referentes al top 10 de OWASP 2017. Todo esto tiene como finalidad identificar los riesgos o vulnerabilidades que se pueden estar presentando.

Se puede destacar también que el uso de los scripts realizados para este proyecto investigativo en el momento son de manera gratuita, lo que es de importancia para empresas que necesiten un análisis de penetración, y de esta manera ahorrar económicamente lo que les costaría efectuar estas pruebas con las herramientas que existen en la actualidad, además de que estas no están enfocadas a ninguna metodología de seguridad sino más bien a manera general, lo que con la prueba de concepto automatizada que se realizó si se puede notar, ya que está alineada a la metodología OWASP.

Cabe añadir que el uso de un dispositivo de bajo costo para la automatización de las pruebas no es suficiente, debido a que en la ejecución de los scripts tarda mucho tiempo en dar respuesta en sitios web por demasiada información. La continuación de este proyecto podría llevar a cabo la automatización de todas las pruebas de la guía cuatro de OWASP y la generación de reportes automáticos, con el uso de un Clúster de Raspberry pi 3 Modelo B, para que se compartan los recursos de los dispositivos y sea más rápida la ejecución y salida de resultados de los scripts a realizar.

Para finalizar, es recomendable utilizar la guía de desarrollo seguro de OWASP, para mitigar las vulnerabilidades en aplicaciones web antes de que estas salgan a producción y se enfrenten a las amenazas de atacantes maliciosos en internet.

Agradecimientos

Agradecimientos a Innovación Cauca, a la Universidad del Cauca en especial a su grupo de investigación GTI y al semillero Beta Bit de la Facultad de Ingeniería de la Institución Universitaria Colegio Mayor del Cauca, por el apoyo brindado para el desarrollo del proyecto.

Referencias

Ascencio mendoza, m., & moreno patiño, p. (2011). Desarrollo de una Propuesta Metodológica para Determinar la Seguridad en una Aplicación Web. Obtenido de http://repositorio.utp.edu.co/dspace/bitstream/handle/11059/2511/0058A811.pdf;jsessionid=9E3CDC5C58DFB673220AA28275CDB018?sequence=1

Aspernas, A., & Simonsson, T. (2015). IDS on Raspberry Pi. Computer Science.

Darknet. (2016). OWASP OWTF – Offensive Web Testing Framework. Recuperado de: https://www.darknet.org.uk/2016/09/owasp-owtf-offensive-web-testing-framework/

Kali Linux. (2018). Obtenido de DOCUMENTACIÓN OFICIAL: https://docs.kali.org/introduction/what-is-kali-linux

Klosowski, T. (3 de marzo de 2016). GIZMODO. El rendimiento de la Raspberry Pi 3, comparado contra los modelos anteriores. Obtenido de http://es.gizmodo.com/el-rendimiento-de-la-raspberry-pi-3-comparado-contra-l-1762657409

Lakhani, A., & Muniz, J. (2015). Penetration Testing with Raspberry Pi. Packt Publishing Ltd.

López Santoyo, R. (Junio de 2015). Propuesta de Implementación de una Metodologia de Auditoria de Seguridad Informatica. Obtenido de https://repositorio.uam.es/bitstream/handle/10486/668900/Lopez_Santoyo_Roberto_tfg.pdf?sequence=1

OWASP. (2017). Guía de Pruebas 4.0. Quito.

OWASP. (18 de enero de 2018). Open Web Application Security Project. Obtenido de Category: OWASP Project: https://www.owasp.org/index.php/Category:OWASP_Project

Peña Torres, I. (06 de Junio de 2016). Auditoría Técnica de Seguridad. Máster Universitario en Seguridad de las Tecnologías de la. Universitat Oberta de Catalunya.

Perafán Ruiz, J. J., & Caicedo Cuchimba, M. (2014). Análisis de Riesgos de la Seguridad de la Información para la Institución. Popayán: Universidad Nacional Abierta y a Distancia.

Pino, F., Piattini, M. & Travassos, G. (2013) Gestión y desarrollo de proyectos de investigación distribuidos en ingeniería del software por medio de investigación-acción. Revista Facultad de Ingeniería de Antioquia.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 11/09/2018
Aceitação/Acceptance: 03/11/2018

Explorando el poder predictivo del modelo TAM: una comparación empírica de la estimación basada en CBSEM, PLS y ANN

Felipe Machorro-Ramos[1], Patricio Ramírez-Correa[2], Giselle Painen-Aravena[2], Jorge Alfaro-Peréz[2], Ari Melo Mariano[3]

felipe.machorro@udlap.mx, patricio.ramirez@ucn.cl, giselle.painen@alumnos.ucn.cl, jalfaro@ucn.cl, arimariano@unb.br

[1] Universidad de las Américas Puebla, Ex Hacienda Sta. Catarina Mártir S/N, 72810, San Andrés Cholula, México.

[2] Universidad Católica del Norte, Larrondo 1281, 1781421, Coquimbo, Chile.

[3] Universidad de Brasilia, Campus Darcy Ribeiro, 70904970, Brasilia, Brasil.

Pages: 660–670

Resumen: El objetivo de este estudio es comparar el poder predictivo de tres diferentes estrategias de estimación (Ecuaciones Estructurales basadas en Covarianzas CBSEM, Mínimos Cuadrados Parciales PLS y Redes Neuronales Artificiales ANN) para un modelo de aceptación tecnológica (TAM) en el contexto de una institución de educación superior en México. Los resultados obtenidos indican que las Redes Neuronales Artificiales tiene un mejor desempeño en la predicción que los Mínimos Cuadrados Parciales y estos a su vez que las Ecuaciones Estructurales Basadas en Covarianzas. Con esto se confirma el mejor poder predictivo de las ANN y el carácter confirmatorio de CBSEM.

Palabras-clave: Redes neuronales artificiales, Ecuaciones estructurales basadas en covarianzas, Mínimos cuadrados parciales, Modelo de Aceptación de Tecnología, México.

Exploring the predictive power of the TAM model: an empirical comparison of the estimation based on CBSEM, PLS and ANN

Abstract: The aim of this study is to compare the predictive power of three different estimation strategies (Covariance Based Structural Equations Models CBSEM, Partial Least Squares and Artificial Neural Networks ANN) for a Technological Acceptance Model (TAM) in the context of an institution of Higher Education in Mexico. The obtained results indicate that the Artificial Neural Networks have a better performance in the prediction than the Partial Least Squares and these in turn that CBSEM. This confirms the predictive power of the ANNs and the confirmatory nature of CBSEM.

Keywords*:* Artificial Neural Network, Covariance Based Structural Equation Model, Partial Least Squares, Technological Acceptance Model, Mexico.

1. Introducción

El constante cambio que caracteriza a la época actual ha traído como consecuencia un incremento sustancial del ritmo en el que nuevas tecnologías de la información se desarrollan y se ponen a disposición de los usuarios. Sin embargo, la implementación de cualquier tecnología resulta costosa en función del tiempo y los recursos que implica asegurar su éxito (Al-Emran, Mezhuyev, & Kamaludin, 2018). Es por ello que cada vez tiene más importancia el estudio de los factores que intervienen en la aceptación de estas nuevas tecnologías.

Las últimas tres décadas han visto surgir modelos teóricos útiles para estudiar la aceptación y el uso de las tecnologías de información - como la Teoría de la Acción Razonada (Fishbein & Ajzen, 1975), de la cual derivan propuestas como la Teoría de Comportamiento Planeado, el Modelo de Aceptación de Tecnología (TAM por sus siglas en inglés), y la Teoría Unificada de Aceptación y Uso de Tecnología – los cuales han sido validados y extendidos ampliamente en estudios académicos en todo el mundo (Davis, 1989; Venkatesh & Davis, 2000; Venkatesh, Morris, Davis, & Davis, 2003). A pesar de la existencia de esta amplia gama de modelos de aceptación de la tecnología, la propuesta TAM se ha convertido en el marco teórico más utilizado y citado en la literatura (Bazelais, Doleck, & Lemay, 2018), esto debido a su adaptabilidad, simplicidad y solidez, ya que ha recibido un amplio soporte empírico, demostrando su robustez en diferentes contextos (Al-Emran et al., 2018; Tarhini, Hone, & Liu, 2014b). El modelo TAM, desarrollado por Davis (1989), propone explicar las razones por la cuales aceptan o rechazan determinada tecnología por medio de dos factores principales: la utilidad percibida (PU por sus siglas en inglés) y la facilidad de uso percibida (PEOU por sus siglas en inglés). Ambas variables influyen a la intención de un comportamiento (BI por sus siglas en inglés) determinado, en este caso, el uso de un sistema de información.

Uno de los factores determinantes en el éxito del TAM como modelo de aceptación tecnológica es su poder predictivo respecto a la intención del comportamiento mientras se conserva la parsimonia del modelo (Ali, Raza, Qazi, & Puah, 2018; King & He, 2006). Sin embargo, se ha demostrado que la estrategia de estimación utilizada puede influir en el poder explicativo que brinda el modelo (López-Bonilla & López-Bonilla, 2017). Tradicionalmente, los estudios utilizan dos estrategias para estimar las relaciones entre las variables del TAM, los Modelos de Ecuaciones Estructurales Basado en Covarianzas (CBSEM por sus siglas en inglés) (Ali et al., 2018; Joo, Lee, & Ham, 2014; Park, 2009; Rafique et al., 2018; Tarhini et al., 2014b; Tarhini, Hone, & Liu, 2014a; Teo & Noyes, 2011) y los Mínimos Cuadrados Parciales (PLS por sus siglas en inglés) (Agudo-Peregrina, Hernández-García, & Pascual-Miguel, 2014; Al-Maroof & Al-Emran, 2018; Bazelais et al., 2018; Mokhtar, Katan, & Hidayat-Ur-Rehman, 2018; Ramirez-Anormaliza, Sabaté, Llinàs-Audet, & Lordan, 2017; Ramirez-Correa et al., 2015). Adicionalmente, en años recientes se han usado las Redes Neuronales Artificiales (ANN por sus siglas en inglés), como estrategia de estimación (Al-Shihi, Sharma, & Sarrab, 2018; Hew, Leong, Ooi, & Chong, 2016; Liébana-Cabanillas, Marinković, & Kalinić, 2017; Tan, Ooi, Leong, & Lin, 2014).

Henseler, Ringle, y Sinkovics (2009) afirman que la selección de la estrategia de estimación adecuada está en función de dos variables: el grado de desarrollo teórico de las relaciones existentes y el poder predictivo del modelo. De acuerdo con estos autores, los modelos

CBSEM son más adecuados para la confirmación de relaciones asumidas teóricamente, mientras que la estimación por PLS se destina al análisis predictivo en situaciones de baja información teórica. Por su parte, las ANN, se han posicionado como un mejor modelo predictivo que CBSEM y PLS; sin embargo, a pesar de existir una amplia literatura al respecto del TAM no se ha detectado algún estudio que haya estimado el modelo con las tres estrategias de estimación mencionadas y compare sus resultados.

De acuerdo con lo anterior, y derivado de la importancia que tiene la precisión en la predicción de los modelos de aceptación tecnológica, el presente estudio tiene como objetivo comparar el poder predictivo del modelo TAM estimado con las tres estrategias CBSEM, PLS y ANN. Para ello se utilizan datos del cuestionario TAM recolectados en estudiantes universitarios en México respecto de la aceptación de un sistema de gestión académica. El estudio se considera relevante en función de no haberse identificado en la literatura antecedentes respecto al análisis comparativos del poder predictivo del TAM en el contexto de la educación superior.

El resto del documento se estructura de la siguiente manera: por principio se revisa la literatura principal del modelo TAM aplicado en el contexto de la educación superior; posteriormente se describe la metodología empleada en el estudio. Seguido a esto se presentan y discuten los resultados obtenidos para finalmente exponer las principales las conclusiones del análisis realizado.

2. Revisión de literatura

2.1. El modelo TAM

El TAM ha demostrado ser un modelo teórico para ayudar a explicar y predecir el comportamiento del usuario de la tecnología de la información, se considera una extensión influyente de la teoría de la acción razonada (TRA), de Ajzen y Fishbein (1980). Davis (1989) propuso al TAM para explicar por qué un usuario acepta o rechaza la tecnología de la información al adaptar la TRA (Park, 2009).

En la formulación inicial del modelo se plantea la hipótesis de que dos variables: la utilidad percibida (PU) y la facilidad de uso percibida (PEOU) son los determinantes fundamentales de la actitud que se tenga hacia el uso de cualquier tecnología y a su vez esta influye al BI (Teo & Noyes, 2011). Sin embargo, los hallazgos empíricos posteriores apuntaron a la exclusión de la actitud como mediadora entre las creencias (PU y PEUO) y el comportamiento (BI), por lo que la PU y la PEOU se establecieron como antecedentes directos de BI (Davis, Bagozzi, & Warshaw, 1992; López-Bonilla & López-Bonilla, 2017).

2.2. Estrategias de estimación del TAM en educación superior

Los institutos de educación superior han logrado importantes oportunidades gracias al desarrollo de la tecnología de la información y la comunicación. Las universidades y otras instituciones educativas pueden reducir sus costos e incrementar los ingresos gracias a la implantación de diversas tecnologías. Es por ello que resulta importante estudiar los factores que influyen en la aceptación de estas tecnologías. En este contexto, el TAM ha sido utilizado en universidades de diversos países para evaluar la aceptación

de tecnologías aplicables a la educación, así como también se han utilizado distintas estrategias de estimación.

Para el caso de PLS, se identifican estudios de aceptación de sistemas de e-learning en contextos como las universidades en Madrid (Agudo-Peregrina et al., 2014), Chile y España (Ramirez-Correa et al., 2015), Ecuador (Ramirez-Anormaliza et al., 2017), Omán (Al-Maroof & Al-Emran, 2018) y Canadá (Bazelais et al., 2018). Con esta estimación también se han estudiado la aceptación de sistemas de aprendizaje por parte de instructores en Arabia Saudita (Mokhtar et al., 2018).

Los CBSEM, son otra estrategia de estimación popular para analizar el TAM en educación superior. Ejemplo de ello lo encontramos en aplicaciones para e-learning y m-learning para estudiantes universitarios de Pakistán (Ali et al., 2018), Inglaterra (Tarhini et al., 2014b), Corea (Joo et al., 2014; Park, 2009) y Singapur (Teo & Noyes, 2011), así como aplicaciones de bibliotecas móviles en Pakistán (Rafique et al., 2018).

Finalmente, también se han utilizado ANN en combinación con CBSEM para estimar la aceptación de sistemas de m-learning en Malasia (Tan et al., 2014) y Omán (Al-Shihi et al., 2018), así como m-enterteinment en estudiantes universitarios de Malasia (Hew et al., 2016).

3. Metodología

3.1. Participantes

El estudio fue realizado en una muestra de estudiantes de ingeniería de una institución pública de educación superior tecnológica del estado de Veracruz en México. Esta institución ha implementado un sistema de gestión de procesos académicos (por ejemplo, consulta de notas, inscripciones, seguimiento de residencias profesionales), llamado Sistema de Gestión Integral (SIG).

Se realizó un muestreo no aleatorio solicitando la respuesta del instrumento en la plataforma del SIG. Se obtuvieron 405 respuestas válidas, participando 228 hombres y 177 mujeres.

3.2. Instrumentos

Los tres constructos (PU, PEOU y BI), fueron medidos con nueve ítems, los cuales fueron adaptados del estudio original de Davis (1989) y trabajos relacionados. La definición conceptual de cada constructo se presenta en la tabla 1.

Constructo	Definición conceptual	Items
Utilidad percibida (PU)	Grado en que una persona cree que usar un sistema en particular mejoraría su desempeño laboral (Davis, 1989).	3
Facilidad de uso percibida (PEOU)	Grado en que una persona cree que usar un sistema en particular estará libre de esfuerzo (Davis, 1989).	3
Intención de comportamiento (BI)	Grado en el que un usuario formula planes conscientes para desarrollar o no un comportamiento futuro particular (Warshaw & Davis, 1985).	3

Tabla 1 – Definición de constructos

3.3. Estrategias de estimación

Para la estimación del modelo se utilizaron las tres estrategias que se han descrito anteriormente:

Modelos de ecuaciones estructurales basados en Covarianzas (CBSEM): la modelación de ecuaciones estructurales (SEM, por sus siglas en inglés), posibilita examinar simultáneamente una serie de relaciones de dependencia, lo que permite contestar a un conjunto de preguntas de investigación interrelacionadas en un simple, sistemático y compresivo análisis (Gefen, Straub, & Boudreau, 2000). Existen dos aproximaciones de SEM, CBSEM y PLS. La CBSEM es una técnica que surgió de la unión de la perspectiva econométrica enfocada en la predicción y el enfoque psicométrico que modela conceptos como las variables latentes (no observadas) que son indirectamente inferidas de múltiples medidas observadas (indicadores o variables manifiestas) (Chin, 1998), es ésta se requiere una fuerte base teórica y supone que los datos analizados presentan una distribución normal.

La literatura explica que CBSEM es esencial para conducir investigación aplicada en Ciencias Sociales, pues se desempeña apropiadamente tanto desde la filosofía del descubrimiento como desde la filosofía de la confirmación (Bagozzi & Yi, 2012). Por tanto cuando la meta de la investigación es confirmar un modelo teórico que incorpora solo variables medidas con indicadores reflectivos, se propone seleccionar CBSEM (J. Hair, Hollingsworth, Randolph, & Chong, 2017; Sarstedt, Hair, Ringle, Thiele, & Gudergan, 2016).

Mínimos cuadrados parciales (PLS): el enfoque de PLS, se ha popularizado en los últimos años entre los investigadores, dadas ciertas ventajas sobre los enfoques basados en covarianzas. El uso de PLSSEM no requiere necesariamente una fuerte base teórica, y por tanto, soporta tanto investigación exploratoria como confirmatoria. Adicionalmente, PLS requiere tamaños de muestra relativamente pequeñas en relación a CBSEM (J. F. Hair et al., 2014).

Si bien existen limitaciones en el uso de PLS (Marcoulides, Chin, & Saunders, 2009), diversos autores propone seleccionar esta aproximación cuando se desea tener una mayor capacidad de predicción (J. F. Hair et al., 2014; Henseler et al., 2009), o cuando se desea examinar la heterogeneidad no observada en un modelo (J. Hair et al., 2017).

Redes neuronales artificiales (ANN): Una red neuronal es un procesador distribuido masivamente paralelo formado por unidades de procesamiento simples, que tienen una propensión neuronal para almacenar conocimiento experimental y hacerlo disponible para su uso (Haykin, 2001). De manera similar al cerebro humano, un modelo de red neuronal es capaz de adquirir conocimiento con la ayuda del proceso de aprendizaje (Chiang, Zhang, & Zhou, 2006). Una red neuronal está compuesta por tres capas: de entrada, de salida y ocultas con fuertes conexiones conocidas como pesos sinápticos que se ajustan a través del proceso de aprendizaje iterativo (Tan et al., 2014)

Mientras que la estimación PLS está destinada principalmente para el análisis causal-predictivo en situaciones de alta complejidad pero baja información teórica, el CBSEM se centra en la confirmación de la relación asumida teóricamente (López-Bonilla & López-Bonilla, 2017).

Por su parte, los modelos de redes neuronales se han establecido como un mejor modelo predictivo para los modelos estadísticos (Henseler et al., 2009) y son capaces de capturar la relación no lineal entre las variables de decisión además de la relación lineal (Al-Shihi et al., 2018).

Para poder realizar la comparación del poder predictivo de las estrategias de estimación, se estimó el mismo modelo para los tres casos utilizando los tres constructos descritos: PU, PEUO y BI. Se buscó conservar el modelo lo más parsimónico posible por lo que se excluyó a la actitud del modelo. Para efectos de poder estimar las mismas relaciones en todos los modelos se eliminó la relación entre PEOU y PE. Esto debido a que en el modelo de ANN no es posible estimar los efectos indirectos originados por la mediación de variables.

4. Resultados

4.1. Confiabilidad

La Tabla 2 muestra los resultados de la confiabilidad del instrumento. Puede apreciarse que todas las dimensiones presentan un coeficiente de Alfa Cronbach y de Varianza Media Extraída (AVE, por sus siglas en inglés) por encima de 0.800 lo cual es indicativo su confiablidad.

Constructo	Alfa de Cronbach	AVE
PU	0.887	0.816
PEOU	0.894	0.825
BI	0.935	0.883

Tabla 2 – Confiabilidad del instrumento

4.2. CBSEM

La Figura 1 muestra los resultados obtenidos del modelo de ecuaciones estructurales basado en covarianzas.

Puede apreciarse que las relaciones entre PU y PEOU resultan significativas lo que resulta adecuado de acuerdo con los estudios empíricos previos. En el caso del poder predictivo del modelo la R^2 es de 0.418.

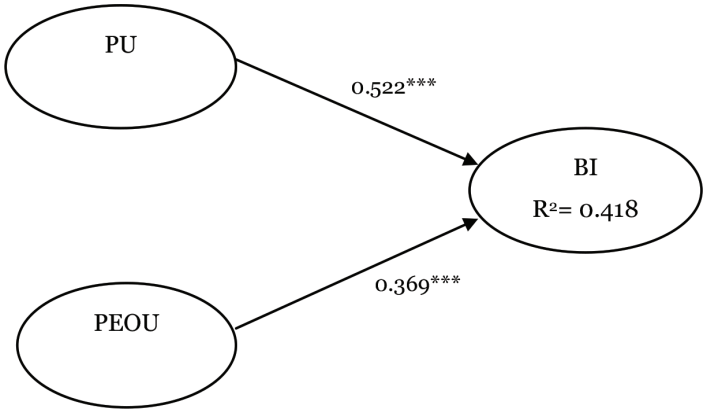


Figura 1 – Resultados CBSEM

4.2. PLS

Para el caso de los resultados de la estimación por PLS (Figura 2), los coeficientes beta de igual manera son significativos, aunque de menor magnitud. En este caso R^2 se incrementó de 0.418 a 0.477.

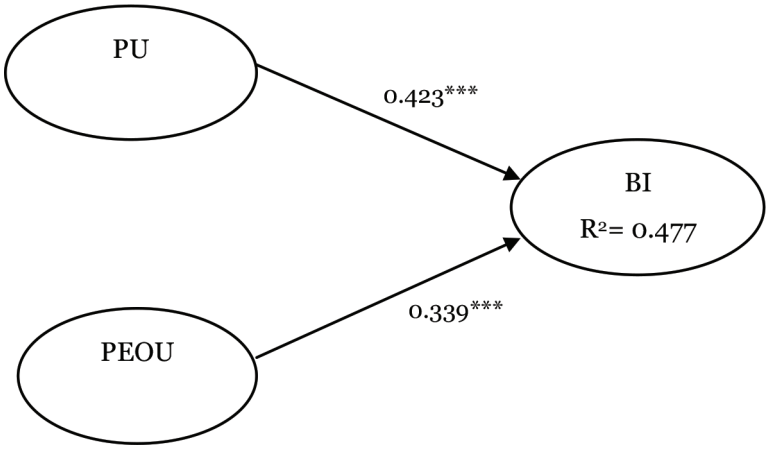


Figura 2 – Resultados PLS

4.3. ANN

Se estimó el modelo de red neuronal mediante un entrenamiento Deep Learning en el software Rapid Minner utilizando el 80% de los datos para el aprendizaje del modelo y el 20% restante para la predicción. Se han utilizado los 3 reactivos de PU y los 3 de PEOU, además del promedio de los 3 reactivos de BI como inputs del modelo (Figura 3). A partir de éstos y después de un proceso de Deep Learning, se obtuvieron las predicciones para BI. Tal predicción al compararse con los valores reales ha reportado un R^2 de 0.520.

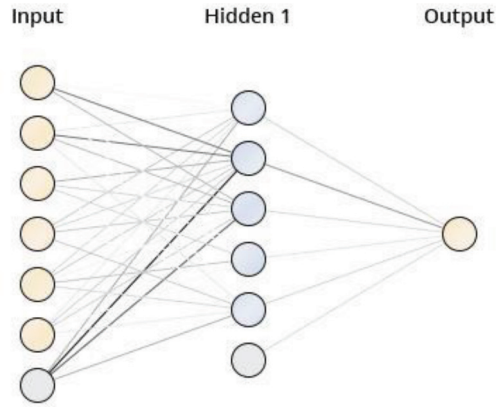


Figura 3 – Red Neuronal

La Tabla 3 muestra un resumen del poder predictivo de los tres modelos estimados.

Estimación	R^2
CBSEM	0.418
PLS	0.477
ANN	0.520

Tabla 3 – Resumen estimaciones

Puede apreciarse que tal como lo exponen Henseler, Ringle, y Sinkovics (2009), la estimación por ANN tiene un mayor poder predictivo que PLS y CBSEM.

5. Conclusiones

El presente estudio tuvo como objetivo comparar el poder predictivo de tres diferentes estrategias de estimación para el modelo TAM. A partir de los resultados podemos concluir lo siguiente: por principio, se valida la influencia tanto de PU como PEOU sobre BI para el caso de un sistema de gestión académica en el contexto de una institución de educación superior en México. Segundo, se confirma que las ANN tienen mejor poder predictivo que los PLS y éstos a su vez son mejores predictores que CBSEM. Con esto se valida la naturaleza predictiva de las ANN y confirmatoria de CBSEM. Como futuros temas de investigación se propone estudiar el poder predictivo de las tres estrategias en otros contextos y con diferentes modelos.

Referencias

Agudo-Peregrina, Á. F., Hernández-García, Á., & Pascual-Miguel, F. J. (2014). Behavioral intention, use behavior and the acceptance of electronic learning systems: Differences between higher education and lifelong learning. *Computers in Human Behavior*, *34*, 301–314. http://doi.org/10.1016/j.chb.2013.10.035

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Al-Emran, M., Mezhuyev, V., & Kamaludin, A. (2018). Technology Acceptance Model in M-learning context: A systematic review. *Computers & Education*, *125*, 389–412. http://doi.org/10.1016/J.COMPEDU.2018.06.008

Al-Maroof, R. A. S., & Al-Emran, M. (2018). Students Acceptance of Google Classroom: An Exploratory Study using PLS-SEM Approach. *International Journal of Emerging Technologies in Learning (iJET)*, *13*(06), 112. http://doi.org/10.3991/ijet.v13i06.8275

Al-Shihi, H., Sharma, S. K., & Sarrab, M. (2018). Neural network approach to predict mobile learning acceptance. *Education and Information Technologies*, *23*(5), 1805–1824. http://doi.org/10.1007/s10639-018-9691-9

Ali, M., Raza, S. A., Qazi, W., & Puah, C.-H. (2018). Assessing e-learning system in higher education institutes. *Interactive Technology and Smart Education*, *15*(1), 59–78. http://doi.org/10.1108/ITSE-02-2017-0012

Bagozzi, R. P., & Yi, Y. (2012). Specification, Evaluation and Interpretation ofStructural Equation Models. *Journal of the Academy of Marketing Science*, *40*(1), 8–34.

Bazelais, P., Doleck, T., & Lemay, D. J. (2018). Investigating the predictive power of TAM: A case study of CEGEP students' intentions to use online learning technologies. *Education and Information Technologies*, *23*(1), 93–111. http://doi.org/10.1007/s10639-017-9587-0

Chiang, W. K., Zhang, D., & Zhou, L. (2006). Predicting and explaining patronage behavior toward web and traditional stores using neural networks: a comparative analysis with logistic regression. *Decision Support Systems*, *41*(2), 514–531. http://doi.org/10.1016/J.DSS.2004.08.016

Chin, W. W. (1998). The Partial Least Squares Approach to Structural Equation Modeling. En G. A. Marcoulides (Ed.), *Modern Methods for Business Research*. Mahwah, NJ: Lawrence Erlbaum Associates.

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319. http://doi.org/10.2307/249008

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1992). Extrinsic and Intrinsic Motivation to Use Computers in the Workplace1. *Journal of Applied Social Psychology*, *22*(14), 1111–1132. http://doi.org/10.1111/j.1559-1816.1992.tb00945.x

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research*.

Gefen, D., Straub, D. W., & Boudreau, M.-C. (2000). Structural equation modelling and regression: guidelines for research practice. *Communications of the Association for Information Systems*, *4*(1), 7 (79pp).

Hair, J. F., Hult, G. T., Ringle, C. M., Sarstedt, M., Ringle, C. M., & Hair, J. F. (2014). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM). Handbook of market research*. SAGE Publications, Inc.

Hair, J., Hollingsworth, C. L., Randolph, A. B., & Chong, A. Y. L. (2017). An updated and expanded assessment of PLS-SEM in information systems research. *Industrial Management and Data Systems*, *117*(3), 442–458. http://doi.org/10.1108/IMDS-04-2016-0130

Haykin, S. (2001). *Neural networks: a comprehensive foundation*. Upper Saddle River, NJ: Prentince Hall.

Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. En *New Challenges to International Marketing* (Vol. 20, pp. 277–319). Emerald Group Publishing Limited. http://doi.org/doi:10.1108/S1474-7979(2009)0000020014

Hew, T. S., Leong, L. Y., Ooi, K. B., & Chong, A. Y. L. (2016). Predicting drivers of mobile entertainment adoption: A two-stage sem-artificial-neural-network analysis. *Journal of Computer Information Systems*, *56*(4), 352–370. http://doi.org/10.1080/08874417.2016.1164497

Joo, Y. J., Lee, H. W., & Ham, Y. (2014). Integrating user interface and personal innovativeness into the TAM for mobile learning in Cyber University. *Journal of Computing in Higher Education*, *26*(2), 143–158. http://doi.org/10.1007/s12528-014-9081-2

King, W. R., & He, J. (2006). A meta-analysis of the technology acceptance model. *Information & Management*, *43*(6), 740–755. http://doi.org/10.1016/J.IM.2006.05.003

Liébana-Cabanillas, F., Marinković, V., & Kalinić, Z. (2017). A SEM-neural network approach for predicting antecedents of m-commerce acceptance. *International Journal of Information Management*, *37*(2), 14–24. http://doi.org/10.1016/J.IJINFOMGT.2016.10.008

López-Bonilla, L. M., & López-Bonilla, J. M. (2017). Explaining the discrepancy in the mediating role of attitude in the TAM. *British Journal of Educational Technology*, *48*(4), 940–949. http://doi.org/10.1111/bjet.12465

Marcoulides, G. A., Chin, W. W., & Saunders, C. (2009). A critical look at partial leasy squares modeling. *MIS Quarterly*, *33*(1), 171–175. http://doi.org/10.2307/4132321

Mokhtar, S. A., Katan, H., & Hidayat-Ur-Rehman, I. (2018). Instructors' Behavioural Intention to Use Learning Management System: An Integrated TAM Perspective. *TEM Journal*, *7*(3), 513–525. http://doi.org/10.18421/TEM73-07

Park, S. Y. (2009). An Analysis of the Technology Acceptance Model in Understanding University Students' Behavioral Intention to Use e-Learning. *Journal of Educational Technology & Society*. International Forum of Educational Technology & Society. http://doi.org/10.2307/jeductechsoci.12.3.150

Rafique, H., Anwer, F., Shamim, A., Minaei-Bidgoli, B., Qureshi, M. A., & Shamshirband, S. (2018). Factors Affecting Acceptance of Mobile Library Applications: Structural Equation Model. *Libri*, *68*(2), 99–112. http://doi.org/10.1515/libri-2017-0041

Ramirez-Anormaliza, R., Sabaté, F., Llinàs-Audet, X., & Lordan, O. (2017). Acceptance and use of e-learning systems by undergraduate students of Ecuador: The case of a state university. *Intangible Capital*, *13*(3), 548–581.

Ramirez-Correa, P. E., Arenas-Gaitan, J., Javier Rondan-Cataluna, F., Ramírez-Correa, P. E., Arenas-Gaitán, J., & Rondán-Cataluña, F. J. (2015). Gender and Acceptance of E-Learning: A Multi-Group Analysis Based on a Structural Equation Model among College Students in Chile and Spain. *Plos One*, *10*(10). http://doi.org/10.1371/journal.pone.0140460

Sarstedt, M., Hair, J. F., Ringle, C. M., Thiele, K. O., & Gudergan, S. P. (2016). Estimation issues with PLS and CBSEM: Where the bias lies! *Journal of Business Research*, *69*(10), 3998–4010. http://doi.org/10.1016/j.jbusres.2016.06.007

Tan, G. W.-H., Ooi, K.-B., Leong, L.-Y., & Lin, B. (2014). Predicting the drivers of behavioral intention to use mobile learning: A hybrid SEM-Neural Networks approach. *Computers in Human Behavior*, *36*, 198–213. http://doi.org/10.1016/J.CHB.2014.03.052

Tarhini, A., Hone, K., & Liu, X. (2014a). Measuring the Moderating Effect of Gender and Age on E-Learning Acceptance in England: A Structural Equation Modeling Approach for An Extended Technology Acceptance Model. *Journal of Educational Computing Research*, *51*(2), 163–184. http://doi.org/10.2190/EC.51.2.b

Tarhini, A., Hone, K., & Liu, X. (2014b). The effects of individual differences on e-learning users' behaviour in developing countries: A structural equation model. *Computers in Human Behavior*, *41*, 153–163. http://doi.org/10.1016/j.chb.2014.09.020

Teo, T., & Noyes, J. (2011). An assessment of the influence of perceived enjoyment and attitude on the intention to use technology among pre-service teachers: A structural equation modeling approach. *Computers & Education*, *57*(2), 1645–1653. http://doi.org/10.1016/J.COMPEDU.2011.03.002

Venkatesh, V., & Davis, F. D. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science*, *46*(2), 186–204. http://doi.org/10.1287/mnsc.46.2.186.11926

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, *27*(3), 425. http://doi.org/10.2307/30036540

Warshaw, P. R., & Davis, F. D. (1985). Disentangling behavioral intention and behavioral expectation. *Journal of Experimental Social Psychology*, *21*(3), 213–228. http://doi.org/10.1016/0022-1031(85)90017–4

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 29/09/2018
Aceitação/Acceptance: 01/12/2018

Pharming y Phishing: Delitos Informáticos Penalizados por la Legislación Ecuatoriana

Hamilton Villón[1], Marlon Sojos[1], Carlos Mendoza[1], Teresa Guarda[1,2,3], Arturo Clery[1]

hamilton.villon@gmail.com, angelsoj32@gmail.com,c_mendoza95@hotmail.com, tguarda@gmail.com, **clery@upse.edu.ec**

[1] Universidad Estatal Península de Santa Elena – UPSE, La Libertad, Ecuador

[2] Universidad de las Fuerzas Armadas-ESPE, Unidad Académica Especial ESSUNA, Salinas, Ecuador

[3] Algoritmi Centre, Minho University, Guimarães, Portugal.

Pages: 671–677

Resumen: La tecnología en la actualidad brinda tantas ventajas y beneficios que es inevitable no emplear alguna de las herramientas que proporciona. El internet, es en la actualidad el recurso tecnológico más usado en el mundo. La navegación en la web por este lado, es la tarea más frecuente realizada por las personas. Ya sea desde un computador o hasta en un dispositivo inteligente. El internet es sumamente ventajoso y en general nos ha facilitado la vida, a través de la creación de plataformas y sitios en línea; razón por la cual nos vemos en la necesidad de dar información personal al internet para hacer uso de estas. Algo que debe hacerse responsablemente puesto que a la par que la tecnología aumenta, los delitos relacionados a estos se incrementan en igualad magnitud. Así pues, surgen un sin número de delitos informáticos, en donde los autores de estos pueden llegar a extorsionar, estafar e inclusive suplantar la identidad de sus víctimas. Problemas que han aumentado, considerablemente a lo largo de recientes años. Dos de los más comunes son el Pharming y el phishing, mecanismos que son usados por los delincuentes informáticos que usan la ingenuidad de las personas para sustraer su información personal y captarlos para un mal uso posterior. Es necesario estar al tanto de esta problemática actual, pues cada vez más están presentes en nuestra sociedad. La legislación ecuatoriana vigente actualmente a formulado medidas y sanciones para aquellas personas que lo cometan. El presente ensayo se abarcará toda esta situación.

Palabras-claves: Extorción, Confidencial, Pharming, Phishing.

Pharming and Phishing: Cybercrime penalized by the Ecuadorian Legislation

Abstract: Technology today offers so many advantages and benefits that it is inevitable not to use any of the tools it provides. The internet is currently the most used technological resource in the world. Web browsing on this side is the most frequent task performed by people. Either from a computer or even in an intelligent device. The internet is extremely advantageous and in general it has made life

easier for us, through the creation of platforms and online sites; reason why we see ourselves in the need to give personal information to the internet to make use of these. Something that must be done responsibly since at the same time as technology increases, the crimes related to these increase in equal magnitude. Thus, a number of computer crimes arise, where the perpetrators can extort money, cheat and even supplant the identity of their victims. Problems that have increased, considerably over recent years. Two of the most common are Pharming and phishing, mechanisms that are used by computer criminals who use the ingenuity of people to steal their personal information and capture them for later misuse. It is necessary to be aware of this current problem, because more and more are present in our society. The Ecuadorian legislation currently in force has formulated measures and sanctions for those who commit it. This essay will cover all this situation.

Keywords: Extortion, Confidential, Pharming, Phishing.

1. Introducción

El uso de las tecnologías hoy en día es prácticamente inevitable, el mundo en el que vivimos está sumamente inmerso en el uso de este recurso que en una perspectiva general todos tenemos acceso a él directa o indirectamente. Independientemente de la nación del mundo de la que provengamos tenemos derecho a emplearlo.

La navegación en la web ha dado la posibilidad a miles de personas en todo el mundo de comunicar sus ideas y provocar cambios en sus sociedades. Razón por la cual en recientes años, la Asamblea General de las Naciones Unidas ha declarado el acceso a internet como un derecho humano. Citamos por ello las palabras del Relator Especial de la ONU, Frank La Rue, el cual nos dice: "La única y cambiante naturaleza de internet no sólo permite a los individuos ejercer su derecho de opinión y expresión, sino que también forma parte de sus derechos humanos y promueve el progreso de la sociedad en su conjunto" (WebEspacio, s.f.).

Desde una perspectiva general el internet ha traído consigo un sin número de ventajas para aquellos que a diario lo usan, permiten la comunicación directa a otros puntos del mundo, nos mantiene al tanto de lo que sucede globalmente respecto a economía, noticias, entretenimiento; y en general nos ha facilitado nuestras vidas automatizando un sin número de procesos que antes eran un tanto tediosos de realizar. Referente a esto, en la actualidad la mayoría de los procesos que realizamos a diario están relacionados con la interacción con la web e internet. Lo que conlleva a que cada vez más nos vemos en la necesidad de proporcionar información personal a cientos de sitios y portales web en línea. Un caso muy conocido son las redes sociales, usado por millones y millones de usuario, en estas se consta de un sinnúmero de datos delicados que si caen en las manos incorrectas representan un gran problema al involucrado, pudiendo en ocasiones estos ser víctimas de extorción, estafas e inclusive suplantación de identidad.

Los delitos relacionados con la informática, no son problemas recientes, ni mucho menos escasos, al contrario, en la última década estos han ido en creciente aumento, con cifras cada vez más alarmantes. Aunque podría decirse que la raíz de aquello es el uso indiscriminado de las redes sociales, esto no es del todo cierto, pues además de estos medios sociales, hay diversos sitios y portales con información personal que debería ser confidencial pero que en ocasiones es de acceso público, tal es el caso de

los registros médicos, cuentas bancarias, información laboral del ministerio de obras públicas, información académica de repositorios de instituciones de educación superior y otras más.

2. Pharming y Phishing ¿Qué son y en qué consisten?

El incremento del uso de los servicios del internet, produce que proporcionemos información personal a las diferentes plataforma, aplicaciones, sitios y portales web. Esto no se realiza controladamente, al contrario, cada vez prestamos menos cuidado posible, por lo que es de esperarse que haya personas que impulsados por diversas motivaciones se quieran hacerse con todos nuestros datos (Competitividad, s.f.). De esta ideología de sustracción de información surgen el "Pharming y el Phishing", delitos considerados los más comunes y de los que estamos más propensos a caer desafortunadamente (Guarda, Bustos, Torres, & Villao, 2018).

Hoy en día en el Ecuador estos delitos informáticos están sucediendo cada vez más, pero contrariamente a esto, son pocos las personas detenidas a causa de estos. Puede deberse a muchas razones, pero creemos que la principal es el desconocimiento como tal. Las definiciones no son del todo claras y no se sabe en que consiste y cuales podrían ser las posibles consecuencias. Por ello es necesaria establecer de manera simple y concisa de que trata cada uno de estos delitos, sus consecuencias, y más que nada, conocer las penas a las que están sujetos los delincuentes, para así incrementar los casos denunciados y cierta medida disminuir gradualmente el cometimiento de estos delitos.

El Pharming y Phishing, dos palabras pocas conocidas pero que se están dimensionando a medida que se obtiene la información de las violaciones efectuadas al usuario final. El Pharming es un modelado de páginas Web que intentan captar los datos ingresados a través de un evento en el navegador. Estos modelos son una réplica casi exacta de la página original, el objetivo de este método es engañar al usuario final, para q este ingrese su información personal (Eduardo Lagaron, s.f.).

A partir de esta mecánica surgen unas varias maneras de usarla, por ejemplo, se ha adaptado al Pharming como un virus informático que trabaja en conjunto con un Keylogger. Este virus captura las pulsaciones del teclado en las páginas oficiales accedidas, con está captura puede darse un fraude múltiple ya que existen usuarios que utilizan una sola contraseña para todo tipo de transacción.

El Phishing por otro lado, cuyo significado en español es "pescar", y si se piensa detenidamente tiene mucho sentido, consiste en utilizar los correos electrónicos a modo de anzuelo, y que sirva como enganche para que las personas procedan a dar su información personal pensando que es una fuente verídica. Se comete haciendo uso de la ingeniería social, y se caracterizada por sacar información confidencial de forma fraudulenta. Su propósito es captar a una posible nueva víctima, creando los escenarios necesarios para hacer creíble todo el engaño (Jaramillo Barea, 2012).

Como se observa las opciones de delinquir empleando herramientas tecnológicas, son casi infinitas, y muy peligrosas para los cibernautas. Algo que resulta sorprendente es la capacidad de los delincuentes informáticos para ingeniarse nuevos métodos de engaños para obtener información para sus propios fines.

Los delitos informáticos mencionados anteriormente son muy frecuentes, pero no se conoce a ciencia cierta cuál es el tamaño de impacto producido en nuestro país, ya que por diferentes razones la víctima no denuncia el hecho y el delito queda impune. Además, los términos Pharming y Phishing aún no son muy conocidos por las personas, por lo tanto, la presente redacción nos permite conocer más sobre estos delitos informáticos y dar nociones para poder evitarlos (Mercedes, s.f.).

3. Impacto del problema y su relación con la Legislación Ecuatoriana

En el Ecuador se empezó a sancionar los delitos informáticos desde el año de 2009, un tiempo considerable hasta la actualidad, pero como es de esperarse los delitos informáticos relativamente recientes no se encuentran tipificados en el vigente Código Integral Penal.

En la legislación ecuatoriana el Pharming y el Phishing están penados de alguna manera, en función de los crímenes informáticos relacionadas en cada uno de estos y que son sujetos a sanciones mediante nuestras leyes locales. Lo que se trata de decir es que la legislación actual como tal, castiga a aquellas personas que incurren en delitos que surgen o parten de uso indebido de las herramientas informáticas. Por ejemplo, la suplantación de identidad, el robo o fuga de información, y la estafa parten de estos delitos informáticos y como tal si son sujetos a sanciones.

El Pharming y el Phishing prácticamente van de la mano al momento de delinquir, pues un mecanismo se apoya en el otro a fin de que el engaño sea evidente. El robo de identidad es solo uno de muchas contravenciones surgidas de estos estos delitos.

Una manera de evitar el robo de la identidad originado por el Phishing es ser muy atento en cuanto a las características de los correos que se reciben mediante este mecanismo y tomar las medidas necesarias para cada caso en particular. Por ejemplo, emplean nombres de compañías prestigiosas, utilizan el nombre algún conocido como emisor del correo, la dirección de respuesta es diferente de la identidad que hace el envío, utiliza dominios gratuitos (@gmail.com, @hotmail.com, @yahoo), el contenido del cuerpo del email es únicamente un enlace, o bien el mensaje contiene solo enlaces a URL cortos (Valle Matute, 2013).

Como recomendaciones ante esto podemos mencionar, evitar al máximo el común mensaje que contiene algo como “haz clic aquí”. También ayuda usar navegadores como Google Chrome pues estos ayudan a mostrar la dirección a la que van los enlaces moviendo el cursor del mouse sobre el este enlace.

El robo de identidad es en síntesis la apropiación de la identidad de otra persona, para asumirla como propia frente a otras en privado o públicamente. Los impactos de este son gravísimos y producen daños graves a la honra de las personas, puesto que van desde el chantaje, amenazas, e inclusive connotaciones sexuales en el caso de las mujeres. El Código Integral Penal vigente desde el 10 de agosto del 2014 sanciona el delito de suplantación de identidad, pero no explícitamente en medios electrónicos (Rosero Puetate, 2015).

Por otro lado, el delito contra la intimidad o privacidad informática también son una consecuencia también del Pharming y el Phishing, y están penadas en nuestro país.

La información confidencial a modo de grandes bases de datos clandestinas, han adquiridito un gran valor comercial en el ámbito delincuencial. Por lo que era necesaria crear medidas legales que lo contrarresten y provean de un respaldo legal a aquellas personas que han sido victima de aquello. La confidencialidad en cuanto a la informática, implica la protección de datos y en general la seguridad de nuestra información en el ciberespacio (Rosero Puetate, 2015).

Respecto a esto, la Constitución de la República el Ecuador indica que "La intimidad personal" es un derecho que se encuentra consagrado en su Capitulo Sexto, "Derechos de Libertad", Art. 66, numerales 19, 20, 21. Así también el Art. 9 de la misma ley, se refiere a la protección de datos otro aspecto importante ligada con la privacidad e integridad de la información personal (Rosero Puetate, 2015).

En cuanto a otras actividades ilícitas sujetas a sanción que surgen del Pharming y el Phishing tenemos las siguientes:

Artículo 186.- Estafa. - La persona que, para obtener un beneficio patrimonial para sí misma o para una tercera persona, mediante la simulación de hechos falsos o la deformación u ocultamiento de hechos verdaderos, induzca a error a otra, con el fin de que realice un acto que perjudique su patrimonio o el de una tercera, será sancionada con pena privativa de libertad de cinco a siete años (Rosero Puetate, 2015).

Artículo 230.- Interceptación ilegal de datos. - Será sancionada con pena privativa de libertad de tres a cinco años la persona que, sin orden judicial previa, en provecho propio o de un tercero, intercepte, escuche, desvíe, grabe u observe (es decir realice Pharming o Phishing), en cualquier forma un dato informático en su origen, destino o en el interior de un sistema informático, una señal o una transmisión de datos o señales con la finalidad de obtener información registrada o disponible (Serrano Espinosa, 2016).

Artículo 231.- Transferencia electrónica de activo patrimonial. - La persona que, con ánimo de lucro, altere, manipule o modifique el funcionamiento de programa o sistema informático o telemático o mensaje de datos, para procurarse la transferencia o apropiación no consentida de un activo patrimonial de otra persona en perjuicio de esta o de un tercero, será sancionada con pena privativa de libertad de tres a cinco años (Rosero Puetate, 2015).

A través de estos artículos se puede notar que en la actualidad el estado ecuatoriano está comprometido en lograr un ambiente de justicia integral, considerando al ámbito informático como un escenario causante de nuevos delitos que deben ser sancionados con medidas fuertes. Vale la pena mencionar entonces que el Ecuador ha iniciado los primeros pasos en la generación de nuevas Leyes y Decretos que contemplan esos aspectos significativos relacionados con las nuevas tecnologías y a la vez que va estableciendo penas y sanciones en el Código de Procedimiento Penal. A pesar de todo ello, la sociedad en general, aún no lo evidencia claramente y se siente la falta de medias legislativas precisas y que vallan de acuerdo con las nuevas modalidades de delincuencia. El Pharming y el Phishing son delitos relativamente nuevos y como estos surgen cada día otros más, por lo que es importante la creación de leyes que contemplen a los delitos informáticos de una manera integral a medida que estos aparecen (Ureta Arreaga & Echeverria Briones, 2009).

4. Conclusiones

El internet es una herramienta muy poderosa y como tal debe ser usada responsablemente. Las repercusiones podrán ser positivas o negativas para con nosotros, y dependerá del cómo es usado. Al trabajar con plataformas o cualquier otro medio que este en línea se debe ser muy precavido al momento de proporcionar nuestros datos personales. Porque como ya se mencionó existen personas malintencionadas que pueden hacerse de estos y darle un mal uso, perjudicándonos gravemente en ocasiones.

A partir de ese hecho, la tecnología posibilita nuevas formas de operación ilícita, que se traduce a la aparición de nuevos delitos relacionados con la informática. Los mismos que aparentemente no pueden ser cubiertos en su totalidad por la legislación actual, debido a que hoy en día no se puede responder a estos nuevos y complejos problemas sociales que se convierten en potenciales delitos que deben ser sancionados y que en la mayoría de los casos no se da.

Concluimos entonces, que en la actualidad no existe una legislación penal adecuada, lo que provoca que cada vez más delitos parecidos al Pharming y el Phishing queden en la impunidad todos los días. Por esta razón creemos necesario, la toma de medidas preventivas propias que eviten que este tipo de delitos se incrementen. Como, por ejemplo, no dar indiscriminadamente nuestros datos personales, pues no sabemos con exactitud, en donde se almacenan y quien va a ser uso de ella. Por lo que ser víctima de estos delitos, es prácticamente decisión propia y del uso responsable que le damos a nuestra información personal.

Referencias

Competitividad, M. d. (n.d.). www.planetic.es. Retrieved 08 23, 2018, from http://www.planetic.es/content/grupo-de-trabajo-micronanotecnolog%C3%ADas-y-sistemas-inteligentes

Eduardo Lagaron, K. R. (n.d.). www.eduardolagaron.com. Retrieved 08 23, 2018, from http://www.eduardolagaron.com/wp-content/uploads/2011/02/phising-pharming-y-spoofing2.pdf

Expansión. (2011, Junio 08). Expansión. Retrieved from https://expansion.mx/tecnologia/2011/06/08/la-onu-declara-el-acceso-a-internet-como-un-derecho-humano

Guarda, T., Bustos, S., Torres, W., & Villao, F. (2018). Botnets the Cat-Mouse Hunting. The 2018 International Conference on Digital Science (pp. 408-416). Montenegro: Springer, Cham.

Jaramillo Barea, F. X. (2012). Delitos informáticos prevención modificación y creación de nuevos tipos penales. Universidad del Pacífico Escuela de Negocios, Facultad de Derecho y Ciencias Políticas. Guayaquil : Universidad Del Pacífico. Retrieved from http://repositorio.upacifico.edu.ec/handle/40000/55

Mercedes, I. (n.d.). Delitos Infomáticos. Retrieved 08 23, 2018, from http://idalnemercedes.blogspot.com/2011/05/delitos-informaticos_06.html

Nacional, P. (n.d.). www.policiaecuador.gob.ec. Retrieved 08 23, 2018, from http://www.policiaecuador.gob.ec/delitos-informaticos-o-ciberdelitos/

Rosero Puetate, M. C. (2015). El derecho Ecuatoriano y la suplantación a la identidad por medios electrónicos respuesta oportuna. Tesis de Pregrado, Universidad Regional Autónoma de los Andes "UNIADES", Facultad de Jurisprudencia, Ambato. Retrieved from http://dspace.uniandes.edu.ec/handle/123456789/405

Serrano Espinosa, K. A. (2016). Los delitos ejecutados mediante la red social facebook y su relación con el Codigo Organico Integral Penal Ecuatoriano. Tesis de Pregrado, UNIVERSIDAD REGIONAL AUTÓNOMA DE LOS ANDES UNIANDES , Facultad de Jurisprudencia, Babohoyo. Retrieved from http://dspace.uniandes.edu.ec/handle/123456789/4944

Ureta Arreaga, L. A., & Echeverria Briones, P. F. (2009). RETOS A SUPERAR LA LA ADMINISTRACIÓN DE JUSTICIA ANTE LOS DELITOS INFORMÁTICOS EN EL ECUADOR. Tesis de Grado, Escuela Superior Politécnica del Litoral, Facultad de Ingeniería en Electricidad y Computación, Guayaquil. Retrieved from http://www.dspace.espol.edu.ec/handle/123456789/5792

Valle Matute, J. C. (2013). El delito informático de Phishing. Tesis de Pregrado, Universidad Regional Autónoma de los Andes UNIANDES, Facultad de Jurisprudencia, Quevedo. Retrieved from http://dspace.uniandes.edu.ec/handle/123456789/2819

WebEspacio. (n.d.). www.webespacio.com. Retrieved 08 23, 2018, from https://www.webespacio.com/internet-derecho-humano/

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 19/09/2018
Aceitação/Acceptance: 11/12/2018

Herramienta para el seguimiento del estrés mental en pruebas de usabilidad

Gabriel Elías Chanchí G.[1], Martha Giraldo G.[2], Wilmar Yesid Campo Muñoz[3]

gchanchi@unimayor.edu.co, marthagir16@unimayor.edu.co, wycampo@uniquindio.edu.co.

[1] Institución Universitaria Colegio Mayor del Cauca, Popayán, 190001, Cauca, Colombia.

[2] Institución Universitaria Colegio Mayor del Cauca, Popayán, 190001, Cauca, Colombia.

[3] Universidad del Quindío, Armenia, 630004, Quindío, Colombia.

Pages: 678–690

Resumen: En un test de usabilidad, la percepción de usuario es obtenido por lo general a partir de la percepción del coordinador sobre el comportamiento del usuario, mientras este interactúa con el software evaluado. Este artículo propone una herramienta automatizada para el seguimiento del estrés mental de un usuario durante una prueba de usabilidad, tomando como variable fisiológica la variación de la frecuencia cardiaca (VFC), como contribución a la problemática de la estimación de la percepción de usuario. Para la construcción de la herramienta de estimación de la percepción de usuario, esta investigación se abordó en 5 fases a saber: caracterización de pruebas de usuario, estudio del estrés mental, diseño de la herramienta, implementación de la herramienta y verificación del funcionamiento de la herramienta. La herramienta construida pretende servir de apoyo en la ejecución de pruebas de usuario, en cuanto a la obtención objetiva de la percepción de usuario a partir del análisis del estrés mental.

Palabras-clave: Estrés mental; percepción de usuario; pruebas de usuario; usabilidad; VFC.

Tool for monitoring mental stress in usability test

Abstract: In a usability test, the user perception attribute is usually obtained from the coordinator's perception of user behavior, while interacting with the evaluated software. In this paper we propose an automated tool for monitoring the mental stress of a user during a usability test, taking as physiological variable the heart rate variability (HRV), as contribution to the problem of estimating the user perception. For the construction of the user perception estimation tool, this research was addressed in 5 phases namely: characterization of user testing, mental stress study, tool design, tool implementation and tool verification. The proposed tool, aims to support the execution of user tests, in terms of the objective achievement of the user perception from the analysis of mental stress.

Keywords*:* HRV, mental stress; usability; user perception; user testing.

1. Introducción

Gracias a los avances tecnológicos entorno a temáticas como Internet de la cosas (IoT), se estima que cerca de 20,8 billones de dispositivos alrededor del mundo estarán conectados a internet en 2020 (Said & Masud, 2013), (Virkki & Chen, 2013). Así mismo, según (Evans, 2011), se estima que para el mismo año se tendrán aproximadamente 6,58 dispositivos conectados a internet por persona, por lo cual la cifra anterior podría incrementar a 50 billones de dispositivos. Las anteriores cifras y tendencias en cuanto a IoT han permitido que también el número de aplicaciones software desplegadas sobre internet y en entornos móviles crezcan, ya que estas dan soporte a nivel de usuario a las diferentes tecnologías. Por ende, es relevante considerar al usuario final como parte fundamental del ciclo del éxito de dichas aplicaciones. A partir de lo anterior, las organizaciones han incluido en sus proyectos requisitos de usabilidad, pues han identificado la importancia que representa desarrollar productos usables que permitan traer una mayor cantidad de usuarios a sus tecnologías o aplicaciones.

De acuerdo al estándar ISO 9241-11, la usabilidad es entendida como: "El grado en que un producto puede ser usado por usuarios específicos, para lograr los objetivos especificados, con eficacia, eficiencia y satisfacción o percepción de usuario, en un contexto de uso específico", es decir que la usabilidad puede ser evaluada mediante el análisis de métricas asociadas a tres atributos fundamentales como son: eficacia, eficiencia y satisfacción de usuario (Perurena & Moráguez, 2013). La eficiencia puede ser determinada en un laboratorio de usabilidad, mediante métricas tales como el número de intercambios por unidad de tiempo que el usuario puede realizar en una prueba usando el sistema (Estayno, Dapozo, Cuenca Pletsch, & Greiner, 2009). De este modo, la eficiencia puede estar asociada con el tiempo que tardan los usuarios en realizar una tarea, tras aprender el funcionamiento básico de una aplicación. En lo que respecta a la eficacia en una prueba de usuario, esta puede obtenerse con indicadores tales como el conjunto de porcentajes relacionados con tareas terminadas, así como con el éxito de dichas tareas terminadas y los tiempos de ejecución de dichas tareas (Mascheroni, Greiner, Petris, Dapozo, & Estayno, 2012). Finalmente, la satisfacción puede ser definida según la ISO 9241-11 como "ausencia de incomodidad y existencia de actitudes positivas hacia la utilización del producto" (Ronda, 2005). Así, de acuerdo a lo anterior, existe una cierta relación entre el comportamiento emocional de un usuario durante una prueba de usabilidad y su grado de percepción (Ferré, 2000).

Estimar la percepción de usuario ha representado un gran reto no solo para la academia, también para la industria, puesto que el estudio del comportamiento humano es complejo. Normalmente para evaluar la percepción de usuario en una prueba de usabilidad, se hace uso de cuestionarios orientados a conocer el nivel de percepción de los usuarios con respecto a su experiencia con la aplicación. Dichos cuestionarios aunque han permitido aportar datos para dar cumplimiento de las métricas de usabilidad, tienen como desventaja el depender de la veracidad del usuario en las diferentes respuestas, lo cual puede afectar el grado de validez de la prueba. Otro de los modos de determinar la percepción de usuario es mediante la percepción por parte del evaluador de la prueba a través del análisis de los gestos, las expresiones faciales y las posturas del usuario,

mientras este interactúa con una aplicación (Montero & Fernández, 2003). Así, es posible concluir que de los tres atributos asociados a la usabilidad, el que cuenta con métricas más subjetivas es la percepción de usuario.

Teniendo en cuenta la problemática anterior y considerando el creciente surgimiento de dispositivos IoT para el seguimiento de variables fisiológicas (ritmo cardiaco, conductividad de la piel, ondas cerebrales, entre otros), se abre la oportunidad de aprovechar dichas variables con el fin de obtener indicadores emocionales que pueden ser relacionados con la percepción de usuario. Lo anterior considerando que existen diferentes estudios en los que se hace uso de variables fisiológicas con el fin de analizar el comportamiento emocional de un usuario (Solarte, Sánchez, Chanchí, & Durán, 2016).

En este artículo se propone una herramienta para el análisis y seguimiento del estrés mental de un usuario en una prueba de usuario. Esta herramienta busca obtener indicadores que posibiliten la caracterización de patrones de comportamiento emocional, los cuales puedan usarse para la estimación de la percepción de usuario en una prueba de usabilidad. Para lo anterior, se considera como variable fisiológica la variación de la frecuencia cardiaca (VFC), la cual una vez muestreada es usada para el cálculo del estrés mental a través de la fórmula de Bayevsky (Bayevsky et al., 2002). La herramienta propuesta pretende servir de guía para el evaluador de una prueba de usabilidad, en cuanto a la estimación de la percepción de usuario, de modo que permite hacer un seguimiento en el tiempo de las diferentes tareas que realiza un usuario dentro del laboratorio de usabilidad.

El resto del artículo está organizado de la siguiente forma: en la sección 2 se definen algunos conceptos que se tuvieron en cuenta para el desarrollo de este artículo. En la sección 3 se describe la estructura y funcionamiento de la herramienta propuesta. Finalmente en la sección 4 se presentan las conclusiones y trabajos futuros derivados de este trabajo.

2. Trabajos Relacionados

La usabilidad es un atributo complejo de estimar, lo que puede generar productos software con un nivel de usabilidad deficiente, cuando una mayor atención hacia este atributo contribuiría a incrementar la calidad del producto percibida por el usuario (Mascheroni et al., 2012). Así, diferentes investigaciones se centran en realizar una cuantificación de la usabilidad tal que les permita obtener mejores resultados en el desarrollo de software. Por ejemplo, en (Georgsson & Staggers, 2015) los autores identifican como los sistemas de salud móvil (mHealth) aún necesitan estudios de usabilidad sobre las perspectivas de los pacientes y el desempeño de interacción de mHealth, por lo que proponen como caso de estudio la cuantificación de la usabilidad para un sistema mHealth de diabetes sobre las métricas de eficacia, eficiencia y satisfacción o percepción de usuario. La identificación del rendimiento de la tarea más difícil indicó las áreas para el rediseño. Usando el estándar de usabilidad ISO 9241-11, el instrumento proporcionó medidas objetivas de la usabilidad experimentada por los pacientes. Por su parte en (Harvey, Koubek, Bégat, & Jacob, 2016) se evalúa la usabilidad de un sistema de control de glucosa en sangre, encontrándose una usabilidad superior del nuevo sistema respecto de los medidores actuales presentes en los Estados Unidos y Francia. De igual modo,

en (Zaki, Wook, & Ahmad, 2017) se presenta una prueba de usabilidad para un juego terapéutico conocido como ASAH-i, que fue diseñado específicamente para estimular las habilidades cognitivas para niños con deficiencia en el habla y el manejo del lenguaje. Los datos recopilados se analizaron en función del tiempo de la tarea, el éxito de la tarea, las tasas de error y la percepción de usuario para medir el nivel de capacidad de aprendizaje, la eficiencia, los errores y la percepción de usuario con respecto al uso de ASAH-i. Los resultados muestran que la capacidad de aprendizaje de los participantes es alta cuando se tiene en cuenta la percepción de usuario de los participantes para que sea utilizable. En (Adhy et al., 2017) por su parte, se presenta una aplicación basada en IoT que permite la recuperación de datos en tiempo real para el monitoreo del clima. El objetivo de esta aplicación es recoger datos y presentar la información meteorológica según las necesidades del usuario. Por lo tanto, es necesario probar la usabilidad de la aplicación para obtener el valor de la eficiencia, la eficacia, la satisfacción y la capacidad de aprendizaje de la aplicación. Los resultados muestran que la aplicación tiene inconvenientes en términos de eficiencia y efectividad con puntajes de 73.4% y 77.8%, respectivamente. En contraste, da ventajas en satisfacción con puntaje de aproximadamente 81.1% y en capacidad de aprendizaje con puntaje de aproximadamente 74.5%. De igual manera, en (Retnani, Prasetyo, Prayogi, Nizar, & Abdul, 2017) se analiza la estimación de la satisfacción de diferentes usuarios al acceder al sitio web de una biblioteca. Para el proceso de medir la percepción de usuario se utilizó el método de Guttmann para compilar las preguntas del cuestionario. Los encuestados para cada nivel se dividen en 3 clases, usuarios no profesionales, usuarios activos y usuarios expertos. Por su parte, en (Mozos et al., 2017) se detecta de manera automática el estrés en una situación social, a partir de la combinación se sensores fisiológicos y sociométricos. La utilización de dichos sensores les permitió la identificación del estrés en tiempo real, así como saber el tipo de sensor a utilizar en ciertas actividades. Los sensores sociométricos permiten ser utilizados en actividades que requieran largos períodos de tiempo. Además, los sensores sociométricos son más fáciles de encontrar en otros dispositivos, entre los que cabe destacar: teléfonos, brazaletes o relojes inteligentes. Finalmente, en (Mcduff, Hernandez, Gontarek, & Picard, 2016) proponen la medición del estrés cognitivo con una cámara libre de contacto, con posibles aplicaciones en las áreas de aprendizaje remoto, monitoreo del estrés y la optimización de las cargas de trabajo al facilitar la medición de la experiencia del usuario. Aunque los autores reconocen la importancia de las medidas del ritmo cardiaco y la frecuencia de la respiración para reconocer el estrés cognitivo; se propone un modelo de reconocimiento basado en señales fotoplestimográficas, medidas desde una distancia de 3 metros. Una de las principales ventajas señaladas por los autores, es que según estudios indican que el contacto directo con elementos sensores pueden alterar las mediciones, dependiendo de la formación y elementos psicológicos particulares de los individuos. Mientras que las mediciones que se realizan por métodos "ópticos" pueden garantizar medidas más precisas de los elementos característicos del estrés como el ritmo cardiaco y respiratorio.

Los anteriores trabajos evidencian la necesidad de estimar de forma automatizada la percepción de usuario en pruebas de usabilidad, dado que por lo general se hace uso de cuestionarios de percepción y/o anotaciones hechas por el coordinador de la prueba al observar el comportamiento del usuario, elementos que son subjetivos al compararlos con los atributos eficacia y eficiencia. De este modo se hace importante aprovechar

el surgimiento de dispositivos de seguimiento IoT, con el fin de obtener indicadores que permitan la estimación de la percepción de usuario de manera más precisa. En el presente artículo se pretende hacer uso de dispositivos wereables IoT para la captura de la variabilidad del ritmo cardiaco durante una prueba de usuario, con el fin de calcular el estrés mental durante esta de forma automatizada, de tal modo que se pueda obtener indicadores objetivos de la percepción de usuario.

3. Metodología

Para el desarrollo del presente trabajo se plantearon cinco fases a saber: caracterización de pruebas de usuario, estudio del estrés mental, diseño de la herramienta, implementación de la herramienta y evaluación de la herramienta, ver Figura 1. En la fase de caracterización de las pruebas de usuario, se identificaron en conjunto con expertos en el área de usabilidad, las características principales de las pruebas de usuario y la estructura funcional de un laboratorio de usabilidad. En la fase de estudio del estrés mental se identificó el concepto de estrés mental, así como la fórmula para el cálculo del índice de estrés mental y las tecnologías IoT adecuadas para obtener datos asociados al estrés mental. En la fase de diseño se definió el esquema de bloques funcionales de la herramienta, teniendo en cuenta la caracterización realizada en la fase 1.

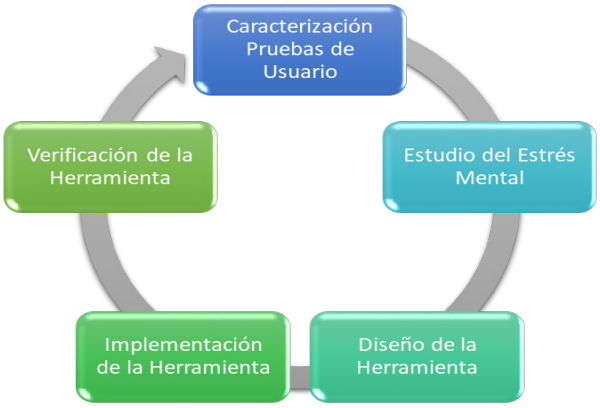


Figura 1 – Metodología propuesta

En la fase de implementación se desarrollaron los diferentes módulos definidos en la fase de diseño, haciendo uso del lenguaje Python. Finalmente, en la fase de verificación se realizó una prueba funcional de la herramienta en conjunto con un grupo de expertos en el área de la usabilidad, quienes simularon una prueba de usabilidad e inspeccionaron las diferentes prestaciones de la herramienta. Las fases de caracterización de las pruebas de usuario y de estudio del estrés mental son presentadas en la sección IV (marco conceptual), las fases de diseño e implementación de la herramienta son abordadas en la sección V (propuesta de la herramienta), finalmente la fase de verificación de la herramienta es abordada en la sección VI (verificación de la herramienta).

4. Marco Conceptual

En esta sección se presentan los conceptos relevantes que fueron considerados para el desarrollo de la presente propuesta. Dentro de estos se encuentran: laboratorio de usabilidad, variabilidad del ritmo cardiaco y nivel de estrés mental.

4.1. Laboratorio de usabilidad

Un laboratorio de usabilidad se puede considerar como una plataforma de observación, donde el objetivo principal es estudiar y analizar diferentes aspectos de la interacción entre una aplicación y un usuario. Un laboratorio de usabilidad debe posibilitar la obtención de indicadores de eficiencia, eficacia y satisfacción o percepción de usuario de acuerdo con lo propuesto por la ISO 9241-11 (Borges, 2002).

Dentro de un laboratorio de usabilidad se ejecutan las pruebas de usuario, las cuales son supervisadas por el coordinador de la prueba y están centradas en tareas. A partir de las tareas de una prueba de usabilidad, es posible obtener los atributos de eficiencia, eficacia y satisfacción. La eficacia puede determinarse mediante el porcentaje de tareas cumplidas, la eficiencia puede hallarse mediante el tiempo empleado por el usuario en ejecutar las diferentes tareas, con respecto al tiempo estimado. Finalmente la satisfacción por lo general es obtenida a través de la observación del comportamiento del usuario durante la interacción (gestos, posturas y expresiones faciales). En este artículo se propuso una herramienta automatizada que permite hacer un seguimiento sobre el estrés mental de un usuario en una prueba de usabilidad, de tal forma que se puedan obtener indicadores objetivos de la percepción de usuario.

4.2. Variabilidad del ritmo cardiaco

La variación de la frecuencia cardíaca (VFC) es definida como el cambio en frecuencia del latido cardíaco durante un intervalo de tiempo [10]. Una manera habitual de medir esta variabilidad es por medio del electrocardiograma donde son detectadas cada una de las ondas R o intervalos R-R, tal como se aprecia en la Figura 2.

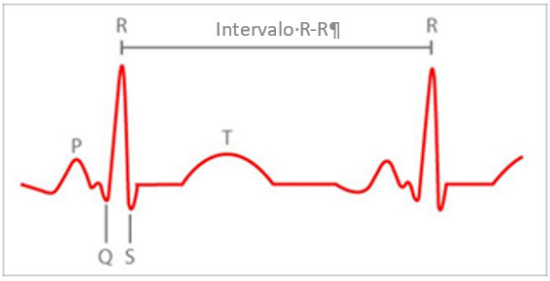


Figura 2 – Variabilidad de la frecuencia cardiaca (Ronda, 2005)

Al momento de detectar cada onda R, es calculado el tiempo entre las distintas ondas consecutivas o intervalos R-R, ver Figura 2. Este intervalo R-R es el encargado de medir el periodo cardíaco, de tal modo que al conjunto de intervalos R-R es lo que se conoce como VFC.

4.3. Índice de estrés mental

Al igual que ocurre con la emoción, no hay una definición unificada para el estrés mental, de este modo, en (Bayevsky et al., 2002) se menciona que el estrés puede ser interpretado como una amenaza a la integridad psicológica de un individuo que da lugar a respuestas fisiológicas y/o de comportamiento. Así mismo, el índice de estrés o índice del sistema de regulación de tensión, puede entenderse como el nivel de tensión de los sistemas de regulación (nivel de prevalencia de la actividad de regulación central por encima de los mecanismos autonómicos) (Solarte et al., 2016).

Existen métodos matemáticos que permiten obtener el estrés mental a través del análisis de la variabilidad de la frecuencia cardiaca o aumento del flujo sanguíneo (Borges, 2002), a este último se conoce como "arousal" que se refiere también a la excitación de las neuronas en el sistema nervioso central (Hall & Guyton, 2016). En el presente artículo se propone hacer uso de la ecuación de Bayevsky, ver Ecuación 1, para el cálculo del índice de estrés mental (SI).

$$SI = \frac{AM_o}{2M_o * (M * DM_n)}$$

Ecuación 1 – Cálculo del índice de estrés mental (Ronda, 2005)

Esta fórmula se basa en métodos geométricos, haciendo uso de la curva de distribución o cardio-histograma dibujada a partir del estudio de las variaciones del ritmo cardiaco (Bayevsky et al., 2002). En la ecuación de Bayevsky se involucran términos tales como: M_o (moda), AM_o (Amplitud de la moda), $M*DM_n$ (rango de variación o varianza) (Bayevsky et al., 2002), (Borges, 2002), (Heilman, 1997).

La Moda (M_oM_o) o nivel presumible en el que se encuentra trabajando el sistema cardiovascular, hace referencia al valor de R-R que se presenta con mayor regularidad en el conjunto de medidas analizado. La amplitud de la moda (AM_oAM_o) en términos fisiológicos, es el índice nominal de actividad de la cadena de regulación simpática, mientras que en términos matemáticos se refiere al porcentaje o porción de los intervalos que corresponden con el valor de la moda en la muestra tomada. Finalmente, el rango de varianza ($M * DM_nM * DM_n$), es la diferencia entre los valores máximos y mínimos de los cardio-intervalos R-R, es decir la varianza. En la Tabla 1, se pueden observar los rangos de índice de estrés que el usuario puede presentar. Estos a su vez están asociados a tres estados de ánimo, como son: estresado, relajado y normal. A partir de lo anterior, se puede concluir que el índice de estrés mental puede ser asociado a tres estados emocionales: estresado, normal y relajado. Estos tres estados pueden aportar indicadores importantes de cara al seguimiento de un usuario en una prueba de usabilidad. La herramienta propuesta en este artículo permite visualizar gráficamente el comportamiento emocional de un usuario a medida que este interactúa con un aplicativo software, de modo que muestra los diferentes cambios en el tiempo por los tres estados mencionados.

Rango de SI	Estado asociado	Nivel de arousal
>150	Estresado	>0,5
≥40 y ≤150	Normal	≥0,14 y ≤0,5
<40	Relajado	<0,14

Tabla 1 – Rangos SI asociados a cada estado

5. Propuesta de la Herramienta

En cuanto al diseño de la herramienta se contó con el apoyo de un conjunto de expertos en el área de usabilidad, quienes manifestaron como elemento fundamental la posibilidad de obtener un análisis general de la prueba y por tarea del estrés mental del usuario. Los anteriores elementos fueron considerados en el esquema funcional de la herramienta propuesta (diagrama de bloques y diagrama de flujo). Así mismo en esta sección se presentan las interfaces finales del prototipo desarrollado.

5.1. Esquema funcional de la herramienta

En la Figura 3 se presenta el esquema funcional de la herramienta para el seguimiento del estrés mental en una prueba de usabilidad. La herramienta está pensada para desplegarse en un laboratorio de usabilidad, por lo que en el módulo de usuario se cuenta con un cinturón adherido al pecho del usuario durante la prueba. En el módulo del evaluador por su parte, la herramienta propuesta se comunica con el cinturón a través de bluetooth y consta de los siguientes módulos funcionales: captura RR, almacenamiento, descripción de tareas, obtención del SI, visualización de resultados.

En el módulo de captura RR, se obtienen a partir del cinturón adherido al pecho del usuario, los intervalos R-R de la VFC. En el módulo de almacenamiento, la herramienta haciendo uso del gestor de bases de datos TinyDB guarda los registros (tiempo, intervalo R-R) de las capturas realizadas en el módulo anterior. Una vez termina la prueba, en el módulo de descripción de tareas el evaluador o coordinador de la prueba se encarga de consignar los tiempos de cada tarea. En el módulo de obtención del SI, la herramienta se encarga de estimar el nivel de estrés en los rangos de tiempo definidos para cada tarea en el módulo anterior. Finalmente, en el módulo generación de gráficas se despliega gráficamente el comportamiento del estrés mental en cada una de las tareas de la prueba, tal como muestra la Figura 3.

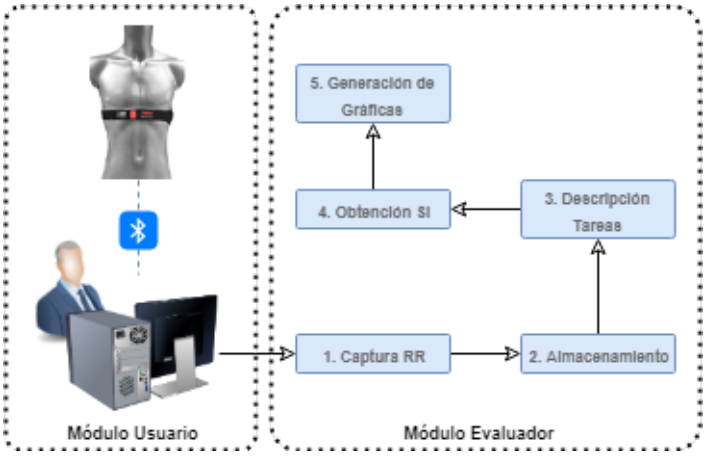


Figura 3 – Esquema funcional de la herramienta

5.2. Prototipo final

A continuación, se presentan las diferentes vistas de la interfaz gráfica correspondiente a la herramienta de seguimiento del estrés mental en pruebas de usabilidad. La herramienta está organizada en tres etapas funcionales a saber: captura de datos fisiológicos de usuario, descripción de tareas de usabilidad y representación gráfica de los resultados. En la etapa de captura de datos se obtiene una muestra de los intervalos R-R de la VFC correspondiente al tiempo total de la prueba de usabilidad. En la etapa de descripción de tareas, se especifica el número de tareas de usabilidad realizadas por el usuario durante la prueba y la duración de cada una de ellas. Finalmente en la etapa de representación gráfica de los resultados, la herramienta se encarga de representar gráficamente el comportamiento emocional del usuario (estresado, relajado, normal) para cada tarea específica de la prueba. A continuación, se describen en detalle cada una de estas etapas funcionales. En la Figura 4.a, se puede observar la interfaz principal de la herramienta de seguimiento propuesta, en la cual se destacan las pestañas de seguimiento y de tareas. En la pestaña de seguimiento se realiza la etapa de captura de datos fisiológicos durante la prueba de usabilidad, mientras que en la pestaña de tareas se ejecutan las etapas de descripción de tareas y representación gráfica de los resultados, una vez ha terminado la prueba de usabilidad.

Al iniciar el test, el coordinador de la prueba procede a ejecutar el paso 1 de la Figura 4.a, por lo que acciona el botón “Iniciar”. La herramienta empieza a obtener vía bluetooth y desde el cinturón Zephyr, variables fisiológicas del usuario tales como: variabilidad de la frecuencia cardiaca (VFC) o intervalos R-R y frecuencia cardiaca (FC) o HR. Los anteriores datos son capturados aproximadamente cada segundo desde que el usuario inicia la prueba de usabilidad hasta que termina la interacción.

Una vez termina la prueba de usabilidad, el coordinador de la prueba procede a ejecutar el paso 2 de la Figura 4, por lo que da clic en el botón “Parar” de la pestaña de seguimiento. La herramienta se encarga de guardar la información de los intervalos R-R capturados en el gestor de bases de datos TinyDB y calcular cada 10 segundos a partir de la información guardada, el índice de estrés (SI) y las métricas de pNN50 y RMSD, las cuales están directamente relacionadas con el estrés mental (Bayevsky et al., 2002). Los anteriores cálculos dan cuenta del comportamiento del estrés mental a lo largo de la prueba de usabilidad, sin discriminar las diferentes tareas realizadas a lo largo del test.

Con el fin de describir las tareas realizadas durante la prueba de usabilidad, el coordinador o evaluador de la prueba procede a ejecutar el paso 3 de la Figura 4.b, para lo cual diligencia la caja de texto de la pestaña tareas. En dicha caja de texto, la herramienta presenta por defecto el tiempo total que duró la prueba en minutos y segundos. El coordinador de la prueba debe entonces llenar separando con punto y coma el tiempo en minutos y segundos que demora cada tarea, de tal modo que la herramienta obtiene el número de tareas de la prueba, de acuerdo con el número de tokens separados por punto y coma.

A modo de ejemplo en la Figura 4.b, la caja de texto de la pestaña tareas contiene el tiempo total de la prueba obtenido por la herramienta, el cual corresponde a 5 minutos y 2 segundos. Por su parte, el coordinador de la prueba se encarga de llenar el tiempo

de las tareas separadas por punto y coma, en este caso hay 2 tareas, la tarea 1 tiene una duración de 2 minutos y la tarea 2 tiene una duración de 3 minutos con 2 segundos.

Una vez descritas las tareas de la prueba de usabilidad en el paso anterior, el coordinador procede a ejecutar el paso 4 de la Figura 4.b, para lo cual acciona el botón "Reporte Resultados". La herramienta se encarga de obtener las gráficas de seguimiento del estrés mental en el tiempo por cada una de las tareas descritas y a lo largo de la prueba de usabilidad, ver Figura 4.b. A modo de ejemplo en la Figura 5.a, se muestra un gráfico de barras que describe los diferentes cambios entre los estados: relajado, normal y estresado, a lo largo de las diferentes tareas de una prueba de usabilidad. En este caso específico, durante la tarea 2 de la prueba, el usuario se mantiene en un estado normal hasta el segundo 90, momento en el cual hay un cambio al estado relajado durante 10 segundos, luego de los cuales regresa al estado relajado hasta el final de la tarea.

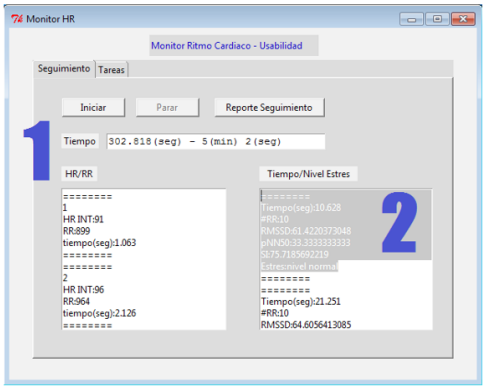


Figura 4.a – Interfaz de captura de datos

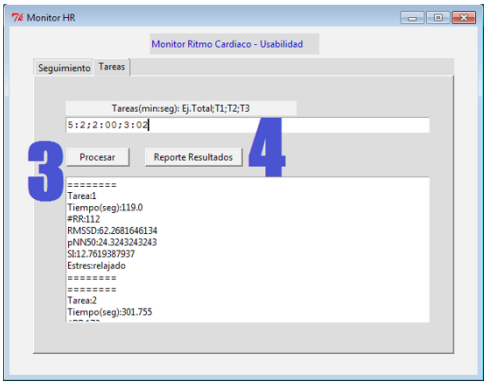


Figura 4.b – Interfaz para descripción de tareas

Del mismo modo en la Figura 5.b se presenta un gráfico de escalón generado por la herramienta propuesta. En el gráfico de la Figura 5.b generado por la herramienta, se puede apreciar de manera más detallada como fluctúa el comportamiento emocional (estresado, relajado, normal) de un usuario a lo largo de una prueba de usabilidad. Este gráfico permite ilustrar de modo claro los cambios bruscos de emoción, de tal modo que el coordinador de la prueba los pueda contrastar con la interacción que va realizando el usuario sobre el software evaluado.

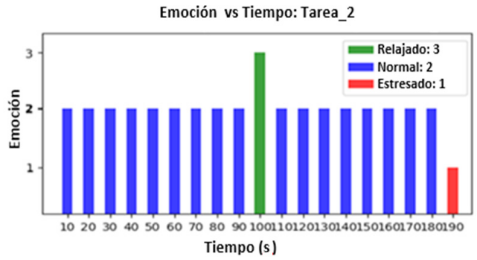


Figura 5.a – Monitoreo gráfico del estrés mental.

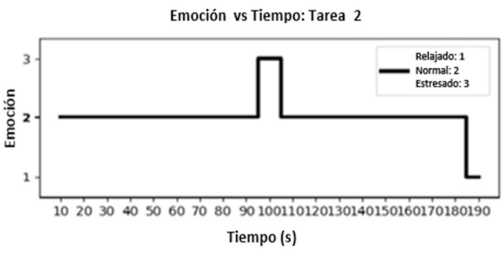


Figura 5.b – Estrés mental escalonado.

6. Evaluación de la Herramienta

A modo de verificación de la herramienta propuesta en este artículo, un grupo de tres expertos en el área de usabilidad inspeccionaron dentro del laboratorio de usabilidad de la Institución Universitaria Colegio Mayor del Cauca, la funcionalidad de la herramienta, en tal sentido que verificaron que la herramienta fuera adecuada para el seguimiento del estrés mental de acuerdo a las características de las pruebas de usabilidad. De acuerdo a lo anterior, los evaluadores manifestaron como elemento positivo el hecho que la herramienta permite hacer un seguimiento gráfico sobre el estrés mental a lo largo de la prueba, discriminando las diferentes tareas de la misma. En este sentido los evaluadores resaltaron la utilidad en cuanto a poder contrastar las gráficas obtenidas con la herramienta con los diferentes elementos de la interacción capturados en los videos de la prueba, lo cual puede arrojar posibles puntos en el tiempo donde el software evaluado no responde a las expectativas del usuario. De igual manera los evaluadores, resaltaron el hecho que la herramienta permite calcular cada 10 segundos el estrés mental, de tal modo que se puede hacer un seguimiento en tiempo real sobre el estrés mental. Finalmente, los evaluadores de la herramienta mencionaron como aspecto a mejorar el hecho de incluir un indicador numérico general o porcentual sobre el estrés mental en las diferentes tareas de la prueba. Es decir, es necesario incluir un gráfico de torta que permita obtener una estimación porcentual de los tres estados de ánimo considerados (relajado, neutral y estresado), asociados a cada tarea de la prueba.

7. Conclusiones y Trabajos Futuros

En este artículo se propuso una herramienta automatizada para el seguimiento en tiempo real del estrés mental en una prueba de usabilidad. Esta herramienta pretende servir de apoyo al coordinador de la prueba en cuanto a la estimación objetiva de la percepción de usuario en cada una de las tareas de un test de usuario.

Dentro de las principales ventajas de la herramienta, los evaluadores de la misma destacaron el seguimiento gráfico a lo largo de la prueba de usuario, sobre el comportamiento emocional del usuario por tres estados posibles: relajado, normal y estresado. Este aspecto posibilita al coordinador de la prueba contrastar las interacciones capturadas en los archivos de video con el nivel de estrés del usuario, lo cual permite detectar posibles elementos donde el software que se evalúa está fallando.

Las emociones son un elemento complejo a evaluar en un usuario, por lo que es conveniente hacer seguimiento sobre diferentes variables fisiológicas en una prueba de usabilidad. En este sentido el uso de este tipo de variables permite mejorar la objetividad de los métodos tradicionales usados en las pruebas, como son los cuestionarios de percepción y la percepción del evaluador al observar el comportamiento del usuario.

La herramienta automatizada propuesta aprovecha la tendencia en cuanto al uso de dispositivos wereables de IoT para la captura de variables fisiológicas de usuario. De este modo, en este artículo se hizo uso de la variable VFC en conjunto con la fórmula de Bayevsky para la estimación del estrés mental en cada una de las tareas de una prueba de usabilidad, representando el comportamiento del usuario haciendo uso de gráficos de barras y en escalón.

Como trabajo futuro derivado de la presente investigación, se pretende vincular dentro de la herramienta la funcionalidad de estimar porcentualmente el comportamiento emocional del usuario por cada tarea. Así mismo, dada la necesidad de integrar la herramienta propuesta en un laboratorio de usabilidad de manera adecuada, se pretende adaptarla y desplegarla sobre dispositivos tipo SBC, tal que permita conocer la experiencia del uso de la herramienta.

Agradecimientos

Al proyecto 839 de la vicerrectoría de investigaciones de Uniquindio.

Referencias

Adhy, S., Noranita, B., Kusumaningrum, R., Wirawan, P. W., Prasetya, D. D., & Zaki, F. (2017). Usability testing of weather monitoring on a web application. In *Informatics and Computational Sciences (ICICoS), 2017 1st International Conference on* (pp. 131–136).

Bayevsky, R. M., Ivanov, G. G., Chireykin, L., Gavrilushkin, A., Dovgalevsky, P., Kukushkin, U., & Fleishmann, A. (2002). HRV analysis under the usage of different electrocardiography systems (Methodical recommendations). *Committee of New Medical Techniques of Ministry of Health of Russia*, *4*, 2–67.

Borges, H. (2002). *Análisis experimental de los criterios de evaluación de usabilidad de aplicaciones multimedia en entornos de educación y formación a distancia*. Universitat Politècnica de Catalunya.

Estayno, M. G., Dapozo, G. N., Cuenca Pletsch, L. R., & Greiner, C. L. (2009). Modelos y Métricas para evaluar Calidad de Software. In *XI Workshop de Investigadores en Ciencias de la Computación*.

Evans, D. (2011). The internet of things: How the next evolution of the internet is changing everything. *CISCO White Paper*, *1*(2011), 1–11.

Ferré, X. (2000). Principios Básicos de Usabilidad para Ingenieros Software. In *JISBD* (pp. 39–46).

Georgsson, M., & Staggers, N. (2015). Quantifying usability: an evaluation of a diabetes mHealth system on effectiveness, efficiency, and satisfaction metrics with associated user characteristics. *Journal of the American Medical Informatics Association*, *23*(1), 5–11.

Hall, J. E., & Guyton, A. C. (2016). Guyton y Hall. Tratado de fisiología médica DECIMOTERCERA EDICIÓN. Retrieved from www.meddics.com

Harvey, C., Koubek, R., Bégat, V., & Jacob, S. (2016). Usability evaluation of a blood glucose monitoring system with a spill-resistant vial, easier strip handling, and connectivity to a mobile app: improvement of patient convenience and satisfaction. *Journal of Diabetes Science and Technology*, *10*(5), 1136–1141.

Heilman, K. H. (1997). The neurobiology of emotional experience. *The Neuropsychiatry of Limbic and Subcortical Disorders*, 133–142.

Mascheroni, M. A., Greiner, C. L., Petris, R. H., Dapozo, G. N., & Estayno, M. G. (2012). Calidad de software e ingeniería de usabilidad. In *XIV Workshop de Investigadores en Ciencias de la Computación*.

Mcduff, D. J., Hernandez, J., Gontarek, S., & Picard, R. W. (2016). COGCAM : Contact - free Measurement of Cognitive Stress D uring Computer Tasks with a Digital Camera. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1–5. https://doi.org/10.1145/2858036.2858247

Montero, H., & Fernández, Y. (2003). Método de test con usuarios. Retrieved from http://www.nosolousabilidad.com/articulos/test_usuarios.htm

Mozos, O. M., Sandulescu, V., Andrews, S., Ellis, D., Bellotto, N., Dobrescu, R., & Ferrandez, J. M. (2017). Stress Detection Using Wearable Physiological and Sociometric Sensors. *International Journal of Neural Systems*, *27*(02), 1650041. https://doi.org/10.1142/S0129065716500416

Perurena, L., & Moráguez, M. (2013). Usabilidad de los sitios Web, los métodos y las técnicas para la evaluación. *Revista Cubana de Información En Ciencias de La Salud (ACIMED)*, *24*(2), 176–194.

Retnani, W. E. Y., Prasetyo, B., Prayogi, Y. P., Nizar, M. A., & Abdul, R. M. (2017). Usability testing to evaluate the library's academic web site. In *Computer Applications and Information Processing Technology (CAIPT), 2017 4th International Conference on* (pp. 1–4).

Ronda, R. (2005). Productos electrónicos: principios y pautas. *Editorial Félix Varela, La Habana*.

Said, O., & Masud, M. (2013). Towards internet of things: Survey and future vision. *International Journal of Computer Networks*, *5*(1), 1–17.

Solarte, L. A., Sánchez, M., Chanchí, G. E., & Durán, D. F. (2016). Video on demand service based on inference emotions user. *Sistemas & Telemática*, *14*(38), 29–45. https://doi.org/10.18046/syt.v14i38.2286

Virkki, J., & Chen, L. (2013). Personal perspectives: Individual privacy in the IoT. *Advances in Internet of Things*, *3*(02), 21.

Zaki, N. A. A., Wook, T. S. M. T., & Ahmad, K. (2017). A usability testing of ASAH-/for children with speech and language delay. In *Electrical Engineering and Informatics (ICEEI), 2017 6th International Conference on* (pp. 1–6).

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 22/10/2018
Aceitação/Acceptance: 10/12/2018

Transferencia de conocimiento tácito entre administrador y trabajador a través de la socialización: estudio en el sector minero chileno

Raúl Caussin[1], Felipe Machorro-Ramos[2], Jorge Alfaro-Pérez[1], Patricio Ramírez-Correa[1], Giselle Painen-Aravena[1], Ari Melo Mariano[3]

jalfaro@ucn.cl, rcaussin@gmail.com, patricio.ramirez@ucn.cl, giselle.painen@alumnos.ucn.cl, felipe.machorro@udlap.mx, arimariano@unb.br

[1] Universidad Católica del Norte, Larrondo 1281, 1781421, Coquimbo, Chile.

[2] Universidad de las Américas Puebla, Ex. Hacienda Sta. Catarina Mártir S/N, 72810, San Andrés Cholula, México.

[3] Universidad de Brasilia, Campus Darcy Ribeiro, 70904970, Brasilia, Brasil.

Pages: 691–702

Resumen: Este estudio se orienta a examinar el proceso de transferencia de conocimiento tácito entre jefe y trabajador en empresas de la gran minería del cobre en Chile. En particular, el objetivo del estudio es detectar las actividades involucradas en este proceso y cuáles son los factores que posibilitan dicha transferencia. Para tal efecto se realizaron entrevistas en formato *storytelling* a trabajadores expertos, las cuales fueron analizadas cualitativamente. Como resultado se obtuvo una propuesta teórica que detalla el proceso de transferencia a través de seis elementos centrales: (i) confianza, (ii) interacción social, (iii) registro de patrones, (iv) imitación, (v) intercambio de conocimiento y (vi) aplicación. Además, se identifican cuatro facilitadores del proceso: (i) equipo de trabajo, (ii) comunicación, (iii) liderazgo y (iv) disciplina.

Palabras-clave: Conocimiento tácito, transferencia de conocimiento, gestión del conocimiento, interacción social.

Tacit knowledge transfer between manager and worker through socialization: study in the Chilean mining sector

Abstract: This study is oriented to exam the process of tacit knowledge transfer between boss and worker in large copper mining companies in Chile. The study aim is to detect the activities involved in this process and the factors that enable the transfer. For this purpose, interviews were conducted in storytelling format to expert workers, which were analyzed qualitatively. As a result, a theoretical proposal was obtained detailing the transfer process through six central elements: (i) confidence, (ii) social interaction, (iii) pattern registration, (iv) imitation, (v) knowledge exchange and (vi) application. In addition, four facilitators of the process are identified: (i) teamwork, (ii) communication, (iii) leadership, and (iv) discipline.

Keywords*:* Tacit knowledge, knowledge transfer, knowledge management, social interaction.

1. Introducción

El conocimiento y su gestión han sido un tema de amplia relevancia en las últimas tres décadas, tanto para académicos como gestores en las organizaciones. Esta importancia ha sido impulsada por la transformación de la sociedad industrial a una basada del conocimiento, en la cual el valor de los activos intangibles es cada vez mayor. Es por esto que el conocimiento, como un activo intangible, se constituye en una fuente de ventaja competitiva sostenible para las empresas (Leonard y Sensiper, 1998; Stewart, 1997; Joia y Lemos, 2010).

La literatura distingue entre dos tipos de conocimiento: el conocimiento tácito y el conocimiento explícito (Nonaka y Von Krogh, 2009; Van den Berg, 2013). Existe además cierto consenso en la idea de que el conocimiento tácito es la fuente principal de ventaja competitiva para las empresas, debido a la dificultad que implica la codificación del mismo, lo que impide una fácil imitación por parte de los competidores (Barney, 1991; Grant, 1996); sin embargo, esta característica al mismo tiempo obstaculiza su transferencia al interior de las organizaciones entre sus miembros (Ranucci y Souder, 2015). Esto implica que, para ser una verdadera fuente de ventaja competitiva sostenible, el conocimiento tácito debe poder transferirse dentro de la empresa. Es por ello que se han buscado formas de transferencia de conocimiento tácito entre empleados y de esta manera evitar la pérdida de conocimiento organizacional (Joia y Lemos, 2010).

Algunos de los estudios previos identificados proponen como facilitadores de la transferencia de conocimiento tácito características personales de los participantes en el proceso, tales como la confianza, participación y diligencia (Foos *et al*., 2006). Por su parte, Holste y Fields (2010) distinguen entre los efectos que tiene la confianza basada en el afecto y aquella basada en el conocimiento. Sus conclusiones sugieren que la primera facilita la voluntad de compartir el conocimiento, mientras que la segunda juega un papel más importante en la voluntad de utilizar el conocimiento transferido. Otros autores estudian la importancia del contexto en la transferencia del conocimiento (Yakhlef, 2007) y características organizacionales tales como la estrategia de gestión y estructura organizacional de la compañía (Joia y Lemos, 2010). Finalmente, Wijetunge (2012) explora el uso de las historias organizacionales (*storytelling*) como un método de transferencia de conocimiento tácito en el ámbito bibliotecario, transfiriendo el conocimiento de empleados con experiencia a nuevos funcionarios.

El presente estudio aborda la transferencia del conocimiento tácito al interior de la organización, específicamente el conocimiento tácito de los jefes expertos a trabajadores novatos. Para ello fue realizada una investigación de carácter cualitativo en el contexto de las empresas del rubro industrial minero en Chile. En este sector no han sido detectados estudios previos acerca de los procesos de transferencia del conocimiento; además, la pérdida de conocimiento en este tipo de organizaciones por la rotación de personal y la falta de mecanismos de transferencia, hacen el tema relevante para los gestores de estas organizaciones.

Por tanto, el objetivo de este estudio es detectar las actividades involucradas en la transferencia de conocimiento tácito entre el jefe y trabajador, y cuáles son los factores que posibilitan dicha transferencia.

2. Marco Teórico

2.1. Conocimiento Tácito y Explícito

Existen en la literatura diversas definiciones del conocimiento. Una concepción habitual es que el conocimiento se deriva de la información, la cual proviene a su vez de datos (Roberts, 2000). Sin embargo, en el concepto debe considerarse además el impacto del factor humano en la transformación de la información en conocimiento. Para ello es necesario tomar en cuenta la experiencia, conocimiento y aprendizaje de las personas (Joia y Lemos, 2010). Es por ello que de acuerdo con Nonaka y Takeuchi (1995) el conocimiento se genera solo a partir del flujo de la información que es fijado en el receptor de tal conocimiento.

El conocimiento se clasifica de acuerdo con el grado en el cual puede ser estructurado y codificado en explícito y tácito. El conocimiento explícito es aquel que se expresa en lenguaje formal, palabas, símbolos y números. Tiene naturaleza impersonal y formal y es susceptible de ser almacenado dentro de una base de datos, documentos, patentes, fórmulas, etc., las cuales se puede transmitir fácilmente dentro de la organización (Holste y Fields, 2010; Joia y Lemos, 2010). Por su parte, el conocimiento tácito es difícil de expresar en lenguaje formal, su fuente es la experiencia, percepciones y valores individuales, además depende del contexto en el cual es generado (Nonaka y Takeuchi, 1995; Joia y Lemos, 2010; Leonard y Sensiper, 1998). Tal como se ha mencionado, estas características lo convierten en una fuente de ventaja competitiva, pero a la vez lo hacen difícil de transferir a los miembros de la organización.

2.2. Transferencia de Conocimiento Tácito

Michael Polanyi definió al conocimiento tácito como aquel que no puede ser articulado o verbalizado: "sabemos más de lo que podemos contar" (1966, p.4). Desde este enfoque, existen autores que argumentan que el conocimiento tácito no es traducible y que únicamente puede ser obtenido a partir de la propia experiencia (Tsoukas, 1996), mientras que existe una corriente de estudios que afirma que el conocimiento tácito puede traducirse y por lo tanto transferirse (Weick, 1995; D'Eredita y Barreto, 2006; Wu *et al.* ,2010). Sin embargo, resulta difícil clasificar dicotómicamente todo el conocimiento existente en las organizaciones, es decir, es raro encontrar conocimiento tácito absoluto o conocimiento explícito absoluto. En otras palabras, como Polanyi (1966) señala, la distinción entre conocimiento explícito y tácito no debe ser vista como una dicotomía, sino como un espectro con los dos tipos de conocimiento como los polos en los extremos. Bajo esta perspectiva, podría afirmarse que todo conocimiento es susceptible de ser codificado y transferido, sin embargo, cuanto más tácito sea el conocimiento más difícil será transferirlo.

Hansen *et al.* (1999) proponen dos estrategias para la transferencia de conocimiento organizacional: la codificación y la personalización. En la primera, el conocimiento debe ser estandarizado, estructurado y almacenado en sistemas de información. De esta manera, es posible acceder al conocimiento a través de un sistema de indización eficiente y puede ser distribuido mediante redes de datos. Esta estrategia de transferencia hace énfasis en la transformación del conocimiento tácito en explícito el cual puede reutilizarse, no existe sin embargo algún proceso de personalización ni adaptación del

conocimiento a nuevos contextos o necesidades de los miembros de la organización o clientes. Por otro lado, en la personalización, el énfasis está en la transferencia de conocimiento tácito de una persona a otra. En este caso, los sistemas de almacenamiento son menos robustos. Las herramientas utilizadas priorizan el contacto personal, ya que el conocimiento tácito se adquiere a través de procesos individuales internos tales como la experiencia, la reflexión, la internalización y el talento individual (Gordon y Grant, 2005; Nonaka, 2007; Nonaka y Takeuchi, 1995). De esta forma, la interacción cara a cara con frecuencia es el método principal para la transferencia de conocimiento tácito. En general, se identifican facilitadores del proceso de transferencia de conocimiento tácito.

2.3. Facilitadores para la Transferencia del Conocimiento Tácito

Para Smedlund (2008), la transferencia del conocimiento tácito es facilitada por la interacción social dentro de las organizaciones, que se caracteriza por vínculos basados tanto en relaciones interpersonales como relaciones de trabajo de largo tiempo. Sin embargo, la importancia del contexto no ha escapado a la atención de los teóricos de la transferencia del conocimiento. Algunos investigadores, en el ámbito de la transferencia de conocimiento organizacional, parecen converger en la idea de la adaptación al contexto del destinatario, que es un mecanismo necesario para aumentar la aceptación (Yakhlef, 2007). Los facilitadores de la transferencia del conocimiento tácito ayudan a establecer ambientes cuyo contexto de trabajo está orientado al conocimiento, transferencia del mismo, la integración socio-cultural entre los trabajadores y gestores de proyectos. Esto permite forjar lazos de trabajo que desarrollan un mayor flujo de experiencia entre las personas y la formación de alianzas estratégicas dentro de las unidades de negocio. También se crean comunidades de práctica donde se aprende a contextualizar dichos conocimientos tácitos para aplicarlos de manera ordenada e inteligente. La importancia que tienen conceptos tales como confianza, equipo de trabajo, comunicación, disciplina y liderazgo está sustentada por investigaciones vinculadas con las relaciones sociales y humanas. De acuerdo a esto último, estos conceptos están presentes en la mayoría de las actividades donde se comparte experiencias laborales. A continuación se enuncian brevemente estos facilitadores.

Equipo de Trabajo: Un equipo de trabajo es primordial para llevar a cabo tareas o proyectos las unidades de negocios. Krause (2001) menciona dos aspectos funcionales que caracterizarían una comunidad saludable, susceptibles de ser aplicadas a los equipos de trabajo. Por un lado, la existencia de actividades conjuntas que facilitan la integración social; por otro lado, en el plano afectivo, se mencionan la amistad, la lealtad, el amor y la responsabilidad.

Comunicación: Es fundamental en las organizaciones, dado que hay que lograr incentivar y fortalecer la comunicación al interior de los equipos de trabajo. El aspecto principal es propiciar un óptimo clima laboral que facilite las relaciones humanas. Martínez y Nosnik (1998) hablan de comunicación como un proceso por medio del cual una persona se pone en contacto con otra a través de un mensaje, y espera que esta última de una respuesta, sea una opinión, actitud o conducta.

Disciplina: Peter Senge (1995) establece que practicar una disciplina equivale a seguir un camino de aprendizaje y desarrollo que no tiene fin. Una disciplina es un conjunto de técnicas basadas en una teoría, cuya práctica exige estudio y concentración. De esa manera, nuestra capacidad evoluciona de tal modo que aprendemos a mirar el mundo de otra manera.

Liderazgo: El líder debe contar con un equipo de trabajo altamente motivado. Debe destacar la personalidad, la cual les permite a los trabajadores comprender y ser comprendidos de manera tal de entablar una sinergia. Mintzberg (1992) define el poder del liderazgo como "la capacidad de afectar el comportamiento de las organizaciones... tener poder es tener la capacidad de conseguir que determinadas cosas se hagan, de causar efecto sobre las acciones y decisiones que se toman".

3. Metodología

3.1. Participantes

Para llevar a cabo la presente investigación se llevó a cabo la selección de una muestra trabajadores expertos de empresas pertenecientes al sector industrial minero en la ciudad de Antofagasta Chile. Para la selección de los sujetos a entrevistar se tuvieron en cuenta dos criterios: primero, cada individuo debería contar con al menos cinco años de experiencia laboral en el rubro industrial minero; segundo, para ser susceptible de selección debería ocupar un cargo de jefatura o directivo (jefe de unidad, gerente, subgerente, superintendente, etc.) dentro de la organización en la que actualmente labora.

La participación fue de manera voluntaria, la muestra final estuvo conformada por once individuos del sexo masculino que cumplen con las características mencionadas los cuales accedieron a contestar la entrevista.

3.2. Instrumento

Se utilizó una entrevista de tipo semiestructurada con modalidad *Storytelling*. Esta modalidad permite al entrevistado sentirse cómodo con el tema, logrando que entregue la mayor información posible sin salirse de los límites establecidos previamente, además cuenta con una variedad de elementos narrativos que son propicios para un estudio cualitativo. De esta manera, la entrevista estaba enfocada a que cada sujeto se remontara a sus comienzos laborales en el rubro industrial minero e hiciera énfasis en los mecanismos de transferencia de conocimiento que fueron eficaces en su experiencia personal para poder llegar a ser expertos en sus actividades.

El procedimiento para la validación del instrumento se basó en realizar entrevistas de prueba la cual contó con los siguientes pasos: se elaboró un conjunto de conceptos relacionados al tema de investigación, para a continuación concretar en algunas preguntas particulares. Luego se desarrolló una prueba piloto que consistió en realizar dos entrevistas a dos personas de prueba en su lugar de trabajo relacionadas al rubro. Posteriormente se estudiaron los resultados para validar el tipo de entrevista seleccionada y que los límites establecidos eran los apropiados. Finalmente, una vez realizado los ajustes necesarios se procedió a aplicar la entrevista a la muestra seleccionada.

3.3. Procedimiento

Debido a la modalidad de elicitación aplicada, se pidió que previo a la entrevista las personas respondieran en una forma narrativa al tema de investigación. El lugar para realizar la entrevista fue cada una de las dependencias donde trabajan los entrevistados.

Cada entrevista tuvo una duración aproximada entre 30 a 45 minutos. El soporte tecnológico utilizado fue una grabadora digital, así como la utilización de notas para complementar la grabación. Las transcripciones de las entrevistas se realizaron en modalidad *verbatim*. Este tipo de transcripción es usado para replicar exactamente cada palabra usada en la entrevista realizada. Una transcripción palabra por palabra garantiza la exactitud de lo que el entrevistado expresó verbalmente.

Posteriormente, las transcripciones fueron almacenadas en la herramienta de análisis cualitativo Atlas.ti. Este programa computacional provee un entorno en el cual las transcripciones y los audios pueden ser analizados bajo diferentes métodos. Integra funciones visuales para elaborar gráficos conceptuales, categorías de familias conceptuales, codificaciones, como también elabora estadísticas de ocurrencia cuando se realizan codificaciones a los audios.

El análisis aplicado a las transcripciones fue de carácter categorial temático, una metodología que actúa sobre el texto, el cual constituye un relato significativo (discurso o narración) para obtener la mayor información posible en relación al tema de investigación. Dicho análisis aporta aspectos relevantes de la relación de palabras, como de entidades significativas con respecto al tema central de la investigación. Los objetivos y los supuestos de esta modalidad tratan de construir categorías y dimensiones (cuando es posible) que ordenen el análisis y produzcan temas o tópicos con sentido o significado en torno a los sujetos de estudio. Esto con el fin de establecer una percepción fidedigna de aquello relatado durante la sesión de entrevista.

4. Resultados

Las observaciones obtenidas a través del estudio de campo describen experiencias laborales en donde los entrevistados pudieron adquirir conocimiento tácito. Asociado a estas experiencias se identifican dos patrones. Por un lado, se observa la descripción que los sujetos dieron principalmente en relación a la personalidad que mostraba el jefe durante el proceso de transferencia y, por otro lado, al ambiente laboral organizacional que imperaba al momento del proceso de transferencia de conocimiento tácito.

Las proposiciones establecidas en base a los elementos encontrados en las entrevistas, conformaron dos ejes principales. El primero, los facilitadores, son el conjunto de factores que influyen en el proceso de transferencia del conocimiento tácito identificadas por los entrevistados. El segundo, denominado componentes de interacción organizacional y social, son elementos inferidos a partir del estudio de campo y definen el marco del proceso social de transferencia de conocimiento tácito, ver Figura 1.

4.1. Facilitadores

Los facilitadores son parte esencial del ciclo de gestión del conocimiento ya que dan como resultado al elemento Confianza.

Equipo de Trabajo: Los equipos crean condiciones de aprendizaje, que facilitan el intercambio de experiencias laborales.

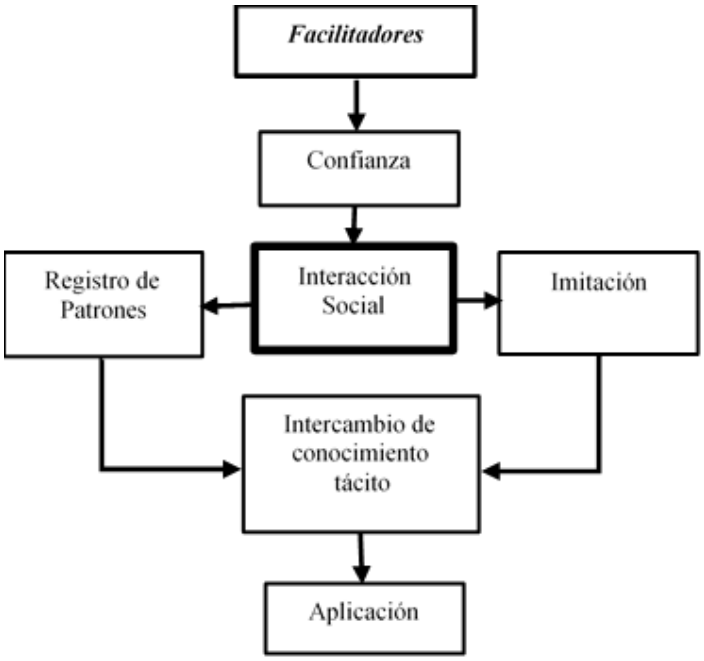


Figura 1 – Propuesta teórica para la transferencia del conocimiento tácito de jefe a trabajadores.

Comunicación: Los equipos generan espacios de comunicación que facilitan la interacción social de los trabajadores.

Liderazgo: Los líderes proveen el ambiente propicio que facilitan la resolución de conflictos y la transferencia de conocimiento tácito.

Disciplina: Los líderes fomentan la disciplina laboral mediante conductas susceptibles de ser imitadas por sus trabajadores.

4.2. Componentes de Interacción Organizacional y Social

Confianza: Los trabajadores que confían en sus jefes aprenden más de ellos, que los trabajadores que no confían en ellos.

Interacción Social: La interacción social afecta positivamente el intercambio del conocimiento tácito, de manera inconsciente, entre los trabajadores.

Registro de Patrones: Los trabajadores que aprenden de sus jefes registran de manera inconsciente patrones de comportamiento que afectan positivamente su trabajo.

Imitación: En ciertas situaciones, los trabajadores imitar conductas o actitudes de ciertas personas de su pasado laboral.

Intercambio de Conocimiento Tácito: La interacción social, sustentada en la confianza, decanta en el registro de patrones y en la imitación de conductas o actitudes, lo que finalmente concluye en el intercambio de conocimiento tácito.

Aplicación: Finalmente, para que el proceso de transferencia de conocimiento tácito esté completo, los aprendices deben ser capaces de aplicar el saber transmitido. De acuerdo con los entrevistados, con base en ciertas formas o estructuras organizacionales, los trabajadores pueden aplicar este activo intangible.

5. Discusión

En el modelo presentado, la interacción social vincula directamente la actividad de transferencia del conocimiento tácito con la relación interpersonal para difundir experiencia y habilidades tácitas, activos de difícil gestión. Esta proposición, en términos generales, ayuda a destacar el contexto donde el conocimiento tácito es fundamental. Así lo destaca el modelo de Boisot (1987), donde su elemento de "aprendizaje social" remarca que la interacción social entre personas es un excelente mecanismo para transferir conocimiento intangible. También Nonaka y Takeuchi (1995) lo vinculan a su modelo, en donde la "socialización" facilita las relaciones humanas en la transferencia de experiencias y habilidades tácitas. Por su parte, McAdam y McCreedy (1999) enfatizan que la interacción humana fomenta la construcción del conocimiento dentro de la organización.

La proposición "Confianza" se relaciona con la opinión favorable o seguridad que un trabajador tiene sobre algo o alguien en una determinada situación. Es una cualidad propia del ser humano que se manifiesta de manera consciente y voluntaria, esto supone trabajo y esfuerzo en conseguirla. Así lo muestran Holste y Fields (2010) que establecen que tanto la confianza basada en el afecto como aquella basada en lo cognitivo entre los trabajadores, tiene una relación positiva en la disposición de usar y compartir el conocimiento tácito. McAllister (1995) también demostró que la confianza basada en el afecto y basada en lo cognitivo son formas de confianza interpersonales, ambas están relacionadas con conductas personales dirigidas a trabajadores de la organización.

Por su parte, "Registro de patrones" es una proposición que tiene como función servir como repositorio de experiencias para la resolución de problemas. La integración de estas experiencias al negocio se debe en gran medida a la interacción social entre los trabajadores. El modelo de Hedlund y Nonaka (1993) describe cuatro niveles de agentes del conocimiento: individual, grupal, organizacional y dominio inter organizacional. Siendo el "grupal" el que cumple la función – entre otras - de almacenar las experiencias de los trabajadores cuando se realiza esta actividad.

La proposición "Facilitadores" está orientada a la organización cuyos elementos son conceptos que están en la mente de los trabajadores, y los utilizan de manera espontánea en sus trabajos. Por ello, el modelo Skandia, desarrollado por Roos y Roos (1997), se sustenta en un fuerte énfasis en la medición asociada a elementos como capital humano, cliente y estructura organizacional. Esto lo enfatiza más en el componente "capital humano" (equipo de trabajo, comunicación, liderazgo), debido a que el flujo principal del conocimiento está estrechamente vinculado al activo tácito del trabajador.

Finalmente, la proposición "Aplicación" tiene como función desarrollar propuestas para integrar el conocimiento tácito al negocio, tener elementos organizacionales desarrollados en las personas y herramientas sociales o tecnológicas. La aplicación de los activos intangibles obedece a la experticia que tienen los trabajadores sobre cuando aplicar un conocimiento en particular. De forma complementaria, el marco de Stankosky (2005) se centra en tres disciplinas: factores habilitantes, gestión del conocimiento y disciplinas de estudios. Pero es en los "factores habilitantes" donde queda de manifiesto que para la aplicación de recursos intangibles es necesario manejar cuatro fundamentos esenciales: el liderazgo, la estructura organizacional, la infraestructura tecnológica y el aprendizaje.

6. Conclusiones

Este estudio examinó un tema de gran importancia para las organizaciones del rubro industrial-minero chileno, como es la transferencia del conocimiento tácito entre jefes y trabajadores. En el trabajo se ha explorado este fenómeno, tanto de la evidencia desde las fuentes primarias de información, como de la literatura de cómo las personas de un determinado contexto transfieren conocimiento de difícil articulación y administración, para su posterior integración al ámbito laboral y social.

En este estudio la propuesta teórica de transferencia del conocimiento tácito utiliza la socialización para entender la naturaleza dinámica del conocimiento y mostrar qué elementos facilitan la interacción en este proceso. Las distintas experiencias narradas por los entrevistados, describen que por una parte las organizaciones no consideraban en aquellos tiempos el conocimiento como un activo primordial ni como una fuente de ventaja competitiva, y por otro lado no fomentaba en los trabajadores el compartir de una manera formal las experiencias con otros trabajadores de manera tal de enriquecer el conocimiento social-organizacional.

Los diferentes patrones en qué el experto aplica su experticia al trabajo, condujeron a explicar detalles de sus experiencias y establecer diferentes realidades en sus carreras profesionales. Las vivencias laborares de los entrevistados permiten identificar elementos que brindan indicios de cómo se transfiere el conocimiento tácito entre jefe y trabajador. En concreto, como resultado del estudio se obtuvo una propuesta que detalla el proceso de transferencia a través de seis elementos centrales: (i) confianza, (ii) interacción social, (iii) registro de patrones, (iv) imitación, (v) intercambio de conocimiento y (vi) aplicación. Además, se identifican cuatro facilitadores del proceso: (i) equipo de trabajo, (ii) comunicación, (iii) liderazgo y (iv) disciplina.

De los hallazgos de estudio se puede inferir que las personas entrevistadas valoran las relaciones humanas y la convivencia en el trabajo, y a su vez, cuentan con la disposición a cumplir las metas como equipo y mantienen una excelente comunicación tanto con sus pares como con sus jefes. Incluso, para algunos, la confianza que facilita obtener conocimiento tácito se asocia a valores familiares.

Dentro del ámbito de futuras investigaciones vinculadas al presente trabajo, se sugiere profundizar el estudio de modelos de gestión del conocimiento inteligente, orientados

a aprovechar los recursos de conocimiento tácito. Como también, explorar como la manera la interacción social en la actividad de transferencia de conocimiento tácito puede facilitar la conformación de comunidades de prácticas al interior de las unidades de negocio.

Referencias

Rocha, Á. (2012). Framework for a Global Quality Evaluation of a Website. *Online Information Review, 36*(3), 374-382. doi:10.1108/14684521211241404

Antunes, A. A. (2004). Sistemas XYZ. In Sousa A. J. (Ed.), *Tecnologias Internet*. Lisboa: Editora Xxxpto.

Barney, J.B. (1991). Firm Resources and Sustained Competitive Advantage. Journal of Management, 17(1), pp.99-120.

Van den Berg, H.A. (2013). Three shapes of organisational knowledge. Journal of Knowledge Management, 17(2), pp.159-174.

Boisot, M. (1987). Information and Organizations: The Manager as Anthropologist, London: Fontana/Collins.

D'Eredita, M.A. y Barreto, C. (2006). How Does Tacit Knowledge Proliferate? An Episode-Based Perspective. Organization Studies, 27(12), pp.1821-1841.

Foos, T., Schum, G. y Rothenberg, S. (2006). Tacit knowledge transfer and the knowledge disconnect. Journal of Knowledge Management, 10(1), pp.6-18.

Gordon, R. y Grant, D. (2005). Knowledge Management or Management of Knowledge? Why People Interested in Knowledge Management Need to Consider Foucault and the Construct of Power. TAMARA: Journal of Critical Postmodern Organization Science, 3(2), pp.27-38.

Grant, R.M. (1996). Toward a knowledge-based theory of the firm. Strategic Management Journal, 17(Winter Special Issue), pp.109-122.

Hansen, M., Nohria, N. y Tierney, T. (1999). What's your strategy for managing knowledge? Harvard Business Review, Marzo-Abri, pp.1-10.

Hedlund, G. y Nonaka, I. (1993). Models of Knowledge Management in the West and Japan. En Implementing Strategic Process, Change, Learning and Cooperation. London: McMillan, pp. 117-144.

Holste, J.S. y Fields, D. (2010). Trust and tacit knowledge sharing and use. Journal of Knowledge Management, 14(1), pp.128-140.

Joia, L.A. y Lemos, B. (2010). Relevant factors for tacit knowledge transfer within organisations. Journal of Knowledge Management, 14(3), pp.410-427.

Krause, M. (2001). Hacia una redefinición del concepto de comunidad. Revista de Psicología de la Universidad de Chile, 10(2), pp.49-60.

Leonard, D.A. y Sensiper, S. (1998). The Role of Tacit Knowledge in Group Innovation. California Management Review, 40(3), pp.112-132.

Martínez, A. y Nosnik, A. (1998). Comunicación organizacional práctica, México, D.F.: Trillas.

McAdam, R. y McCreedy, S. (1999). A critical review of knowledge management modelsnull. The Learning Organization, 6(3), pp.91-101.

McAllister, D.J. (1995). Affect- and Cognition-Based Trust as Foundations for Interpersonal Cooperation in Organizations. The Academy of Management Journal, 38(1), pp.24-59.

Mintzberg, H. (1992). El poder de la negociación, Barcelona: Ariel.

Nonaka, I. (2007). The Knowledge-Creating Company. Harvard Business Review, 85(7/8), pp.162-171.

Nonaka, I. y Von Krogh, G. (2009). Perspective-tacit knowledge and knowledge conversion: Controversy and advancement in organizational knowledge creation theory. Organization science, 20(3), pp.635-652.

Nonaka, I. y Takeuchi, H. (1995). The knowledge-creating company: How Japanese companies create the dynamics of innovation, New York: Oxford University Press.

Polanyi, M. (1966). The Tacit Dimension, New York: Anchor Books.

Ranucci, R.A. y Souder, D. (2015). Facilitating tacit knowledge transfer: routine compatibility, trustworthiness, and integration in M y As. Journal of Knowledge Management, 19(2), pp.257-276.

Roberts, J. (2000). From Know-how to Show-how? Questioning the Role of Information and Communication Technologies in Knowledge Transfer. Technology Analysis y Strategic Management, 12(4), pp.429-443.

Roos, J. *et al.* (1997). Capital Intelectual. El valor intangible de la empresa., Barcelona: Paidós.

Senge, P.M. (1995). La quinta disciplina en la práctica: cómo construir una organización inteligente, Buenos Aires: Granica Ediciones.

Smedlund, A. (2008). The knowledge system of a firm: social capital for explicit, tacit and potential knowledge. Journal of Knowledge Management, 12(1), pp.63-77.

Stankosky, M. (2005). Creating the discipline of Knowledge Management, Burlington, M.A.: Elsevier Butterworth-Heinemann.

Stewart, T.A. (1997). Capital Intelectual: la nueva riqueza de las organizaciones, Barcelona: Ediciones Granica S.A.

Tsoukas, H. (1996). The firm as a distributed knowledge system: A constructionist approach. Strategic Management Journal, 17(S2), pp.11-25.

Weick, K.E. (1995). Sensemaking in organizations. Thousand Oaks California: Sage Publications.

Wijetunge, P. (2012). Organizational *storytelling* as a method of tacit-knowledge transfer: Case study from a Sri Lankan university. The International Information y Library Review, 44(4), pp.212-223.

Wu, C.-H., Kao, S.-C. y Shih, L.-H., 2010. Assessing the suitability of process and information technology in supporting tacit knowledge transfer. Behaviour y Information Technology.

Yakhlef, A. (2007). Knowledge transfer as the transformation of context. The Journal of High Technology Management Research, 18(1), pp.43-57.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 21/09/2018
Aceitação/Acceptance: 09/11/2018

Enfoques Inteligentes para Identificar Estilos de Aprendizaje de los estudiantes mediante las Emociones en un salón de clases

Jorge Cordero[1], Jose Aguilar[2], Kristell Aguilar[2]

jmcordero@utpl.edu.ec, aguilar@ula.ve, kristell153@gmail.com

[1] DCCE, Universidad Técnica Particular de Loja, San Cayetano Alto, 11-01-608, Loja, Ecuador.

[2] CEMISID, Universidad de Los Andes, La Hechicera, 5101, Mérida, Venezuela.

Pages: 703–716

Resumen: En este artículo, se propone un patrón jerárquico para identificar el estilo de aprendizaje de los estudiantes en un salón de clases, que se compone de dos niveles, uno para reconocer el estado emocional, y otro para identificar el estilo de aprendizaje. Cada nivel está definido por diferentes tipos de descriptores, que se perciben desde un enfoque multimodal. Además, utilizando el patrón jerárquico, se analizan dos enfoques, para modelar y reconocer el estilo de aprendizaje, el cual puede ser visual, auditivo, lectura/escritura, kinestésico. Uno de los enfoques es basado en lógica difusa, y el otro basado en crónicas, los cuales explotan la idea de recursividad y aprendizaje en el proceso de reconocimiento. Finalmente, considerando el entorno dinámico en el que se desarrollan las clases, se comparan los enfoques, en cuanto a sus capacidades para aprender y determinar los estilos de aprendizaje; y para comunicar dicha información claramente en un salón de clases inteligentes, basada en las emociones reconocidas.

Palabras-clave: Patrones jerárquicos, estilos de aprendizaje, reconocimiento de emociones, lógica difusa, crónicas, reconocimiento dinámico de patrones.

Intelligent Approaches to identify Student Learning Styles through Emotions in a Classroom

Abstract: In this article, a hierarchical pattern is proposed to identify the learning style of students in a classroom, which is composed of two levels, one to recognize the emotional state, and another to identify the learning style. Each level is defined by different types of descriptors, which are perceived from a multimodal approach. In addition, using the hierarchical pattern, two approaches are analyzed to model the learning style, which can be visual, auditory, reading/writing, kinesthetic. One of the approaches is based on fuzzy logic, and another based on chronicles, which exploit the idea of recursion and learning in the recognition process. Finally, considering the dynamic environment of a classroom, the approaches are compared, in terms of their abilities to learn and to determine the learning styles;

and to communicate that information clearly in a smart classroom, based on the recognized emotions.

***Keywords**:* Hierarchical patterns, learning style, emotion recognition, fuzzy logic, chronicles, dynamic pattern recognition.

1. Introducción

Actualmente, en la mayoría de las situaciones, el profesor actúa como facilitador entre el estudiante y el curso de aprendizaje, y a través de un análisis del comportamiento, de los estilos de aprendizaje, de las emociones y de los logros del estudiante, realiza ajustes constantes en el proceso de enseñanza, con el fin de satisfacer las necesidades y objetivos de los alumnos. En particular, en el proceso de aprendizaje es en una de las actividades en las cuales las emociones tienen una influencia directa. (Coelho, Sousa, & Figueira, 2014; Javanbakht et al., 2017; Willingham, Hughes, & Dobolyi, 2015) argumentan que la emoción juega un papel importante en el proceso de decisión y la adquisición de conocimiento de una persona. Por lo tanto, las emociones pueden influir en el proceso de aprendizaje. Así, comprender la manera en que las emociones cambian el proceso de aprendizaje de los estudiantes, no solo es muy importante, sino que también puede permitir mejorar los procesos de enseñanza-aprendizaje existentes.

Existen trabajos iniciales acerca del reconocimiento de las emociones en el aula, por ejemplo, en (Cordero & Aguilar, 2016) proponen un modelo de reconocimiento del estado emocional, utilizando crónicas y patrones estáticos. Por otro lado, en (Faria et al., 2017; Magdin, Turčáni, & Hudec, 2016) se muestra cómo las emociones son un tema clave, no solo en una interacción humano-computadora, sino también para determinar los estilos de aprendizaje. En (Fuentes et al., 2017) presentan una revisión de los trabajos en reconocimiento de emociones, centrándose en aquellos que influyen en el rendimiento del estudiante. Existen otros trabajos acerca del reconocimiento de emociones de los alumnos, pero en general, proponen modelos unimodales, o solo estudian las emociones, sin considerar aplicabilidad en el contexto educativo (Bettadapura, 2012; Pekrun & Linnenbrink-Garcia, 2012).

Por otro lado, se reconoce que los rasgos de personalidad influyen en el rendimiento de las personas en varias áreas de la vida (Kumar, Bakhshi, & Rani, 2009; Santos, 2010). En particular, los estilos de aprendizaje se definen como las distintas maneras en que una persona puede aprender (Morgan, 2012). Por lo tanto, diferentes personas aprenden de diferentes maneras. Todos tienen una combinación de estilos de aprendizaje, pero algunas personas pueden descubrir que tienen un estilo de aprendizaje dominante. Otros pueden descubrir que tienen un estilo de aprendizaje diferente en diferentes circunstancias. Existen algunos modelos que representan la forma en que las personas aprenden, como el modelo Felder-Silverman (Felder, Silverman, & others, 1988), el modelo Learning Style Inventory (LSI) (A. Y. Kolb & Kolb, 2005), o el modelo VARK presentado por Neil Fleming (Fleming & Baume, 2006), entre otros.

La emoción puede influir en el estilo de aprendizaje del alumno durante el proceso de aprendizaje (D. A. Kolb, 2014; Suliman, 2010). Los factores emocionales pueden influir en la motivación del alumno, así como en el resultado del aprendizaje en general (Shen, Wang, & Shen, 2009). Los alumnos con problemas emocionales pueden tener dificultades para aprender, a menudo son infelices, tienen pocos amigos, casi nunca

se sienten halagados, son inmaduros y siempre están irritados (Shen et al., 2009). La emoción tiene un papel importante en la capacidad de aprendizaje de los estudiantes. Un estudiante que siente ira, tiende a ser incapaz de enfocarse en el aprendizaje (Goralnik, Millenbah, Nelson, & Thorp, 2012; Mahasneh, 2013).

En ese sentido, la principal contribución de esta investigación es proponer un patrón jerárquico para modelar el estilo de aprendizaje, que considera dos niveles de reconocimiento, uno para reconocer el estado emocional, y otro para identificar el estilo de aprendizaje basado en el modelo VARK (Fleming & Baume, 2006). Además, se analizan dos técnicas para reconocer el estilo de aprendizaje, una basada en lógica difusa y otra basada en crónicas. El artículo compara las técnicas, y evalúa sus capacidades para definir estrategias de aprendizaje adaptables al usuario en un salón de clases.

La estructura del artículo es la siguiente: en la sección 2 se presenta la arquitectura del sistema. La sección 3 describe la definición formal del patrón jerárquico para reconocer las emociones e identificar el estilo de aprendizaje de los estudiantes en un salón de clases. En la sección 4 se presenta los enfoques para modelar los estilos de aprendizaje, uno basado en crónicas y otro basado en lógica difusa. La sección 5 presenta una comparación de enfoques considerando las capacidades de cada técnica, y la sección 6 presenta una comparación con otros enfoques, para identificar los estilos de aprendizaje y reconocer emociones en el contexto educativo. Finalmente, en la sección 7 se presentan las conclusiones y trabajos futuros derivados de la investigación.

2. Arquitectura del sistema.

La arquitectura propuesta se compone de un patrón jerárquico para relacionar los dos aspectos principales a analizar en el proceso de enseñanza-aprendizaje de un estudiante: uno para identificar sus estilos de aprendizaje, y el otro para reconocer sus emociones. La Figura 1 muestra la arquitectura. El patrón para identificar los estilos de aprendizaje se basa en el modelo VARK (Visual, Aural, Read/Write, Kinesthetic) (Morgan, 2012; Fleming, 2008). Dicho patrón está compuesto por diferentes descriptores relacionados entre sí, que requieren un enfoque multimodal (facial, acústico, lenguaje corporal, entre otros) para ser percibidos.

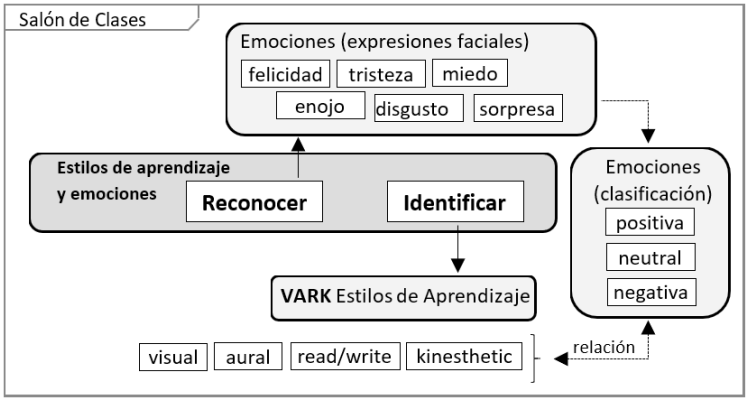


Figura 1 – Arquitectura del sistema

La elección de la estrategia VARK se debe al hecho de que permite descubrir las preferencias de aprendizaje de una manera muy simple. Para identificar el estilo de aprendizaje (visual, auditivo, de lectura/escritura, cinestésico), los alumnos deben responder el cuestionario VARK, versión 7.1, que está conformado por 16 preguntas con cuatro opciones de respuesta (Fleming, 2008). Cada opción está asociada con una preferencia de estilo de aprendizaje diferente, la persona puede elegir más de una opción de respuesta, de acuerdo con su percepción. El estilo de aprendizaje del encuestado es determinado por las respuestas más seleccionadas para cada elemento.

Dependiendo de las emociones, se afecta el comportamiento del sujeto. En general, éstas se pueden clasificar en positivas, neutras o negativas. Las emociones positivas como la felicidad, favorecen la manera de pensar, razonar y actuar de las personas. Las emociones negativas como la tristeza, miedo, y enojo, afectan al bienestar de las personas. Las emociones neutras o ambiguas, como la sorpresa, ni favorecen ni afectan la manera de pensar, sentir y actuar de las personas.

Los estudiantes con tendencias de emociones positivas pueden relacionarse con todos los tipos de estilos de aprendizaje. Por otro lado, los alumnos con emociones neutrales tienden a relacionarse con estilos de aprendizaje auditivo, lectura/escritura y cinestésico. Finalmente, los estudiantes con tendencias emociones negativas solamente tienen una relación con el estilo de aprendizaje visual.

Los profesores pueden diseñar un tratamiento apropiado para alumnos con tendencias emocionales negativas. El profesor puede minimizar el impacto negativo que puede resultar de esta tendencia emocional particular. En general, puede proporcionar métodos de enseñanza individuales apropiados de acuerdo con las necesidades de los estudiantes. También, los resultados se pueden usar como base para crear grupos de estudio que prefieran un estilo de aprendizaje similar.

3. Patrones para reconocer emociones y estilos de aprendizaje

En esta sección se presenta la definición formal del patrón jerárquico, para identificar el estilo de aprendizaje de los alumnos en un salón de clases. El patrón está basado en dos niveles, uno para reconocer las emociones, y otro para identificar los estilos de aprendizaje. En esta sección se define el patrón de cada nivel, los cuales están compuestos por diferentes descriptores, que requieren un enfoque multimodal para ser percibidos en un contexto real, como lo es un aula de clases.

3.1. Definición formal del patrón para reconocer las emociones en los estudiantes.

En este artículo se propone la definición de un patrón de las emociones de los alumnos, el cual debe poder ser usado en procesos de reconocimiento de la emoción actual de un estudiante. Particularmente, se utilizan las seis emociones básicas definidas en la literatura (Ekman, 1992; Ekman & Friesen, 2003): felicidad, tristeza, miedo, enojo, disgusto y sorpresa. Los descriptores que definen este patrón se describen en la Tabla 1.

Descriptor	Descripción
Expresión facial	Caracteriza la expresión facial actual. Ejemplo: características de los ojos, boca, nariz, cejas, etc.
Voz	Caracteriza el tono de voz actual. Ejemplo: tono (grave, agudo), intensidad (alta, media, baja), velocidad de locución.
Lenguaje corporal	Define el lenguaje corporal actual del estudiante. Ejemplo: características de los brazos, cabeza, puños, cuerpo, gestos.
Texto	Define el resultado del análisis sentimental que se hace de los textos redactados por los alumnos, ya sea en sus trabajos, redes sociales, correos, etc. Para ello, se usa minería de texto, diccionarios afectivos, procesamiento del lenguaje natura, entre otros. Los sentimientos se clasifican en positivos, negativos o neutros.
Emoción del estudiante	Define el estado emocional del estudiante. Ejemplo: ira, disgusto, miedo, felicidad, tristeza, sorpresa.

Tabla 1 – Descriptores para reconocer emociones del estudiante

De acuerdo con los valores actuales de los descriptores, se determina la emoción actual del estudiante. La Tabla 2 muestra ejemplos de posibles combinaciones de los valores de los descriptores para el reconocimiento de las emociones.

Identificador de eventos	Emoción	Expresión facial	Voz	Lenguaje corporal	Texto
ED1	Felicidad	Ojos abiertos, Boca con una sonrisa	Tono agudo, e Intensidad media	Cabeza levantada	Positivo (texto agradable)
ED2	Tristeza	Ojos semiabiertos, Boca cerrada	Tono grave e Intensidad baja	Cabeza baja	Negativo (desagradable)

Tabla 2 – Emociones reconocidas en función de los descriptores

3.2. Definición formal del patrón para identificar los estilos de aprendizaje en los estudiantes.

Las evaluaciones del estilo de aprendizaje pueden beneficiar tanto a los profesores como a los aprendices (Felder & Spurlin, 2005). Particularmente, los descriptores propuestos en ese nivel, permiten definir el estilo de aprendizaje de un estudiante basado en el modelo VARK. Según ese modelo, los estilos pueden ser:

- *Visual: aprender viendo*. Personas con una fuerte preferencia visual para aprender como: diferentes formatos, espacio, gráficos, tablas, diagramas, mapas mentales, algoritmos, videos, mapas y planos.
- *Aural: aprender escuchando*. Personas con una fuerte preferencia auditiva para aprender como: discusiones, historias, oradores invitados, grabaciones de audio, chat.
- *Read/write: aprender leyendo/escribiendo*. Personas con una fuerte preferencia de lectura/escritura para aprender como: artículos, diapositivas, listas, notas y textos en todos sus formatos, ya sea en forma impresa o en línea.

- *Kinesthetic aprender haciendo*. Personas con una fuerte preferencia kinestésica para aprender como: sentidos, ejercicios prácticos, ejemplos, simulaciones, casos, prueba y error.

Para deducir esos estilos, se proponen los 5 descriptores definidos en la Tabla 3.

Descriptor	Descripción
Emoción del estudiante	Define el estado emocional del estudiante. Ejemplo: ira, disgusto, miedo, felicidad, tristeza, sorpresa.
Actividades de aprendizajes	Define las actividades que prefiere realizar un estudiante para alcanzar sus objetivos de aprendizaje, tales como: Discusión en grupo, Exposiciones, Charlas, Aprendizaje Basado en problemas, Lecturas, simulaciones, etc.
Estrategias de aprendizaje	Describe el conjunto de actividades, técnicas y medios que preferido por el estudiante, con la finalidad de hacer efectivo el proceso de aprendizaje, tales como: Estrategias de ensayo, de organización, control de la comprensión, afectivas, entre otras
Herramientas de aprendizaje	Define los medios preferidos por los estudiantes para llevar a cabo su proceso de enseñanza y aprendizaje, tales como: Mapas Conceptuales, Texto Narrativo, Videos, Laboratorios, Juegos, Blog
Estilo de aprendizaje	Define el estilo de aprendizaje reconocido: visual, auditivo, de lectura/ escritura, cinestésico

Tabla 3 – Descriptores para identificar los estilos de aprendizaje

El objetivo principal del patrón jerárquico es identificar el estilo de aprendizaje. El estado actual de los descriptores está determinado por los eventos que se capturan en el aula en un momento dado. Para eso, se utiliza la información de los diferentes sensores y software que dispone el aula, para caracterizar estos eventos.

4. Enfoques para modelar y reconocer los estilos de aprendizaje de un estudiante

En esta sección se presentan dos enfoques para la modelización de los estilos de aprendizaje, uno basado en crónicas y otro basado en lógica difusa.

4.1. Basado en Crónicas.

Una crónica se puede definir como un conjunto de eventos con restricciones temporales entre ellos y con respecto al contexto, que representa una interpretación de lo que está ocurriendo en la dinámica del sistema en estudio en un momento dado (Aguilar, 2011; Dousson & Le Maigat, 2007). Para definir una crónica, normalmente se usan dos predicados: *evento* (*event*) y *mantener (hold)*. Un *evento* expresa un cambio en un atributo, por ejemplo: evento (voz (tono): (agudo, grave), t2). *Mantener* especifica que un atributo mantiene un valor durante un intervalo de tiempo. Por ejemplo: mantener (expresión facial (estudiante, sonrisa), (t2, t4)).

En general, un modelo crónica C se define por un par (S, T), donde S es el conjunto de eventos y T las restricciones temporales entre los eventos. Una instancia de crónica c de

un modelo de una crónica C es un conjunto de ocurrencias de eventos, que es consistente con las restricciones de tiempo de C.

El sistema jerárquico de reconocimiento de patrones basado en el paradigma de crónicas consta de 2 tipos de crónicas: (i) una que representa los patrones emocionales del estudiante. Su objetivo es describir las emociones del estudiante; (ii) otra que representa los patrones de los estilos de aprendizaje. Su objetivo es identificar los estilos de aprendizaje.

Cada emoción y estilo de aprendizaje del estudiante será modelado por una crónica diferente, que contiene los eventos y las relaciones temporales para reconocerlos. Cada emoción y estilo de aprendizaje puede ser reconocido por varias crónicas, y cada crónica está definida por el conjunto de descriptores que se requieran para describir los eventos del patrón modelado. Un ejemplo de una crónica del primer tipo, para reconocer una emoción, es:

Cronica felicidad {
evento (EF3, T_1),
evento (V1, T_3),
evento (LC1, T_4),
evento (TX1, T_2)
$T_1 < T_3$,
mantener (EF3, (4, 10)),
mantener (LC1, (6, 20)),
cuando reconoce {reportar evento (ED1)}}

Según la crónica, el patrón "felicidad" puede reconocerse cuando el evento de expresión facial «ojos abiertos y boca sonrisa» (EF3) ocurre en el tiempo T_1, y se mantiene entre 4 y 10 unidades de tiempo; el evento de voz «tono agudo e intensidad media» (V1) ocurre en el momento T_3, el cual es mayor a T_1; el evento lenguaje corporal "cabeza levantada" (LC1) ocurre en el momento T_4, y se mantiene entre 6 y 20 unidades de tiempo; y el evento "texto agradable" (TX1) ocurre en el tiempo T_2.

Un ejemplo de una crónica del segundo tipo, para identificar el estilo de aprendizaje visual, es:

Cronica visual {
evento (ED1, T_{ED1}),
evento (AA3, T1),
evento (EA1, T_2),
evento (HE2, T_3)
$T_{ED1} < T_1$,
$T_3 < T_2$,
mantener (AA3, (5, 15)),
cuando reconoce {reportar estilo de aprendizaje}}

La estructura de las crónicas de la segunda capa es similar. Por ejemplo, para detectar el estilo de aprendizaje visual, el evento de la emoción "felicidad" (ED1) debe ocurrir en el momento TED1, el evento que determina que la actividad de aprendizaje es "simulación" (*AA3*) debe ocurrir en el tiempo T1, mayor a T_{ED1}, y durar entre 5 a 15 unidades de tiempo,

el evento que determina que la estrategia de aprendizaje es "basado en ensayos" (EA1) debe ocurrir en el momento T2, y así sucesivamente, la.

Esta es solo una muestra de las posibles crónicas utilizadas en el artículo, donde: (i) Las emociones (enojo, felicidad, miedo, entre otras) conforman las crónicas que representan los patrones emocionales del estudiante (primer nivel); (ii) Los estilos de aprendizaje (visual, auditivo, de lectura/escritura, cinestésico) son las crónicas que representan los patrones del segundo nivel (estilos de aprendizaje).

Las crónicas para reconocer emociones, se componen de los eventos primarios capturados a través de diferentes tipos de sensores (cámaras, micrófonos, entorno virtual de aprendizaje, entre otros). Las crónicas para identificar los estilos de aprendizaje, son una mezcla de los eventos primarios y los eventos reconocidos en el contexto educativo.

4.2. Basado en Lógica Difusa

Un controlador difuso es un sistema difuso basado en reglas, compuesto por un conjunto de reglas de inferencia del tipo *SI <condición> ENTONCES <acción>*, que define las acciones de control de acuerdo con varios rangos de las variables controladas en el problema. En general, la estructura básica de un sistema de inferencia difusa consiste en tres componentes: una base de reglas, que contiene las reglas difusas; un conjunto de variables difusas, cada una definida por un conjunto de funciones de membresía; y un mecanismo de razonamiento que realiza el proceso de inferencia.

Se propone modelar el modelo jerárquico multimodal de estilo de aprendizaje, utilizando un sistema clasificador difuso multicapa (MFCS). En (Camargo & Aguilar, 2014) se presenta un MFCS que consiste en una serie de sistemas difusos distribuidos jerárquicamente, que tienen la ventaja de que el número total de reglas de la base de conocimiento es más pequeño, y más simple, que un sistema difuso convencional. La salida de un sistema clasificador difuso (FCS) es la entrada al próximo FCS.

En la Figura 2, se presenta el modelo MFCS para el reconocimiento del estilo de aprendizaje, que está compuesto por dos FCS: a) un FCS para reconocer el estado emocional, b) un FCS para reconocer el estilo de aprendizaje. Las entradas son los descriptores definidos en la sección anterior para cada nivel, pero en este caso se definen como variables difusas. Con estas variables difusas, se describe el conjunto de reglas difusas de cada FCS_i. Los conjuntos difusos propuestos para cada variable difusa de salida de cada nivel son: emoción={felicidad, tristeza, miedo, enojo, disgusto y sorpresa}, estilo de aprendizaje={visual, auditivo, de lectura/escritura, cinestésico}

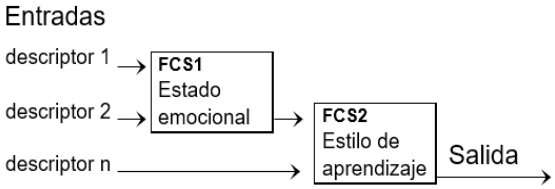


Figura 2 – Modelo MFCS para reconocer el estilo de aprendizaje

Por ejemplo, para el caso del FCS1, algunas de las posibles reglas difusas son:

- *Si (ojos abiertos) y (boca sonrisa) y (voz_tono agudo) y (cabeza levantada) y texto (positivo), entonces (emoción es felicidad).*
- *Si (boca cerrada) y (voz_tono grave) y (voz_intensidad baja) y (cabeza baja) y texto (negativo), entonces (emoción es tristeza).*

Para el caso del FCS2, algunas de las posibles reglas difusas son:

- *Si (emoción del estudiante es positiva) y (actividades de aprendizajes es simulaciones) y (estrategias de aprendizaje es basado en ensayos) y (herramientas de aprendizaje es videos), entonces (estilo de aprendizaje es visual).*
- *Si (emoción del estudiante es positiva) y (actividades de aprendizajes es simulaciones) y (estrategias de aprendizaje es basado en afectos) y (herramientas de aprendizaje es laboratorios), entonces (estilo de aprendizaje es cinestésico).*

5. Comparación de enfoques

En esta sección, se realiza una comparación de enfoques considerando las capacidades de cada técnica. Así, para la estrategia de adaptación se analiza la capacidad de aprendizaje, y para la comunicación del estado emocional del estudiante se analiza la capacidad de comunicación (esta última es importante en los ambientes inteligentes y el Internet de las cosas).

5.1. Estrategias de adaptación (capacidad de aprendizaje).

Consiste en la capacidad de poder aprender rápidamente la personalidad y estilos de aprendizaje de un estudiante.

Crónicas: Una misma situación (una emoción, un estilo de aprendizaje, etc.) puede describirse por diferentes crónicas, para expresar la diversidad de contextos donde puede ocurrir una misma situación (por ejemplo, un estilo de aprendizaje visual). Sin embargo, el principal problema es aprender el conjunto de crónicas requeridas. En la literatura, hay dos tipos de procesos de aprendizaje en el paradigma de las crónicas (Aguilar, 2011) aprender la estructura de una crónica o parametrizar una crónica general. Este es un problema abierto, en el cual se han desarrollado recientes propuestas (Aguilar et. Al 2018), que permiten descubrir las crónicas que describen los estados a reconocer (emociones, etc.) desde los datos. En un sistema real, se pueden definir crónicas generales para cada descriptor del modelo propuesto, y luego parametrizar estas crónicas para cada estudiante. Este enfoque requiere una sólida base de datos de crónicas, que se actualice y aprenda constantemente para adaptarla al estudiante y a nuevas situaciones.

Lógica Difusa: Un FCS puede aprender las reglas y las estructuras de las variables difusas. Esto significa que las funciones de membresía de las variables difusas se pueden adaptar al contexto, y las reglas de la base de datos se pueden modificar (sus componentes antecedentes y consecuentes) (Camargo & Aguilar, 2014). Por ejemplo, cuando la definición difusa de las emociones no es adecuada, las funciones de membresía

se pueden modificar. De manera similar, para el caso de las reglas difusas, las reglas deben adaptarse para reflejar los patrones específicos de cada. Para lograr esto, el FCS permite la modificación de las reglas cuando se presenta información nueva.

La tabla 4 muestra las métricas de la capacidad de aprendizaje con estos enfoques, usando una base de datos artificial, con todos los descriptores de cada nivel del modelo. En ella, cada registro corresponde a un estilo de aprendizaje de un estudiante, donde la cedula del estudiante es el identificador del registro. Esa base de datos contiene 5.000 registros etiquetados con las correspondientes emociones y estilos de aprendizaje. En general, las crónicas proporcionan mejores resultados, son más precisas para reconocer los patrones, posibilitan procesos de aprendizaje en línea desde los datos con bajo error, sin hacer reconocimientos inesperados.

Enfoque	Exactitud	Memoria	Medida F	Error cuadrático
Crónicas	0.87	0.91	0.89	0.12
Lógica difusa	0.79	0.82	0.80	0.69

Tabla 4 – Resultados de la capacidad de aprendizaje

En particular, las crónicas pueden describir una misma situación (una emoción, un estilo de aprendizaje) utilizando diferentes crónicas, pero este enfoque requiere una sólida base de datos de crónicas, que aprende constantemente, para adaptarse al estudiante y a las nuevas situaciones.

5.2. Comunicación del estado emocional o estilo de aprendizaje del estudiante (capacidad de comunicación).

En este caso, queremos evaluar el escenario de un salón de clases donde el intercambio de información útil es natural y necesario entre dispositivos y el sistema. Esto es algo muy común en el Internet de las Cosas, y en general, en los ambientes inteligentes. En ese sentido, los enfoques de reconocimiento deben poder sintetizar la información del patrón reconocido (ya sea de una emoción o estilo de aprendizaje), para poder trasmitirlo rápidamente al dispositivo que lo requiera en el contexto.

Crónicas: la comunicación entre diferentes crónicas son eventos. Eso es válido para el caso en que las crónicas están en el mismo salón de clases o en ambientes diferentes. Estos eventos pueden incluir información específica requerida por las crónicas, pero es la única información requerida. El modelo jerárquico de patrones de emociones comunica las crónicas reconocidas (la información relevante). Por ejemplo, cuál es el estilo de aprendizaje de un estudiante. Eso es suficiente para generar las acciones relacionadas con el estilo de aprendizaje reconocido. También, ellas simplemente necesitan recibir los mensajes que contengan información de los eventos a reconocer. Por ejemplo, el sistema de reconocimiento de emociones, envía un evento para informar que la emoción que prevalece en el aula es "ira", el cual puede ser recibido por el profesor para que modifique le metodología de enseñanza y favorezca el aprendizaje, o cualquier dispositivo en el ambiente.

Lógica Difusa: En este caso, existen dos posibilidades: enviar un valor discreto, que debe ser *desfuzzificado* en el otro sistema difuso, o enviar los valores de las variables

difusas (pero en el otro lado, el sistema difuso debe ser similar). El principal problema es que podemos tener múltiples salidas (múltiples reglas activas, que pueden representar, por ejemplo, varias emociones) y deben enviarse al otro sistema difuso para tener una idea real del contexto.

Además de los aspectos específicos de cada técnica señalados anteriormente como necesidades de información, el tiempo que duran en inferir (tiempo para reconocer), y el tiempo que duran en transmitir la información que generan (cuando reconocen), son fundamentales en los ambientes inteligentes y en el Internet de las cosas.

La tabla 5 muestra las métricas de la capacidad de comunicación en un entorno simulado de transmisión de 1000 mensajes enviados con eventos, procesados por cada enfoque. Los tiempos de respuesta y transmisión son mejores para las crónicas, ya que solo envían los eventos involucrados en el proceso de reconocimiento, y con ellos, se realiza el proceso de inferencia. En el caso del FCS, se requiere un tiempo de preprocesamiento de las variables difusas intercambiadas, y el sistema de reglas tiene un proceso arduo de razonamiento causal difuso, lo que implica mayores tiempos de respuesta y comunicación.

Enfoques	**Tiempo de respuesta**	**Tiempo de transmisión**
Crónicas	0.21	0.16
Lógica difusa	0.76	0.57

Tabla 5 – Resultados de la capacidad de comunicación (normalizadas en una escala de 0 a 1)

6. Comparación con otros enfoques

En esta sección se presenta una comparación con otros trabajos, para la identificación de los estilos de aprendizaje y el reconocimiento de las emociones en el contexto educativo. Para eso, se consideran los siguientes criterios:

1. Usan las emociones para identificar el estilo de aprendizaje en los alumnos.
2. Son capaces de aprender el estilo de aprendizaje.
3. Usan técnicas de inteligencia artificial (redes neuronales, lógica difusa, arboles bayesianos, etc.) para reconocer los estilos de aprendizaje y emociones.
4. Plantean un modelo de aprendizaje emocional para ambientes inteligentes.

Trabajo	**Criterios**			
	1	**2**	**3**	**4**
Nuestra propuesta	x	x	x	x
(Sun, 2016)		x	x	
(Suliman, 2010)	x			
(Faria et al., 2017)	x		x	
(Zatarain-Cabada, Barrón-Estrada, & Zatarain-Cabada, 2013)	x	x	x	

Tabla 6 – Comparación de diferentes trabajos para identificar estilos de aprendizaje

En particular, (Sun, 2016) propone un sistema de aprendizaje personalizado, basado en los estilos de aprendizaje, donde se aplican los enfoques de agrupamiento y árbol de decisiones para dicha tarea. (Suliman, 2010) explora las habilidades o estilos de aprendizaje preferidos de los aprendices, así como las relaciones entre el éxito académico y los estilos de aprendizaje, y la inteligencia social emocional. (Faria et al., 2017) diseñaron un prototipo de software para reconocer las emociones de los alumnos, mediante técnicas de computación afectiva, y así realizar ajustes a sus procesos de aprendizaje individual. (Zatarain-Cabada et al., 2013) presentan un sistema difuso que reconoce los estilos de aprendizaje y las emociones. En general, a excepción del modelo propuesto en este trabajo, ninguno ha sido diseñado para ser usados en ambientes inteligentes, y en partículas en aulas inteligentes, que requieren tiempos de respuesta reducido (ver (Sánchez et al. 2015a; Sánchez et al. 2015b), para más detalles sobre los requerimientos de esos entornos). Además, aunque varios trabajos reportan procesos de aprendizaje (Sun, 2016; Faria et al.,2017), los mismos son fuera de línea.

7. Conclusiones

En esta investigación, se ha propuesto un patrón jerárquico para identificar el estilo de aprendizaje, que considera dos niveles, uno para reconocer el estado emocional, y, el otro para identificar el estilo de aprendizaje. El modelo del patrón jerárquico es flexible porque permite incorporar nuevos descriptores, por ejemplo, para reconocer emociones considerar aspectos psicológicos.

Los dos enfoques analizados para modelar los estilos de aprendizaje, uno basado en crónicas y otro basado en lógica difusa, contribuyen a identificar el estilo de aprendizaje en un salón de clases. En general, las métricas de calidad con las crónicas son mucho mejor, por ser un paradigma que permite fácilmente aprender, con un mecanismo de razonamiento temporal simple.

La relevancia de la incorporación de modelos como el que se propone en este trabajo, es que permite estudiar el impacto de poner en evidencia la identificación de estilos de aprendizaje asociado a las emociones, en particular, si favorece el aprendizaje del estudiante. Esto es útil tanto en ambientes inteligentes como en las aulas clásicas de clase, dado que el profesor o el entorno inteligente puede orientar la clase para lograr mejores resultados.

Como trabajo futuro, se pretende implementar el modelo en entornos reales de clases, y en un aula inteligente, para obtener métricas vinculadas a su efecto en el proceso de enseñanza-aprendizaje. De esta forma, realizaremos una evaluación cuantitativa del impacto del uso en línea de este modelo, lo que es complementaria a la comparación analizada en este artículo.

Referencias

Affectiva. (2016). How it Works. *Emotion SDK*. Retrieved from https://www.affectiva.com/product/emotion-sdk/

Aguilar, J. (2011). Temporal Logic from the Chronicles Paradigm : learning and reasoning problems , and its applications in Distributed Systems. *Ed. Lambert*.

Aguilar, J, Aguilar, K, Gutiérrez, J. (2018) "An Approach for the Structural Learning of Chronicles", Contemporary Engineering Sciences, 11 (17), 793-806.

Bettadapura, V. (2012). Face expression recognition and analysis: the state of the art. *ArXiv Preprint ArXiv:1203.6722*.

Camargo, E., & Aguilar, J. (2014). Hybrid intelligent supervision model of oil wells. In *IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)* (pp. 934–939). https://doi.org/10.1109/FUZZ-IEEE.2014.6891725

Coelho, V., Sousa, V., & Figueira, A.-P. (2014). The impact of a school-based social and emotional learning program on the self-concept of middle school students. *Revista de Psicodidáctica*, *19*(2).

Cordero, J., & Aguilar, J. (2016). Reconocimiento multimodal de emociones en un entorno inteligente basado en crónicas. *COISINT*, 525–541.

Dousson, C., & Le Maigat, P. (2007). Chronicle Recognition Improvement Using Temporal Focusing and Hierarchization. In *IJCAI* (Vol. 7, pp. 324–329).

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, *6*(3–4), 169–200.

Ekman, P., & Friesen, W. V. (2003). *Unmasking the face: A guide to recognizing emotions from facial clues*. Ishk.

Faria, A. R., Almeida, A., Martins, C., Gonçalves, R., Martins, J., & Branco, F. (2017). A global perspective on an emotional learning model proposal. *Telematics and Informatics*, *34*(6), 824–837.

Felder, R. M., Silverman, L. K., & others. (1988). Learning and teaching styles in engineering education. *Engineering Education*, *78*(7), 674–681.

Felder, R. M., & Spurlin, J. (2005). Applications, reliability and validity of the index of learning styles. *International Journal of Engineering Education*, *21*(1), 103–112.

Fleming, N., & Baume, D. (2006). Learning Styles Again: VARKing up the right tree! *Educational Developments*, *7*(4), 4.

Fuentes, C., Herskovic, V., Rodriguez, I., Gerea, C., Marques, M., & Rossel, P. O. (2017). A systematic literature review about technologies for self-reporting emotional information. *Journal of Ambient Intelligence and Humanized Computing*, *8*(4), 593–606.

Goralnik, L., Millenbah, K. F., Nelson, M. P., & Thorp, L. (2012). An environmental pedagogy of care: Emotion, relationships, and experience in higher education ethics learning. *Journal of Experiential Education*, *35*(3), 412–428.

Javanbakht, A., Duval, E. R., Cisneros, M. E., Taylor, S. F., Kessler, D., & Liberzon, I. (2017). Instructed fear learning, extinction, and recall: additive effects of cognitive information on emotional learning of fear. *Cognition and Emotion*, *31*(5), 980–987.

Kolb, A. Y., & Kolb, D. A. (2005). Learning styles and learning spaces: Enhancing experiential learning in higher education. *Academy of Management Learning & Education*, *4*(2), 193–212.

Kolb, D. A. (2014). *Experiential learning: Experience as the source of learning and development*. FT press.

Kumar, K., Bakhshi, A., & Rani, E. (2009). Linking the big five personality domains to organizational citizenship behavior. *International Journal of Psychological Studies*, *1*(2), 73.

Magdin, M., Turčáni, M., & Hudec, L. (2016). Evaluating the Emotional State of a User Using a Webcam. *International Journal of Interactive Multimedia & Artificial Intelligence*, *4*(1).

Mahasneh, A. M. (2013). Learning Styles as a Predictor of Emotional Intellegence Among Sample of Jordanian University Students. *European Journal of Business and Social Sciences*, *2*(2), 46–55.

Morgan, R. (2012). A Comparison of VARK Analysis Studies and Recommendations for Teacher Educators Ron Morgan and Fred Baker University of South Alabama.

Pekrun, R., & Linnenbrink-Garcia, L. (2012). Academic emotions and student engagement. In *Handbook of research on student engagement* (pp. 259–282). Springer.

Sánchez M., Aguilar J., Cordero, J., Valdiviezo, P (2015a) A Smart Learning Environment based on Cloud Learning. International Journal of Advanced Information Science and Technology, 39 (39), 39-52.

Sánchez M., Aguilar J., Cordero, J., Valdiviezo, P (2015b) Basic features of a Reflective Middleware for Intelligent Learning Environment in the Cloud (IECL), In Asia-Pacific Conference on Computer Aided System Engineering.

Santos, R. J. da S. (2010). Sistema de Apoio à Argumentação em Grupo em Ambientes Inteligentes e Ubíquos considerando Aspectos Emocionais e de Personalidade.

Shen, L., Wang, M., & Shen, R. (2009). Affective e-Learning: Using" Emotional" Data to Improve Learning in Pervasive Learning Environment. *Journal of Educational Technology & Society*, *12*(2).

Suliman, W. A. (2010). The relationship between learning styles, emotional social intelligence, and academic success of undergraduate nursing students. *Journal of Nursing Research*, *18*(2), 136–143.

Sun, H. (2016). Design of personalized learning system based on learning styles. *Revista Ibérica de Sistemas e Tecnologias de Informação*, (E11), 66.

Willingham, D. T., Hughes, E. M., & Dobolyi, D. G. (2015). The scientific status of learning styles theories. *Teaching of Psychology*, *42*(3), 266–271.

Zatarain-Cabada, R., Barrón-Estrada, M. L., & Zatarain-Cabada, R. (2013). Learning Styles and Emotion Recognition in a Fuzzy Expert System. *Research in Computing Science*, *65*, 87–94.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 17/10/2018
Aceitação/Acceptance: 04/11/2018

Sistema para la Detección y Seguimiento de Afecciones Cardíacas Soportado en SBC

Darly Mildred Delgado Agudelo[1], Diego Fernando Girón Timaná[2], Gabriel Elías Chanchí Golondrino[3], Katerine Márceles Villalba[4], Stivens Antonio Dionizio Solarte[5]

darlydelgado@unimayor.edu.co, dfgiron@unimayor.edu.co, gchanchi@unimayor.edu.co, kmarceles@unimayor.edu.co, sdionizio@unimayor.edu.co

[1, 2, 3, 4, 5] Facultad de Ingeniería, Institución Universitaria Colegio Mayor del Cauca, Carrera 7 N° 2-34 Edificio Bicentenario, 190003, Popayán- Cauca, Colombia.

Pages: 717–728

Resumen: Según la Organización Mundial de la Salud, las enfermedades cardiovasculares son la principal causa de muerte en el mundo. De ahí que se hace fundamental la detección y el tratamiento temprano de este tipo de enfermedades. Una de las problemáticas asociadas a estos padecimientos es la no detección a tiempo, debido en parte al elevado costo de contar en centros de salud con estos dispositivos, que permiten realizar un seguimiento oportuno. Ante la necesidad de posibilitar un mejor acceso a las tecnologías, en los últimos años los SBC (Single Board Computer) han surgido como alternativa a la implementación de múltiples servicios en diferentes contextos, teniendo como ventajas el costo económico, portabilidad, interoperabilidad, etc. En este artículo se propone el desarrollo de un sistema soportado en SBC y sensores cardíacos para la detección de padecimientos cardiacos, como la obtención y clasificación del índice de estrés mental de una persona.

Palabras-clave: Afecciones cardiacas; detección; SBC; seguimiento; sistema.

System for the Detection and Monitoring of Heart Disease Supported in SBC

Abstract: According to the World Health Organization, cardiovascular diseases are the leading cause of death in the world. Thus, the detection and early treatment of this type of disease is fundamental. One of the problems associated with these conditions is the failure to detect on time, due in part to the high cost of having these devices in health centers, which allow timely monitoring. Given the need to enable better access to technologies, in recent years the SBC (Single Board Computer) has emerged as an alternative to the implementation of multiple services in different contexts, taking advantage of the economic cost, portability, interoperability, etc. In this article propose the development of a system supported in SBC and cardiac sensors for the detection of cardiac conditions, such as the obtaining and classification of the mental stress index of a person.

Keywords*:* Detection; heart disease; monitoring; SBC; system.

1. Introducción

Según la Organización Mundial de la Salud (OMS), las enfermedades cardiovasculares son la principal causa de muerte en el mundo. Del mismo modo, se estima que en 2012 murieron por esta causa 17,5 millones de personas, lo cual representa un 31% de todas las muertes registradas en el mundo. Ejemplos de este tipo de padecimientos cardiovasculares son la bradicardia y la taquicardia (Pardo, et al., 2011). La bradicardia es entendida como un ritmo sinusal con una frecuencia inferior a los 60 latidos por minuto. Por su parte la taquicardia es definida como un ritmo sinusal con una frecuencia superior a los 100 latidos por minuto, el cual surge como respuesta a un factor de estrés físico, emocional, patológico o farmacológico y que tiene como característica no ser no paróxica (Consejo de Salubridad General, s.f.). Por otra parte, el estrés mental puede ser considerado como un factor principal que ejerce influencia sobre el desarrollo de la hipertensión arterial, siendo el estrés entendido como una amenaza a la integridad psicológica de un individuo que da lugar a respuestas fisiológicas y/o de comportamiento (Solarte, et al., 2018).

En ese mismo orden de ideas, se hace necesario considerar que para las personas que tienen un alto riesgo cardiovascular, es fundamental la detección y el tratamiento temprano de este tipo de enfermedades. Una de las dificultades tecnológicas asociadas a este tipo de padecimientos es la no detección a tiempo de las mismas, debido en parte al elevado costo contar en centros de salud con estos dispositivos, los cuales permiten realizar un seguimiento y control oportuno a este tipo de afecciones. En este sentido, ante la necesidad de mejorar el acceso a las tecnologías y propiciar el desarrollo inclusivo de soluciones, en los últimos años han surgido con fuerza los ordenadores de placa reducida o también conocidos como SBC (Single Board Computer), los cuales constituyen una posible alternativa a la construcción de servicios con costos más adecuados en diferentes contextos de aplicación.

A pesar de que los SBC tienen un tamaño reducido, estos dispositivos pueden llegar a tener las características de un ordenador de sobremesa, lo que significa que es posible ejecutar un sistema operativo consumiendo una cantidad de energía considerablemente menor. Esto permite la posibilidad de crear sistemas automáticos, independientes y distribuidos, prescindiendo así de los computadores tradicionales que vienen siendo utilizados en este tipo de sistemas. Así, este tipo de dispositivos tienen como principales ventajas, el costo económico, el estar basados por lo general en software y hardware libre, la portabilidad, la interoperabilidad con diferentes tipos de sensores y su adecuada relación entre costo y capacidad de procesamiento (Pérez, 2015).

En este trabajo se presenta como aporte el diseño y construcción de un sistema para la detección y seguimiento de afecciones cardiacas, a partir de las tecnologías de software libre más adecuadas asociadas a dispositivos tipo SBC. Así, de manera específica, en este artículo se propone el desarrollo de un sistema soportado en SBC y sensores cardíacos para la detección de padecimientos como la taquicardia, y la bradicardia, además de la obtención y clasificación del índice de estrés mental de una persona. El sistema construido y presentado en este trabajo pretende servir de apoyo en centros de salud y campañas médicas rurales, de tal modo que sea posible incrementar el número de

dispositivos para la detección y seguimiento de padecimientos cardiovasculares, dada la problemática del difícil acceso a los dispositivos convencionales.

El resto del artículo está organizado de la siguiente manera: en la sección 2 se presentan un conjunto de conceptos que se tomaron en consideración para el desarrollo del presente trabajo; en la sección 3 algunos trabajos relacionados relevantes en el desarrollo de esta investigación, en la sección 4 se describe el diseño del sistema de detección y seguimiento de afecciones cardiacas; en la sección 5 se presentan los resultados obtenidos mediante la construcción del sistema propuesto; seguidamente en la sección 6 se presenta la evaluación del sistema, y finalmente en la sección 7 se dan las conclusiones y trabajos futuros derivados de la presente investigación.

2. Marco Conceptual

A continuación, se presentan un conjunto de conceptos que se tuvieron en cuenta para el desarrollo del presente proyecto. Dentro de estos se encuentran: bradicardia, taquicardia, estrés mental.

2.1. Bradicardìa

La bradicardia es entendida como un ritmo sinusal con una frecuencia cardiaca inferior a los 60 latidos por minuto. Se considera que la bradicardia puede traer complicaciones mayores si el corazón no bombea suficiente sangre con oxígeno al cuerpo. Generalmente el uso de marcapasos es una de las soluciones médicas a esta anomalía cardiaca (MayoClinic, 2017).

2.2. Taquicardìa

La taquicardia al igual que la bradicardia, es un tipo frecuente de trastorno o padecimiento del ritmo cardíaco, en este caso, el corazón late más rápido de lo normal cuando se encuentra en reposo. El corazón late más acelerado de lo conveniente para una persona, en las cavidades superiores o en las cavidades inferiores del corazón, o en ambas, al encontrarse en estado de reposo. Cuando se tiene este tipo de anomalía, los latidos del corazón sufren una variación superior que la establecida para un individuo sano, del orden mayor a los cien latidos por minuto (MayoClinic, 2017).

2.3. Estrés Mental

El estrés mental puede ser interpretado como una amenaza a la integridad psicológica de un individuo que da lugar a respuestas fisiológicas y/o de comportamiento. Según (Armario, et al., 2012) el estrés mental puede ser considerado uno de los principales factores que ejerce influencia sobre el desarrollo de la hipertensión arterial. Para el análisis del estrés mental en (Bayevsky, et al., 2012) se hace uso de la ecuación de Bayevsky, la cual permite la estimación del llamado índice de estrés mental (SI) (ver figura 1). La ecuación de Bayevsky surge del trabajo desarrollado en la unión soviética en cuanto al análisis de la variabilidad de la frecuencia cardiaca (VFC) para la medicina espacial. Esta fórmula se basa en métodos geométricos haciendo uso de la curva de

distribución cardio-histograma dibujada a partir del estudio de las variaciones del pulso o ritmo cardíaco. Esta identifica características como: Mo (moda), Amo (Amplitud de la moda), M*DMn (rango de variación o varianza (Solarte, et al., 2018)).

$$SI = \frac{AMo}{2Mo * (M * DMn)}$$

Figura 1 –Ecuación del índice de estrés mental

Dónde Mo es la moda o el nivel presumible en el que se encuentra trabajando el sistema cardiovascular y hace referencia al valor de RR que presenta el usuario con mayor regularidad en el conjunto de medidas analizado. Amo es la amplitud de la moda, en términos fisiológicos, es el índice nominal de actividad de la cadena de regulación simpática y se refiere al porcentaje de los intervalos que corresponden con el valor de la moda en la muestra tomada. M*DMn es el rango de varianza o diferencia entre los valores máximos y mínimos de los cardio-intervalos, es mejor conocido como varianza. SI es el índice de estrés, o índice del sistema de regulación de tensión, el cual representa el nivel de tensión de los sistemas de regulación (nivel de prevalencia de la actividad de regulación central por encima de los mecanismos autonómicos). En la Tabla 1, se pueden observar los rangos de índice de estrés que el usuario puede presentar, los cuales están asociados a tres estados de ánimo, tales como: estresado, relajado y normal.

Rango del SI	**Estado asociado**
>150	Estresado
>=40 y <=150	Normal
<40	Relajado

Tabla 1 – Rangos SI asociados a cada estado

3. Trabajos Relacionados

En este espacio se presentan los trabajos relacionados con la investigación, los cuales se encuentran enmarcados en la temática del uso de sensores con relación a la medicina.

En el trabajo presentado por (Íñiguez de la Torre, 2013) se muestra un proyecto investigativo que se basa en un sistema para la medida del pulso cardíaco usando diodos LED, cuyo principio físico hace referencia a la iluminación de la huella de un dedo por medio de la radiación de un LED emisor de infrarrojo. El cambio en el volumen sanguíneo con los latidos del corazón se ve reflejado por tanto en una señal pulsada a la salida del foto– detector. Al igual que el proyecto que se plantea en este artículo, en (Íñiguez de la Torre, 2013) se hace uso de instrumentación electrónica para la medición de pulsos cardiacos, pero la diferencia entre ambos está dada en que el primero a su vez proporciona un posible diagnóstico de alteraciones cardiacas y mide el índice de estrés mental de un usuario. Por el contrario, el segundo, solo da como información la medida del pulso y muestra la señal emitida por este a traves de un osciloscopio.

Por otro lado, en (Santos, 2010), se presenta una investigación cuyo propósito es la implementación de una red inalámbrica de sensores biomédicos que permitan realizar el monitoreo de parámetros fisiológicos a pacientes de una forma no invasiva usando el estándar de radio 802.15.4. Dentro de los parámetros monitoreados se encuentran: la presión sanguínea, el pulso cardiaco, la temperatura, entre otros. En este proyecto también se hace uso de software y hardware libre, sin embargo, solo se hace un monitoreo de las condiciones fisiológicas del paciente, más no se hace el análisis y procesamiento de los datos de cara a la generación de un diagnóstico.

Y finalmente, en (Atienza, 2008) que es producto de una tesis doctoral se plantea una investigación integrada dedicada al estudio de los mecanismos involucrados en la generación de las arritmias cardiacas. Para llevar cabo esto se hace uso de la combinación de herramientas de modelado y técnicas de procesado de señal para establecer una correspondencia directa entre la dinámica del sustrato cardiaco, las señales eléctricas registradas y la caracterización clínica mediante índices cardiacos extraídos a partir del procesado de las señales eléctricas cardiacas. En esta investigación no se hace uso de dispositivos electrónicos para el cumplimiento de su objeto, sino que se estudian las señales digitales en las tradicionales formas de medición de aspectos cardiacos como los electrocardiogramas y a esto se le relaciona la metodología de estudio implementada por el autor.

Como se puede observar, en los trabajos anteriormente mencionados hay un gran avance en las investigaciones de ingeniería dedicadas a la medicina, haciendo uso de componentes electrónicos como sensores. Sin embargo, en gran parte estos se encargan del monitoreo de factores fisiológicos de los usuarios o la implementación de metodologías para su perfeccionamiento encaminados en el mismo. Por lo tanto, el factor distinguible de dichas investigaciones con la que se plasma en este artículo es la detección de un posible diagnóstico.

4. Diseño del Sistema

En esta sección se presenta la estructura funcional del sistema propuesto en este artículo, el cual se ilustra en la figura 2. El sistema se encarga fundamentalmente de realizar un seguimiento sobre la frecuencia cardiaca y la variabilidad de la frecuencia cardiaca con el fin de evaluar si un paciente tiene padecimientos de bradicardia o taquicardia. Así mismo, se muestra como el sistema a partir del cálculo del índice de estrés mental permite establecer tres posibles estados: normal, relajado y estresado. Para realizar estas acciones, el sistema inicialmente, por medio de un sensor, realiza la captura de la señal de pulso, el cual permite obtener los datos del ritmo cardíaco, del usuario o individuo, luego el dispositivo realiza la graficación de la señal mostrada en la pantalla LCD dispuesta para esta tarea, de los datos obtenidos de la frecuencia cardiaca, a continuación, se hace un cálculo de las variables BPM e IBI, luego se establecen las condiciones para los que la persona ha sido evaluada como el nivel de estrés y las características de su condición cardiaca. Finalmente, después de calculadas y analizadas estas variables, se muestra las relaciones de estos datos para permitir establecer la anomalía o no, para el paciente evaluado.

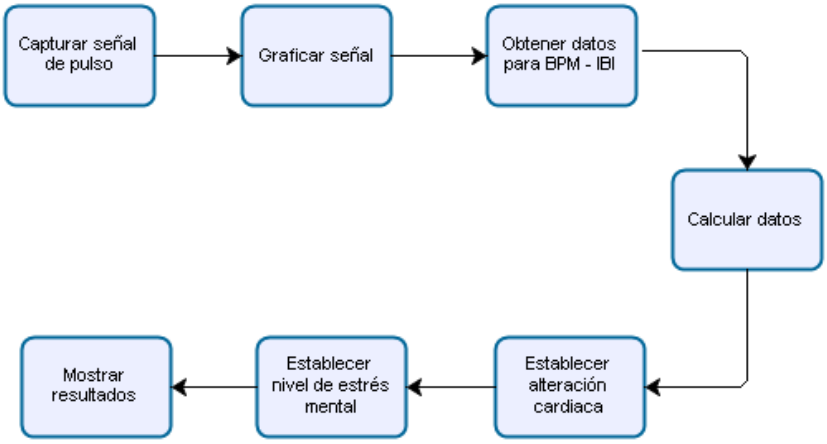


Figura 2 – Esquema funcional del sistema

5. Resultados

A continuación, en este apartado se exponen los resultados obtenidos a partir del desarrollo del "SISTEMA PARA LA DETECCIÓN Y SEGUIMIENTO DE AFECCIONES CARDIACAS SOPORTADO EN SBC", así como también se hace una presentación de los módulos de acoplamiento empleados para la construcción del hardware y software del dispositivo portable, y finalmente la explicación y verificación del funcionamiento (Ver figura 3).

Figura 3 –Tecnologías empleadas

En cuanto a las tecnologías usadas para la construcción del sistema se hizo uso de un dispositivo SBC denominado Arduino y otros módulos electrónicos que se acoplan adecuadamente al mismo. La razón de elección de este tipo de hardware está relacionada con su tamaño reducido que le da portabilidad al sistema además de ser una tecnología de bajo costo que le aporta la característica de fácil acceso al usuario que lo adquiera. Así, dentro de los componentes del sistema son: una placa Arduino Mega 2560 encargada de la ejecución del sistema, un sensor de pulso (pulseSensor) cuya función es recibir y emitir esa señal de pulso, un módulo de pantalla LCD TFT de 2'4 pulgadas que tiene como característica la emisión a traves de su pantalla de la información de entrada y salida , una batería externa recargable que le suministra energía al sistema, y finalmente un led como elemento opcional de indicador de señales de pulso (Ver figura 3).

Ya obtenidos los sub-sistemas eléctricos, electrónicos y mecánicos se procede a la programación, acople y verificación del funcionamiento del prototipo en ejecución por medio de la interfaz visual del dispositivo que en este caso es nuestra pantalla LCD la que permite la interacción máquina-humano, al ensamblar este conjunto de piezas ya con su correspondiente código de programación, se obtiene el dispositivo que se muestra en la figura 4 y 5 donde se hace evidente el tamaño y la potabilidad del dispositivo. En este mismo sentido, ya realizando la respectiva codificación del prototipo además del debido acople haciendo uso de todos los módulos mencionados anteriormente, se obtiene el montaje y la interfaz de interacción basada en la pantalla LCD (Ver figuras 4 y 5).

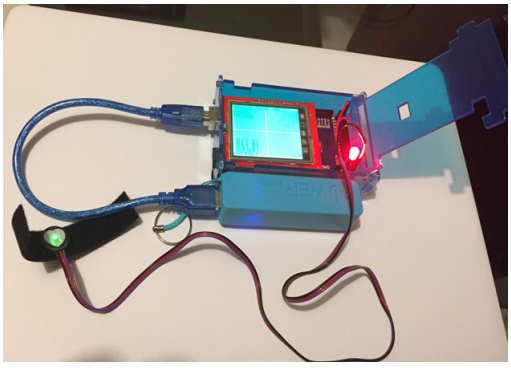


Figura 4 – Montaje final del prototipo

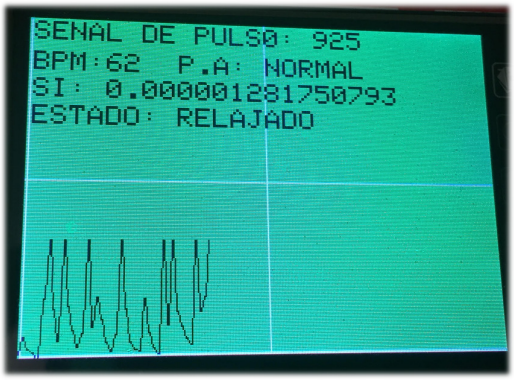


Figura 5– Interfaz del monitor

Seguidamente, en cuanto al contexto de funcionamiento del prototipo desarrollado, es posible observar en la figura 4 el montaje final del mismo, en el cual se ve la conexión del módulo de pantalla (desde la cual se visualiza toda la información) con la tarjeta Arduino seleccionada (Arduino Mega); y de este último con el sensor de pulso (sensor de frecuencia cardiaca) que es el encargado de recibir una señal análoga que luego es procesada y codificada a una señal digital para la obtención de la frecuencia cardiaca (FC) , una posible alteración cardiaca, la variabilidad de la frecuencia cardiaca (VFC) y el cálculo del índice de estrés (SI). De igual modo, se puede apreciar la conexión de la tarjeta Arduino con la batería externa, la cual proporciona la carga y le da el atributo de portabilidad al prototipo. Finalmente se encuentra también un led rojo encendido que

funciona como un indicador de los pulsos altos que pueda recibir el sensor a partir de la señal de entrada.

El sistema funciona de la siguiente forma; el sensor de pulso cardiaco óptico consta de un foto sensor y un led de alta intensidad que debe ser fijado en la punta del dedo o en el lóbulo de la oreja de un usuario (Ver figura 6), el led emite la luz sobre la piel de estas áreas donde gran parte de energía de esta luz es absorbida y otra muy mínima por lo capilares del tejido subcutáneo en donde la cantidad de sangre varía en función del ritmo cardiaco, esto genera una variación energética de la luz reflejada sobre la piel haciendo que el foto sensor lo note dando paso a amplificar y modular mediante el software hasta obtener el ritmo cardiaco, es decir, lee la luz que es reflejada por la piel donde se obtiene la señal análoga de entrada conocida también como señal de pulso, la cual es registrada cada 2 ms y se evidencia en el módulo de pantalla, al mismo tiempo que se va graficando la forma de onda (Ver figura 7). A su vez, la misma señal es transformada para obtener datos como la frecuencia cardiaca (FC) y la variabilidad de la frecuencia cardiaca (VFC), el primero designa la frecuencia cardiaca por minuto y el segundo la variabilidad de esa frecuencia. Al obtener ambos valores por un minuto es posible tener resultados de alteraciones cardiacas como bradicardia, taquicardia o indicar el estado normal. Del mismo modo en pantalla se aprecia el valor del índice de estrés mental y su posible clasificación en los tres posibles estados: relajación, estrés y nivel normal (Ver figura 5).

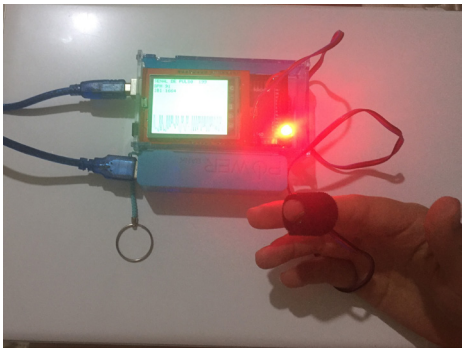


Figura 6–Funcionamiento del prototipo

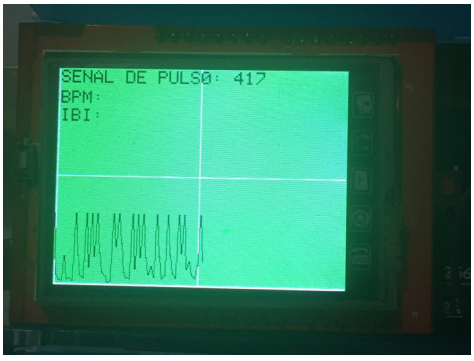


Figura 7 – Visualización de señal de pulso

En la figura 8, se muestran algunas de las pruebas realizadas del sistema construido con usuarios de la Institución Universitaria Colegio Mayor del Cauca. Estas pruebas fueron realizadas con el fin de verificar el funcionamiento adecuado del prototipo desarrollado.

Figura 8 – Pruebas realizadas

Las pruebas fueron realizadas a usuarios con diferentes rangos de edad, y condiciones físicas, es decir, fueron aplicadas a usuarios que practican algún deporte como también a otros con condiciones de sedentarismo, en este caso se obtuvo que la gran mayoría de los resultados de frecuencia cardiaca (FC) o BPM (número de bits/pulsaciones por minuto) se presentaron sin alteraciones cardiacas, en otras palabras una condición normal, sin embargo los usuarios que no practican deporte su FC fue mayor a 85 y aunque se encuentra entre los rangos de normalidad son elevados, y esto se debe a que los mismos al no practicar algún tipo de actividad física sus pulsaciones serán mayores al de una persona que sí lo haga. También se evidencio que en casi todos no hubo registro de algún índice de estrés, a excepción de un usuario que manifestó tener acumulación laboral, y su SI estuvo por encima de 150.

6. Evaluación del Sistema

Para la evaluación del *Sistema para la detección y seguimiento de afecciones cardiacas soportado en SBC*, se le aplicó una prueba de rendimiento, con el propósito de determinar la velocidad que se toma el sistema en procesar una petición, como se muestra en la figura 9, donde se presentan los resultados de las pruebas de tiempo de respuesta o de procesamiento realizadas al sistema construido.

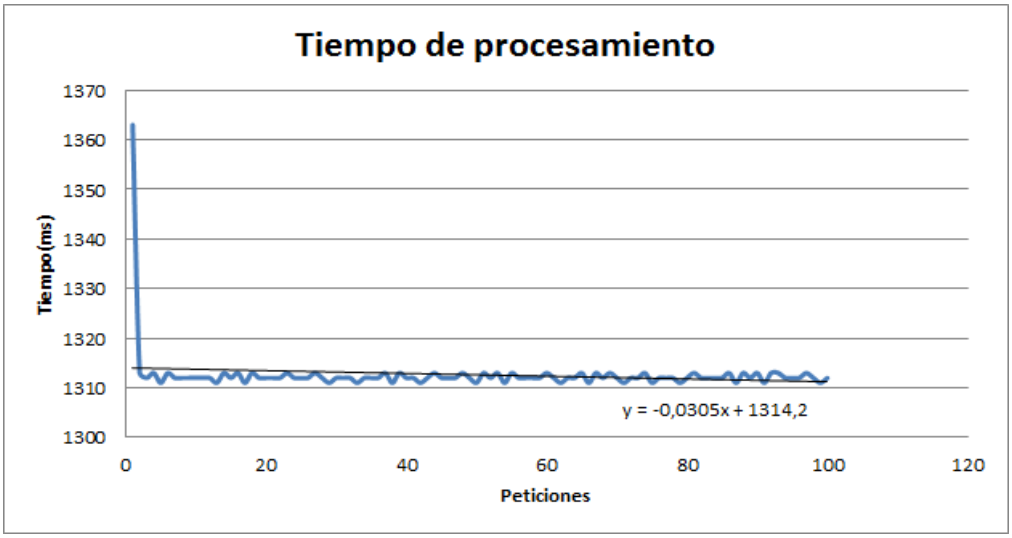


Figura 9 – Pruebas de rendimiento

Para realizar estas pruebas se contó con 100 peticiones vs el tiempo que emplea el sistema en determinar la afección cardiaca (bradicardia, taquicardia, normal) y en calcular el índice de estrés mental; como se muestra en la figura esto denota que la tarjeta Arduino Mega emplea un tiempo máximo cercano a los 1360 milisegundos para la primera petición, el cual se estabiliza para el resto de las 99 repeticiones con un tiempo promedio cercano a los 1314 milisegundos según lo indica la línea de tendencia. Lo anterior determina que el sistema emplea menos de segundo y medio para realizar las estimaciones estadísticas y visualizar el resultado en pantalla, lo que demuestra una rapidez y eficiencia en el funcionamiento del dispositivo. De ese modo se puede concluir que las tecnologías escogidas para la construcción del sistema satisfacen los requisitos funcionales que inicialmente fueron planteados.

7. Conclusiones y Trabajos Futuros

En esta sección se presentan las conclusiones y trabajos futuros derivados de la investigación, entre ellas se destacan:

El sistema construido, permite integrar las herramientas del hardware y software libre más adecuadas para realizar el seguimiento y detección de afecciones cardiacas sobre entornos SBC. En este orden de ideas las herramientas seleccionadas (sensores, LCD, etc.) son compatibles con la tarjeta de hardware libre Arduino. De acuerdo con las pruebas realizadas sobre el sistema se logró demostrar que estas herramientas son las más adecuadas para cumplir con los requisitos funcionales planteados.

El sistema para la detección y seguimiento de afecciones cardiacas soportado en SBC es una herramienta construida con dispositivos como sensores, módulo de pantalla LCD,

tarjeta Arduino, batería, entre otros, los cuales fueron acoplados de manera adecuada encaminados a enfocar su funcionalidad a realizar seguimiento y detección de afecciones cardiacas (taquicardia, la bradicardia y clasificación del índice de estrés mental de una persona).

Otro aspecto a resaltar del sistema construido es la portabilidad del dispositivo y así mismo su adecuado costo económico en lo que concierne a su adquisición, éste es uno de los aspectos más sobresalientes, dado que se tiene conocimiento que por ser un dispositivo muy costoso algunos centros de salud no disponen de los medios para adquirir herramientas como ésta, por ende es difícil que se encuentren en los hogares, para aquellas personas que sufren de trastornos cardiovasculares y por tanto requieren ser monitoreadas, pero con la llegada de este sistema se pretende servir de apoyo a estas entidades para tratar de disminuir el riesgo mediante la constante monitorización, permitiéndole al usuario tener una alerta temprana.

Como trabajo futuro derivado de la presente investigación, se pretende integrar dentro del sistema construido un conjunto adicional de variables para el seguimiento del comportamiento fisiológico de un usuario. Así mismo, como aporte

a futuro se considera importante llevar este producto de nuevo conocimiento y de desarrollo tecnológico e innovación a ser parte de una Spin- noff, de esta manera lograr impactar a esos sectores que requieran hacer uso de esta herramienta.

Referencias

Armario, P.; Hernández del Rey, R.; Almendros Rivas, M.C. (2012). Estrés mental como factor de desarrollo de hipertensión arterial. Revista Hipertensión y Riesgo Vascular. Vol. 19. No. 4.

Atienza, F. (Mayo de 2008). Estudio de los mecanismos de las arritmias cardiacas mediante modelado y procesado robusto digital de señal.

Bayevsky, R.M.; Ivanov, G. G.; Chireykin, L. V. (2002). HRV analysis under the usage of different electrocardiography systems (Methodical recommendations). Committee of New Medical Techniques of Ministry of Health of Russia. Moscow.

Blank, W. (1997). Promising practices for connecting high school to the real world. Tampa, Florida, USA: University of South Florida.

Consejo de Salubridad General, Tratamiento de la taquicardia supraventricular.S.f. Guía de práctica clínica del Gobierno Federal de los Estados Unidos Mexicanos.

Gros, B. (2009). Certezas e interrogantes acerca del uso de los videojuegos para el aprendizaje. Comunicación, 1(7).

Institución Universitaria Colegio Mayor del Cauca. (2008). Proyecto Educativo Institucional P.E.I. Popayán.

Íñiguez de la Torre, I. (Junio de 2013). Medida del pulso cardiaco basado en diodos emisores de luz: acondicionamiento de señales electronicas en sensores. Salamanca, España.

Macario, M. (2003). Documentales de divulgación científica on-line. Obtenido de http://www.upf.edu/pdi/dcom/xavierberenguer/cursos/interact/treballs/MARIA/texto.pdf.

MayoClinic (Agosto de 2017) Bradicardia síntomas y causas. Obtenido de https://www.mayoclinic.org/es-es/diseases-conditions/bradycardia/symptoms-causes/syc-20355474.

MayoClinic (Enero de 2017) Taquicardia síntomas y causas. Obtenido de https://www.mayoclinic.org/es-es/diseases-conditions/tachycardia/symptoms-causes/syc-20355127.

Pardo Fresno, Mónica; Fernández Barbeira, Saleta; González Bermúdez, Inmaculada. (2017). Valoración y manejo de las Bradicardias en urgencias de Atención Primaria. ABCDE en Urgencias Extrahospitalarias. Volume 18. Pág. 107-110.

Pérez, D. (2015). Registrador de datos autónomo con Arduino. Trabajo de fin de carrera en Ingeniería Informática de la Universidad de Cantabria.

Santos, M. (2010). Módulo inhalambrico para redes de sensores biomédicos. *Revista Científica de Ingeniería Electrónica, Automática y Comunicaciones, 31*(2), 16–22.

Sepe. (Junio de 2012). Competencias para la inserciòn laboral - Guía para el profesorado. Obtenido de http://www.sepe.es/LegislativaWeb/verFichero.do?fichero=09017edb800f8507.

Solarte, L.A.; Sánchez, M.; Chanchí, G. (2018). Sistema hardware-software para el reconocimiento de emociones a partir de señales fisiológicas. Memorias de las IV Jornadas Iberoamericanas de Interacción Humano Computador.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 24/10/2018
Aceitação/Acceptance: 19/12/2018

Desenvolvimento de Aplicações Acessíveis no Contexto de Sala de Aula da Disciplina de Interação Humano-Computador

Valéria Farinazzo Martins[1,2], Cibelle Albuquerque de La Higuera Amato[2], Glaucia Roxo Ribeiro[2], Maria Amélia Eliseo[1]

Valeria.farinazzo@mackenzie.br, cibelle.amato@mackenzie.br, caribeiro1@hotmail.com, mariaamelia.eliseo@mackenzie.br

[1] Faculdade de Computação e Informática - Universidade Presbiteriana Mackenzie, Rua da Consolação, 930, CEP 01302-000, São Paulo – SP, Brasil.

[2] Programa de Pós-Graduação em Distúrbios do Desenvolvimento - Universidade Presbiteriana Mackenzie, Rua da Consolação, 930, CEP 01302-000, São Paulo – SP, Brasil.

Pages: 729–741

Resumo: Estima-se que no mundo existam milhões de crianças com deficiências em idade escolar. Em vários países, as leis de inclusão preveem que estes alunos estejam matriculados em escolas regulares. Porém, para garantir um aprendizado universal, são necessárias certas ações, como a elaboração de material acessível a todos os estudantes. Este trabalho apresenta um estudo de caso com os resultados da experiência referente à prática de projeto na área de acessibilidade, desenvolvido no contexto da sala de aula, dentro da disciplina de Interação Humano-Computador (IHC) em cursos de graduação na área de Computação. Durante as aulas de IHC, os estudantes desenvolveram um material de apoio ao aprendizado, voltado para crianças do ensino básico, baseado no conceito de *Design* Universal para a Aprendizagem, contemplando as habilidades de cada estudante, independente de suas deficiências cognitivas ou físicas. A experiência mostrou o engajamento e envolvimento dos estudantes, não apenas por se tratar de um projeto real a ser implementado, mas também, pelo desafio na busca de soluções na elaboração de material acessível que atenda às diferentes necessidades do público-alvo.

Palavras-chave: Interação Humano-Computador; Processo de Ensino-Aprendizagem; Design Universal para a Aprendizagem; Sistemas Educacionais Acessíveis.

Accessible Applications Development in the Classroom Context of the Human-Computer Interaction Course

Abstract: It is estimated that there are millions of children with school-age disabilities. In several countries, basic laws provide for these students to be enrolled in regular schools. However, for universal learning, action is needed, including the development of a material accessible to all students. This paper presents the development results of a learning support material aimed to include children with

disabilities. This work presents a case study with the results of the project practice experience in the accessibility area, developed in the context of the classroom, within the Human-Computer Interaction (HCI) undergraduate course in the Computing area. During the HCI course, students developed a learning support material for children in elementary education, based on the Universal Design for Learning concepts, which addresses each student's abilities, regardless of their cognitive or physical disabilities. Experience has shown student engagement and involvement, not only because it is a real project to be implemented, but also because of the challenge in finding solutions in the production of accessible material that meets the different needs of the target audience.

Keywords*:* Human-Computer Interaction; Teaching-Learning Process; Universal Design for Learning; Accessible Educational Systems.

1. Introdução

De acordo com a Organização Mundial da Saúde (World Health Organization, 2015), no mundo, 15% da população mundial enfrenta alguma deficiência. Estima-se que quase 100 milhões de crianças - ou uma em cada 20 com menos de 15 anos de idade - apresentem uma incapacidade moderada ou grave. Uma informação relevante é que a deficiência é maior em países de baixa renda do que nos países de renda mais alta.

Essas pessoas enfrentam barreiras importantes no acesso a serviços, tais como saúde, educação, emprego e serviços sociais, incluindo moradia e transporte. Muitas questões têm sido responsáveis pelo cenário mundial de falta de inclusão dessas pessoas com deficiência: legislação, políticas e estratégias inadequadas; a falta de prestação de serviços; atitudes negativas e discriminação; falta de acessibilidade; financiamento inadequado; entre outros. Essas barreiras aumentam as desvantagens experimentadas por essas pessoas. Especialmente nos países em desenvolvimento, as pessoas com deficiência são acometidas por taxas mais altas de pobreza, taxas mais baixas de realização educacional e emprego, entre outros.

No âmbito da inclusão escolar, a Lei Brasileira de Inclusão da Pessoa com Deficiência (BRASIL, 2015) obriga as escolas a acomodar os alunos com deficiência no ensino regular e a adotar as medidas de adaptação necessárias, sem que nenhum ônus financeiro seja repassado às taxas.

De acordo com os Direitos das Pessoas com Deficiência, aprovado pelas Nações Unidas, em 2006, foi estabelecido que os Estados Partes deveriam assegurar um sistema de educação inclusivo em todos os níveis de ensino, em ambientes que maximizem o desenvolvimento acadêmico e social compatível, a participação plena e a inclusão. As Políticas Nacionais de Educação Especial na perspectiva da Educação Inclusiva (BRASIL, 2017) atendem o público-alvo de alunos de Educação Especial com deficiência, transtornos globais do desenvolvimento e altas habilidades / superdotação.

Assim, a acessibilidade e a inclusão digital devem ser implementadas por meio do aprimoramento de habilidades e da valorização das habilidades de cada um e pelo acesso aos recursos computacionais fornecidos àqueles que possuem pouca ou nenhuma possibilidade de uso fora do espaço escolar. Torna-se, portanto, fundamental medir a importância dos recursos computacionais no mundo contemporâneo, em termos de acessibilidade e da inclusão digital.

Esses recursos estão presentes no cotidiano das pessoas, inclusive nos espaços educacionais e sociais. No entanto, o maior desafio atual da educação no Brasil - e em muitos outros países - não é apenas incluir pessoas com deficiência em salas de aula regulares, mas sim, fazer com que estes estudantes tenham acesso à aprendizagem (hoje elas são incluídas na escola e excluídas do conhecimento); ou seja, o material didático (livros) não foi projetado para esta inclusão. Esta é a razão pela qual há uma necessidade para o desenvolvimento de ferramentas (digitais) e metodologias que considerem as diferentes necessidades de cada deficiência, mas que é aplicável dentro de um contexto comum de sala de aula.

Desta maneira, resolver esse problema envolve tanto profissionais de Informática quanto de Educação. Cabe aos profissionais da informática conhecer as estratégias para oferecer soluções flexíveis e amplamente utilizáveis, promovendo a autonomia das pessoas com deficiência em suas atividades diárias. Para os profissionais da Educação, é importante conhecer essas soluções, a fim de promover as habilidades desses alunos com deficiência (Melo e Silva, 2013).

Neste contexto, investigou-se a utilização do Design Universal para a Aprendizagem (DUA) para mitigar estes problemas supracitados. A estrutura para a DUA é baseada em descobertas da neurociência cognitiva que fornece informações sobre as necessidades de aprendizes individuais. Incorpora pedagogia acessível em três considerações específicas e centrais no ensino: os meios de representar informação, os meios para expressão de conhecimento dos estudantes e os meios de compromisso em aprender (Meyer et al., 2014).

O objetivo central deste projeto está intrinsecamente ligado ao processo de ensino-aprendizagem da disciplina de Interação Humano-Computador (IHC), no sentido de, não apenas, utilizar projetos reais, mas também o de projetar sistemas acessíveis, para um público-alvo com diferentes necessidades. Assim, DUA foi utilizado como base, pelos estudantes da disciplina de IHC, para o desenvolvimento de conteúdos educacionais digitais interativos e acessíveis. Vale ressaltar que trata-se de uma disciplina de cursos de graduação na área de Computação, cujo foco está nos fatores humanos em sistemas digitais interativos, com ênfase nas necessidades dos usuários e análise de tarefas dos usuários em relação ao sistema.

Este artigo está estruturado como segue. Na seção 2 estão as definições dos conceitos de Design Universal para a Aprendizagem, das deficiências cobertas pelo projeto e, também, dos trabalhos relacionados. Na seção 3 está a contextualização do projeto, apresentando a metodologia utilizada na disciplina para o desenvolvimento do projeto. Na seção 4 são abordados os resultados e discussões da aplicação da metodologia. Finalmente, na seção 5 são encontradas as conclusões do trabalho.

2. Referencial Teórico

2.1. Design Universal para a Aprendizagem

Os meios habituais de ensino, não são capazes de lidar com todos os tipos de alunos, com ou sem deficiência. O Desenho Universal para a Aprendizagem (DUA) é uma proposta que visa garantir o acesso ao conteúdo a todos os alunos. O objetivo da educação passou

da aquisição do conhecimento para a experiência do aprendiz. O lema principal do DUA é: o que é essencial para alguns, é bom para todos (Meyer et al., 2014). Ele não se baseia nos déficits, mas nas características gerais dos alunos (Cruz e Nascimento, 2018). Sendo assim, cabe à escola providenciar os meios e recursos para que cada aluno tenha acesso ao currículo escolar e seja um especialista em aprender.

DUA se baseia no Desenho Universal, conceito que veio da arquitetura e foi criado em 1970 por Mace (1998), que postulou a criação de produtos e ambientes para serem utilizados por todas as pessoas, sem que houvesse a necessidade de adaptações. A ideia de acessibilidade da arquitetura e design, foi utilizada na área educacional, garantindo aos alunos com e sem deficiência o acesso à aprendizagem. A base conceitual do DUA é a neurociência. Para aprender três redes neuronais são convocadas: a) afetivas: o porquê da aprendizagem; b) reconhecimento: o quê da aprendizagem; c) engajamento: o como da aprendizagem.

Para que os alunos tenham acesso ao conhecimento, é necessário que existam algumas mudanças em quatro aspectos do currículo: 1) Objetivos: listar conhecimentos e habilidades necessários para os alunos dominarem; 2) Avaliação: monitorar a evolução do aluno e propor alterações nas propostas sempre que necessário; 3) Métodos: deve estar de acordo com cada rede: afetiva, reconhecimento e engajamento. Para a rede afetiva, oferecer diversos contextos de aprendizagem; para a rede de reconhecimento, oferecer diversos recursos de acesso à aprendizagem e para a rede de engajamento manter o aluno motivado e proativo na tarefa. O uso da informática é um dos possíveis recursos a serem utilizados no auxílio à rede de reconhecimento. 4) Materiais: estar de acordo com as metas de aprendizagem.

Além disso, o uso da tecnologia é fundamental não só para garantir o acesso ao conteúdo, mas por tornar o aluno mais independente e autônomo nas tarefas acadêmicas (Alnahdi, 2014). A tecnologia também diminui as barreiras metodológicas, oportunizando o mesmo currículo para todos os alunos, mas com objetivos, métodos, avaliações e materiais personalizados. Consequentemente, ao diminuir as barreiras metodológicas e de aprendizagem, o DUA permite que alunos desenvolvam os processos mentais superiores e metacognição, pois favorece a interação entre instrumentos (objetos externos) e signos (objetos internos) fazendo com que sejam capazes de aprender a aprender, que é um dos grandes objetivos da educação (Plestch, 2009).

2.2. Deficiências e os Sistemas Web

De acordo com Nielsen (2000), entre as deficiências, a maior barreira de acesso aos sistemas web se dá para os deficientes visuais, pois a maioria destas páginas web é altamente visual. Entre os deficientes visuais, estão os cegos, as pessoas com baixa visão e os daltônicos. Pessoas deste grupo têm dificuldades em identificar imagens, mapas, links e animações, com o uso de cores, tamanho de fontes, entre outros. Elas necessitam de textos que descrevam uma imagem representada na tela, e na maioria das vezes utilizam o teclado ao invés do *mouse* (Macedo, 2010).

Já as pessoas com deficiência auditiva necessitam de apresentação visual para as informações, tais como legendas para textos sonoros. Os usuários surdos têm dificuldade no acesso de *audioclips* ou avisos sonoros. Em sala de aula presencial,

geralmente precisam de intérpretes de língua de sinais e nos cursos online usam seus próprios recursos de assistência. De acordo com Nielsen (2000), o uso de legendas em vídeos e transcrições de apresentação em áudio tornam o conteúdo mais acessível aos mecanismos de busca e facilitam a tradução. No entanto, segundo Slatin e Rush (Slatin e Rush, 2003), poucos indivíduos nascidos surdos são completamente alfabetizados na língua comum falada na sua comunidade; geralmente eles têm sua própria cultura e linguagem (de sinais). Para acesso digital, necessitam de versão de sinais, além de Captions e legendas (Macedo, 2010).

Finalmente para as pessoas com deficiência intelectual e ou dificuldades de aprendizagem, considera-se aqui o impacto que estas condições têm no contexto escolar da sala de aula e do aprendizado (Carneiro e Costa, 2017). Para este grupo podem ser usados recursos visual e sonoro simultâneos, conversão de texto em voz, leitura acompanhada pelo cursor, e principalmente a simplicidade do texto.

2.3. Trabalhos Correlatos

Foram encontrados, na literatura, alguns trabalhos relacionados ao tema utilizado nas aulas de IHC. No que diz respeito ao uso do DUA em sala de aula, Hansen et al. (2016) desenvolveram e refinaram um ambiente de programação para as salas de aula do ensino fundamental (alunos de 9 a 12 anos). Foi desenvolvida e testada uma progressão de aprendizagem para o pensamento computacional no nível do ensino fundamental utilizando uma versão modificada do Scratch para programar uma história digital num ambiente de programação baseado em blocos. Os autores discutem as acomodações e modificações como ferramenta para diferenciar o currículo e instrução em sala de aula para melhor atender às necessidades de diversos alunos que estudam computação na escola primária.

Já Gamboa et al. (2017) apresentam diretrizes para a implementação do DUA na criação de conteúdos educacionais flexíveis no contexto do *t-learning* (aprendizado baseado em Televisão Digital Interativa - iDTV) para atingir o objetivo da rede de reconhecimento. A criação e desenvolvimento de conteúdos educativos para a iDTV deve considerar três aspectos: técnico, pedagógico e customização. Os autores concluem que é essencial fornecer múltiplas opções de representação, facilitando a concepção de conteúdos personalizados de acordo com as preferências e interesses dos estudantes. Assim, as vantagens oferecidas pelo t-*learning* na utilização de conteúdos audiovisuais podem ser aplicadas neste contexto com sucesso.

Goldowsky e Coyne (2016) utilizam os princípios do DUA para o desenvolvimento de um sistema online, utilizado no contexto acadêmico, dotado de um ambiente de discussão para pessoas com dificuldade de leitura. Este ambiente oferece múltiplas opções para realizar cada *post*: escrever um comentário, desenhar uma imagem ou gravar um áudio comentário. O objetivo do projeto é capacitar os leitores com dificuldades a se envolver e se sentirem bem-sucedidos ao ler e compreender textos. Os resultados mostraram que alunos em todos os níveis de habilidade de leitura e escrita puderam participar das discussões. A capacidade de gravar áudio foi particularmente benéfica para os alunos com deficiências intelectuais, que raramente são capazes de se envolver com seus colegas para discutir conteúdos sérios e relevantes para a idade. Além disso, o sistema tem potencial de oferecer não só mais opções para alunos com deficiência, para que possam

se envolver mais completamente em discussões com seus colegas, mas também, para permitir a todos os alunos um conjunto flexível de opções dentro do qual eles possam mostrar seus pontos fortes, expressar sua individualidade e se divertir.

Quanto a trabalhos que discutem o desenvolvimento de aplicações práticas no processo de ensino-aprendizagem da disciplina de IHC, destacam-se Marques e Melo (2015), Falcão (2014), Dantas et al. (2013), Pittarello e Pelegrini (2015). Todos estes trabalhos relatam experiências com alunos da graduação em cursos de Computação que aprenderam na prática a importância dos conceitos estudados em sala de aula.

Marques e Melo (2015) apresentam uma abordagem interdisciplinar entre três disciplinas: Modelagem e Projeto de Banco de Dados, Projeto de Sistemas e Projeto de Interface Homem-Computador. Os estudantes desenvolveram um projeto multidisciplinar que integra visões sobre os dados, funcionalidades e interface de um sistema, possibilitando a vivência de uma prática próxima do real.

O desenvolvimento de um aplicativo destinado a grávidas para auxiliar na prática de exercícios físicos durante as aulas de IHC foi a abordagem utilizada por Dantas et al. (2013). Após a realização das atividades propostas, os estudantes aplicaram técnicas estudadas em sala de aula para melhorar a usabilidade do aplicativo. Da mesma forma, Pittarello e Pelegrini (2015) focaram no desenvolvimento de interfaces de *eco-feedback* para melhorar a consciência ambiental e otimizar o consumo de energia em contextos domésticos. Os estudantes criaram projetos conceituais que foram além das soluções disponíveis comercialmente. A experiência foi uma oportunidade de reflexão sobre os métodos e objetivos de um curso de IHC para estudantes de graduação.

Para motivar os alunos no aprendizado de IHC, Falcão (2014) relata uma experiência prática do ciclo de *design* centrado no usuário como forma de promover a construção de conhecimentos pelos estudantes. A proposta foi o desenvolvimento de um novo aparelho de celular, onde o desafio pedagógico, além de trabalhar o conteúdo teórico da disciplina, foi estimular a criatividade e inovação durante o processo de geração de soluções tecnológicas e fomentar a iniciativa nos alunos.

Os recursos computacionais possuem um grande potencial para o desenvolvimento de materiais didáticos baseados em DUA. No entanto, pouco tem sido realizado nesta área. Os trabalhos baseados em DUA realizados pelos estudantes da disciplina de IHC, além de serem uma forma de motivar o aprendizado, contribuem para a percepção da importância da elaboração de materiais didáticos inclusivos, à medida que os conteúdos desenvolvidos atendam às necessidades de todos os usuários, independentemente de suas limitações. Embora o uso de trabalhos práticos no processo de ensino-aprendizagem de IHC não seja novidade, a abordagem utilizada levou os alunos a colocarem em prática os conceitos vistos em aula e a se conscientizarem da importância da acessibilidade.

3. Metodologia do Trabalho

A experiência relatada neste trabalho (desenvolvimento do material acessível) se deu em quatro turmas da disciplina IHC de cursos da área de Computação de uma universidade particular localizada na cidade de São Paulo, no primeiro semestre de 2018. No total, estavam matriculados, nas disciplinas, 98 alunos.

A disciplina de IHC é semestral e composta por aulas teóricas com 2h/aula semanais e tem um caráter teórico-prático. Todas as aulas são ministradas em salas de aulas com notebooks disponibilizados, um por aluno, o que proporciona o desenvolvimento de exercícios práticos, tornando as aulas mais interativas e dinâmicas. Para o projeto de sistemas, é uma parte essencial do processo de design. A partir dessa perspectiva, a IHC envolve o projeto, a implementação e a avaliação de sistemas interativos no contexto da tarefa e do trabalho do usuário. O conteúdo da disciplina envolve questões ligadas aos usuários finais e aos sistemas computacionais, por isso leva em conta aspectos culturais, sociais, organizacionais e cognitivos. A avaliação foi realizada com base na participação de cada estudante nas atividades e na qualidade dos materiais produzidos.

Para o desenvolvimento, foi utilizada a abordagem de DUA, no entanto, devido a limitações de tempo (três meses para finalizar o projeto) e de pouca experiência dos estudantes em desenvolvimento, foi realizado um recorte de modo que somente as crianças típicas e com deficiência auditiva, visual e intelectual foram consideradas potenciais usuários para este estudo. Por isso, foram desenvolvidos conteúdos educacionais digitais interativos a partir de um capítulo de um livro didático impresso.

O objetivo do material acessível é o de reforçar o aprendizado das crianças em sala de aula junto ao professor, permitindo a inclusão de todos os alunos da sala de aula, independentemente de suas limitações. O assunto tratado no material desenvolvido foi o capítulo “Sistema Monetário Brasileiro”, para crianças do Ensino Fundamental I.

As etapas do desenvolvimento do projeto foram:

1. Entendimento do problema, proposta de possíveis soluções e documentação. Neste ponto houve fornecimento de parte do material sobre o capítulo (Sistema Monetário Brasileiro) pelas professoras da disciplina em consonância com a pedagoga envolvida.
2. Desenvolvimento do protótipo de baixo e alto nível, ou seja, protótipo em papel ou utilizando uma ferramenta de mockup (baixo nível) e também página HTML (alto nível).
3. Desenvolvimento do sistema funcional integrado entre todas as partes. Os estudantes tiveram que decidir questões de arquitetura do sistema, linguagem e design das páginas.
4. Avaliações de usabilidade: os estudantes aplicaram técnicas de avaliação de usabilidade ao sistema funcional (técnicas de inspeção e teste de usabilidade), a fim de diminuir os erros e verificar a aceitabilidade com crianças.

Para o processo de desenvolvimento foi utilizado o ciclo de *design* centrado no usuário, passando por todas as etapas: análise das necessidades dos usuários, das tarefas e dos requisitos; busca de soluções para o problema; prototipação e as avaliações de usabilidade. Com base no DUA, os aplicativos desenvolvidos deveriam ser compostos com as seguintes características:

- Página de conteúdo com o material teórico sobre o Sistema Monetário Brasileiro de maneira a ser compreendido pelas crianças. Uma característica desta página é fornecer, como alternativa, o áudio correspondente ao conteúdo escrito, links para vídeos, animações, exercícios e jogos. Também disponibiliza controles de aumento e diminuição de fontes e contraste.

- Página de atividades para disponibilizar exercícios abordando o tema proposto com diferentes níveis de dificuldade.
- Jogos para abordar o tema proposto de forma lúdica, com diferentes níveis de dificuldade.

Diferentes opções de representação permite a personalização conforme as preferências e interesses dos estudantes. Por isso, uma necessidade deste projeto é contemplar as diferentes mídias na aplicação:

- Textual - para atender, principalmente, alunos com dislexia e deficiência visual, além daqueles que preferem aprender através da leitura.
- Sonoro e áudio-descrição das imagens - para atender, principalmente, alunos com deficiência visual, além daqueles que preferem aprender através de áudio.
- Vídeo com legenda – para atender, principalmente, alunos com dificuldade no aprendizado e deficiência auditiva, além daqueles que preferem aprender através de vídeos.
- Imagens/animação - para atender, principalmente, alunos com dificuldade no aprendizado e deficiência auditiva, além daqueles que preferem aprender através de imagens.
- Jogos - para atender, principalmente, alunos com dificuldade no aprendizado, além daqueles que preferem aprender através de forma lúdica.

4. Resultados e Discussões

Cada uma das quatro turmas apresentou um projeto diferente relacionado ao mesmo tema (Sistema Monetário Brasileiro). A seguir são apresentadas algumas partes destes projetos, separados por tipos de conteúdo.

4.1. Conteúdo didático

O conteúdo didático foi elaborado de maneira a ser inclusivo: voltado para crianças com ou sem deficiências. Para atender este público-alvo, tendo em vista as necessidades especiais de crianças com deficiência auditiva, visual e intelectual, o conteúdo, baseado no DUA, contém os seguintes elementos:

- Calculadora, régua, dicas e níveis de dificuldade diferentes para facilitar o uso e absorção de conhecimento de crianças com dificuldade de aprendizado ou deficiência intelectual.
- Funcionalidades para o leitor de tela, mudança de contraste, aumento e diminuição da fonte para atender a crianças com deficiência visual e, também, crianças com dificuldade de aprendizado e dislexia.
- Inserção de legenda nos vídeos para atender a crianças com deficiência auditiva, assim como de crianças com deficiência intelectual.
- Escrita em letra bastão porque as crianças foco do projeto ainda não sabem ler com letras em outro formato.
- Vídeos (com legenda) com explicação sobre o assunto para atender a crianças com deficiência auditiva, assim como com deficiência intelectual.
- Links para crianças que queiram saber mais informações além do que foi apresentado.

As Figuras 1 e 2 apresentam parte destas necessidades.

Figura 1 – Sistema contendo régua, calculadora, aumento e diminuição de fonte, letra bastão

4.2. Atividades de fixação

As atividades de fixação propostas atendem aos mesmos elementos indicados para o conteúdo didático, além da validação do exercício para que a criança saiba se acertou ou errou a atividade.

Como pode ser visto na Figura 2 também apresenta níveis de dificuldade diferentes, calculadora e áudio. A Figura 3 também apresenta outros destes elementos.

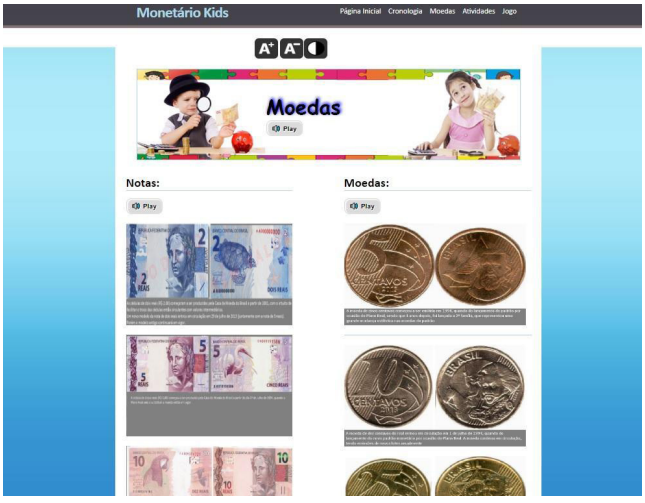


Figura 2 – Sistema contendo aumento e diminuição de fonte e leitor de tela

Figura 3 – Atividades com níveis diferentes, calculadora e letra bastão

4.3.Jogos

Os jogos atendem crianças que preferem aprender de forma lúdica, mas nem sempre conseguem atender às necessidades de uma criança com deficiência visual, conforme mostram as Figuras 4(a) e 4(b). Cada um destes projetos foi testado com duas crianças típicas para verificar a sua aceitação. No entanto, testes com crianças com deficiências deverão ainda ser realizados para verificar pontos de melhorias e correções.

(a)

(b)

Figura 4 - Jogo sobre compras num supermercado, trabalha conceito de troco, apresenta níveis (a). Jogo da memória com intuito de memorizar as cédulas e moedas (b).

4.4.Avaliação de Usabilidade

Além da produção do material didático, os estudantes realizaram um planejamento para aplicação de avaliações de usabilidade por meio de testes de usabilidade e técnicas de inspeção para verificar a eficácia, eficiência e satisfação ao utilizar o material didático. As avaliações buscaram analisar tarefas como: realizar e completar as atividades de fixação; utilizar a régua ou calculadora para realizar cálculos durante as atividades de fixação; trocar as cores da tela para alterar o contraste; alternar o tamanho da fonte; utilizar áudio enquanto realiza a leitura do conteúdo didático; concluir o jogo; verificar a aparência das telas; navegar no sistema.

Os testes de usabilidade foram realizados pelos próprios colegas da sala, devido ao tempo da disciplina, que realizaram as atividades solicitadas e em seguida, responderam o questionário elaborado pela equipe que desenvolveu o material didático. O mesmo ocorreu com a técnica de inspeção, em que, novamente, os próprios colegas da sala, desta vez em pares, analisaram as tarefas solicitadas baseado na Avaliação Heurística. No final, os resultados foram analisados pela equipe, onde puderam perceber o que poderia ser melhorado no material proposto. Desta forma, os estudantes puderam passar por todo o processo de desenvolvimento de um sistema interativo: análise das necessidades dos usuários, tarefas e requisitos; concepção da solução; prototipação e avaliação.

5. Conclusões

Este trabalho apresentou um relato de experiência do desenvolvimento de material didático digital e acessível para apoio ao aprendizado do Sistema Monetário Brasileiro voltado à inclusão de crianças com deficiências, baseado na abordagem de DUA. Este é um assunto tradicionalmente estudado no Ensino Fundamental I nas escolas brasileiras.

Resultados informais coletados pelas professoras da disciplina apontam que os estudantes se sentiram motivados em desenvolver um sistema partindo de um escopo real. Mais ainda, a motivação se deu ao aprenderem e aplicarem conceitos de acessibilidade, assimilando a importância destes conceitos na construção de materiais educativos, bem como outros sistemas interativos. Outro ponto a destacar foi a importância das avaliações de usabilidade, que mostraram aos estudantes, que mesmo com todo o cuidado na realização das análises dos requisitos que deveriam ser atendidos na elaboração do material didático, a análise dos testes de usabilidade e a inspeção indicou a necessidade de melhorias no sistema. Os resultados obtidos, apresentados na seção 3 e discutidos na seção 4, mostram que é possível criar conteúdos educativos digitais interativos acessíveis no contexto de salas de aulas de IHC em cursos de graduação em Computação.

Embora os recursos utilizados na construção dos materiais educativos para promover acessibilidade não sejam novos, seguir um guia a partir do DUA foi importante para entender as necessidades de adaptação de conteúdo didático. Este tema, dentro dos materiais didáticos utilizados nas escolas de Ensino Fundamental é ainda pouco explorado.

Repensar a forma de ensinar, a partir da compreensão do que significa aprender, requer um esforço conjunto de diversos profissionais e de saberes específicos, principalmente quando consideradas as especificidades de cada aluno e a diversidade de cada sala de aula. Talvez este seja o grande desafio a ser enfrentado por todos os profissionais

preocupados com a educação, transformar as escolas em espaços para todos. Ambientes inclusivos e facilitadores à aprendizagem de todos.

Como trabalhos futuros, pretende-se testar os projetos desenvolvidos com crianças em salas de aulas inclusivas.

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de nível superior - Brasil (CAPES) - Programa de Excelência - Proex 0653/ 2018.

Referências

Alnahdi, G. (2014). Assistive Technology in Special Education and the Universal Design for Learning. *The Turkish Online Journal of Educational Technology*, 13, 18-23.

BRASIL (2015). Lei Nº 13.146, DE 6 de julho de 2015. Institui a Lei Brasileira de Inclusão da Pessoa com Deficiência (Estatuto da Pessoa com Deficiência). Brasília, DF.

BRASIL (2017). Ministério da Educação. Secretaria de Educação Especial. Política Nacional de Educação Especial na Perspectiva da Educação Inclusiva. Brasília: MEC/SECADI.

Carneiro, R. U. C.; Costa, M. C. B.. Tecnologia e deficiência intelectual: práticas pedagógicas para inclusão digital. *Revista on line de Política e Gestão Educacional*, [S.l.], p. 706-719, oct. 2017. ISSN 1519-9029. doi:https://doi.org/10.22633/rpge.v21.n.esp1.out.2017.10449.

Cruz, M.M., Nascimento, F.F. (2018). Acessibilidade ao currículo através do uso do computador para estudantes com autismo. *Rev. Interinstitucional Artes de Educar*. Rio de Janeiro, Vol. 4, nº 1- pp 43-65.

Dantas, A., Ferreira, B., Costa, I., Gomes, L., Aoki, M., Conte, T., Rodrigues, R. (2013). Da Teoria à Prática: Aplicando técnicas de IHC em um Aplicativo Móvel voltado para Grávidas. In *Anais do IV Workshop sobre Ensino de IHC - XII Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais (WEIHC 2013)*. Manaus, AM, 31-36.

Falcão, T. P. (2014). Aprender Fazendo: o Protagonismo Estudantil como Prática Pedagógica de Design Centrado no Usuário. In *Anais do V Workshop sobre Ensino de IHC - XIII Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais (WEIHC 2014)*. Foz do Iguaçi, PR, 01-05.

Gamboa, A. X. R.; Builes, J. J.; Monsalve, H. P. (2017). Universal design for learning (UDL) applied to T-learning. *2017 Computing Conference*, London, pp. 1160-1164.

Goldowsky, B. N., Coyne, P. (2016). Supporting engagement and comprehension online through multiple means of expression. In *Proceedings of the 13th Web for All Conference* (W4A '16). ACM, New York, NY, USA, Article 39, 4 pages. DOI: https://doi.org/10.1145/2899475.2899488.

Hansen, A. K., Hansen, E. R., Dwyer, H. A., Harlow, D. B., Franklin, D. (2016). Differentiating for Diversity: Using Universal Design for Learning in Elementary Computer Science Education. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education* (SIGCSE '16). ACM, New York, NY, USA, 376-381.

Mace, Ronald L. (1998). Universal design in housing. *Assistive Technology* 10.1: 21-28.

Macedo, C. M. S. de (2010). *Diretrizes para criação de objetos de aprendizagem acessíveis*. Tese (Doutorado em Engenharia e Gestão do Conhecimento) - Universidade Federal de Santa Catarina, Florianópolis.

Marques, A. B. dos S. M., Melo, A. H. da S. (2015). Facilidade de Aprendizagem, eficiência, memorabilidade, prevenção de erros, satisfação. In *Anais do VI Workshop sobre Ensino de IHC - XIV Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais (WEIHC 2015)*. Salvador, BA, 13-16.

Melo A.M., da Silva J.G. (2013) Online Digital Libraries at Universities: An Inclusive Proposal. In: Stephanidis C., Antona M. (eds) Universal Access in Human-Computer Interaction. Applications and Services for Quality of Life. UAHCI 2013. *Lecture Notes in Computer Science,* vol 8011. Springer, Berlin, Heidelberg.

Meyer, A., Rose, D.H., Gordon, D. (2014). *Universal design for learning: Theory and Practice*. Wakefield, MA: CAST Professional Publishing.

Nielsen, J. (2000). *Designing Web Usability: The Practice of Simplicity*. USA: New Readers Publishing.

Pittarello F., Pellegrini, T. (2015). Designing and evaluating interfaces for domestic eco-feedback: a blended educational experience. In *Proceedings of the 11th Biannual Conference on Italian SIGCHI Chapter* (CHItaly 2015). ACM, New York, NY, USA, 18-25. DOI: http://dx.doi.org/10.1145/2808435.2808463.

Pletsch, M. D. (2009). *Repensando a inclusão escolar de pessoas com deficiência mental: diretrizes políticas, currículo e práticas pedagógicas*. Tese (Doutorado em Educação) - Universidade do Estado do Rio de Janeiro, Rio de Janeiro.

Slatin, J. M.; Rush, S. *Maximum Accessibility. Making your Web site More Usable for Everyone*. USA e Canadá: Pearson Education, inc., 2003.

World Health Organization (2015). *WHO global disability action plan 2014-2021: Better health for all people with disability*. World Health Organization.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 15/09/2018
Aceitação/Acceptance: 07/11/2018

Diseño de un Framework de Arquitectura Empresarial para Instituciones Públicas de Educación Superior

Inés Meriño Fuentes[1], Wilson Nieto Bernal[2], Silvia Moreno Trillos[3], Yamid Hernández Julio[4]

imerino@unimagdalena.edu.co, wnieto@uninorte.edu.co, scmoreno@uninorte.edu.co, yamidh@uninorte.edu.co

[1] Profesora Universidad del Magdalena-Santa Marta, Estudiante de doctorado Ingeniería de Sistemas y Computación- Universidad del Norte, Barranquilla, Colombia.

[2] Profesor e Investigador Universidad del Norte, Barraquilla, Colombia.

[3] Profesora Universidad Simón Bolívar-Estudiante de doctorado Ingeniería de Sistemas y Computación-Universidad del Norte, Barranquilla, Colombia.

[4] Profesor Universidad del Sinú-Montería, Estudiante de doctorado Ingeniería de Sistemas y Computación-Universidad del Norte, Barranquilla, Colombia.

Pages: 742–755

Resumen: Este trabajo propone un framework de Arquitectura Empresarial (AE) para Instituciones de Educación Superior (IES) públicas, basado en sus procesos académico-administrativos. Para ello, se realizó una revisión de conceptos relacionados con AE en general y de estudios previos que aplicaron la AE en educación o en universidades. Se identificaron las características de algunos frameworks de AE validados y construidos específicamente para universidades. El Framework de AE propuesto modela aspectos claves de una organización: estrategias, procesos, aplicaciones, datos, infraestructura tecnológica y canales de entrega. También propone un modelo de madurez basado en el desarsrollo y gestión de la AE dentro de la organización. El Framework de AE propuesto permite alinear la información académico-administrativa de las IES con su misión y objetivos, teniendo como insumo sus procesos administrativos, los recursos existentes; permitiendo que dicha información se pueda almacenar y recuperar para la toma de decisiones, soportada en una infraestructura de Tecnologías de Información (TI).

Palabras-clave: Arquitectura Empresarial; Modelo de Madurez; Procesos académico-administrativos; Instituciones de Educación Superior.

Design of an Enterprise Architecture Framework for Public Higher Education Institutions

Abstract: This work proposes a framework of Enterprise Architecture (EA) for Higher Education Institutions (HEI), based on their academic-administrative processes. In order to reach such a proposal, a review of concepts related to EA in

general and previous studies that implement the EA in education or universities was performed. We identified the features of some validated EA frameworks, and other frameworks built specifically for universities. The proposed AE framework that models the key aspects of an organization: strategies, processes, applications, data, technological infrastructure and delivery channels. It also proposes a maturity model based on the development and management of the EA within the organization. The proposed AE framework allows aligning the academic-administrative information of the HEI with its institutional mission and objectives, taking as input their administrative processes and existing resources, allowing such information to be stored and retrieved for the decision-making process, supported by an IT infrastructure.

Keywords: Enterprise Architecture, Maturity Model; academic-administrative processes; Institutions of Higher Education.

1. Introducción

Las empresas de hoy en día deben enfrentarse a un mercado cada vez más competitivo con el fin de obtener una ventaja sobre sus pares, debido a esto, resulta imprescindible para las organizaciones sacar el máximo provecho de las tecnologías de la información. Según (Riascos, *et al.*, 2016). "Las TIC confluyen en la interrelación de los aspectos de software, hardware y telecomunicaciones, facilitando la generación de estrategias direccionadas al progreso de la organización". Una de las claves para un mayor éxito se logra cuando las metas y las estrategias empresariales están correctamente alineadas con las soluciones informáticas de todas las partes que integran la organización. Con este fin ha surgido la Arquitectura Empresarial (AE), la cual busca la sincronía entre el mundo del negocio y la solución de TI de una organización. Las instituciones de Educación superior – IES no son ajenas a esta necesidad de obtener el mejor provecho de las tecnologías informáticas (TI). Este artículo muestra una propuesta de Framework de Arquitectura Empresarial (AE) para Instituciones Públicas de Educación Superior, en particular sus procesos académico-administrativos. El *framework* de AE propuesto se soporta en algunos existentes que ya han sido validados y evaluados.

El uso de computadores, el desarrollo y aplicación de software (CAI-Enseñanza Asistida por Computador y SIG (Sistemas de Gestión de Información) en las IES, se dio desde los años ochenta, como soporte a la docencia y la gestión. Son muchos los desafíos a los que las IES se enfrentan, buscando un modelo aplicable para la puesta en marcha de una AE acorde a las normas, estructuras organizativas, procesos y políticas que una organización lleva a cabo; sobre todo teniendo presente la dinámica de los procesos al interior de una Universidad. Las universidades y colegios han adoptado una serie de medidas para afrontar ese reto (Chen, Tang, & Li, 2016).

La revisión de la literatura indica que se han propuesto modelos para describir la Arquitectura de Universidades, con las siguientes denominaciones: Campus Digital, Smart Campus, etc. El rápido desarrollo de las tecnologías de la información ha dinamizado también los cambios en las arquitecturas de TI de la universidad. Los departamentos u oficinas de TI, deben estar a la vanguardia de lo que están ocurriendo y adaptarse rápidamente (Chen *et al.*, 2016).

2. Metodología de la Investigación

Para el diseño del Framework de Arquitectura Empresarial para Instituciones Educativas fueron desarrolladas varias fases: Fase 0: Revisión de Literatura; Fase 1: Framework propuesto de Arquitectura Empresarial (SISG); Fase 2: Descripción Detallada del Modelo; Fase 3: Modelo de Madurez para Framework propuesto SSCI; Fase 4: Conclusiones y recomendaciones

3. Fase 0: Revisión de Literatura

Se inició con una revisión sistemática de la literatura sobre el tema. Se utilizaron como palabras claves "Enterprise Architecture Universities" e "Enterprise Architecture Education" en las Bases de Datos IEEE Explorer y ScienceDirect. Estas búsquedas se hicieron a cohorte, el 31 de mayo de 2017. Para la búsqueda realizada con estos dos conjuntos de palabras clave, se tienen los resultados de las Bases de Datos por tipo de publicación mostrados en la Tabla 1.) Luego se seleccionaron artículos para extraer conceptos que contextualizaron al lector sobre la temática tratada. 2) De manera paralela se analizaron procesos misionales de instituciones de educación superior en Colombia, especialmente relacionados con la parte académico-administrativo

Palabras clave:	**"Enterprise Architecture Universities"**		**"Enterprise Architecture Education"**	
	IEEE Explorer	**ScienceDirect**	**IEEE Explorer**	**ScienceDirect**
Conference Publications	1816		174	
Journals & Magazines	168	22489	13	8242
Early Access Articles	7		1	
Books & eBooks	6	3876	0	2307
Reference Work		347		156
TOTAL	**1997**	**25972**	**188**	**10375**

Tabla 1 – Resultados de búsqueda bibliográfica Fuente: Elaborado por los Autores, a partir de los resultados de las consultas a las Bases de Datos

La Figura 1 muestra cómo la producción ha ido avanzando a una tasa promedio de 10% cada año. En el año 2017 se observa una disminución, teniendo presente que la consulta fue a corte de mayo del 2017, en comparación con el 2016 que incluye los 12 meses del año. Cabe anotar, que, al momento de la búsqueda, ya aparecieron algunas publicaciones del año 2018. Se seleccionaron 22 artículos de IEEE y dos (2) de ACM, que guardaban relación directa con el trabajo que se desarrolló. De igual forma se hizo un análisis de los procesos misionales principales que tiene una institución de Educación Superior: Academia, Investigación y Extensión.

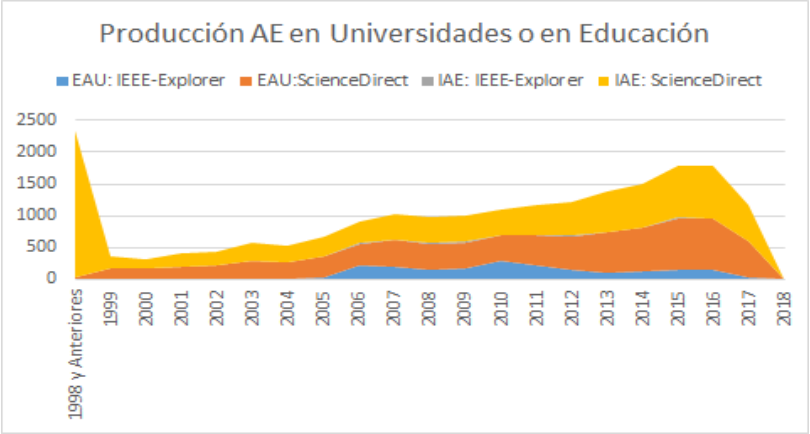


Figura 1 – Producción AE en Universidades y en Educación
Fuente: Elaborado por los Autores, a partir de los resultados de las consultas a las Bases de Datos

3.1. Arquitectura Empresarial (AE)

Varias son las definiciones comunmente aceptadas, iniciaremos por la propuesta por la Federal Enterprise Architecture Framework (FEAF), que define EA como "Un procedimiento administrativo para maximizar la contribución de los recursos de una agencia, las inversiones en TI y las actividades de desarrollo de sistemas para lograr sus objetivos de rendimiento. La arquitectura describe relaciones claras de metas y objetivos estratégicos a través de inversiones para mejorar el rendimiento de toda la empresa o un segmento de esta" (Gao, 2001)enterprise architectures are \u0192blueprints for systematically and completely defining an organizations current (baseline.

La definición AE, según Lapalme (2008) se asocia principalmente a tres escuelas de pensamientos: "(1) La plataforma de TI de la empresa, (2) La Integración empresarial y (3) La empresa y su entorno".

Es importante identificar el marco de arquitectura empresarial (EAF) que la empresa debe aplicar, teniendo presente las necesidades y prioridades de la organización y cómo crearlo (modelos, enfoques, normas), es decir la gobernanza de la AE. De allí la importancia de la AE para gestionar la complejidad del sistema, alinear los negocios y recursos de TI (Shah & Kourdi, 2007), es decir, para integrar el negocio y la infraestructura tecnológica debido a que comprime las funcionalidades y modelos de información, y su relación con la administración y las estructuras organizacionales de negocio(Cabrera, et al. 2015).

Otro tema de interés en el área de la AE es el desarrollo de metodologías que permitan evaluar la AE de una organización. Algunos autores como (Javanbakht, Pourkamali, & Feizi, 2009) buscan métodos automatizados de realizar esta valoración. En su artículo el autor propone un sistema multi-agente para determinar el estado de la arquitectura empresarial en una organización. El sistema propuesto comienza con un enfoque de arriba

hacia abajo para identificar los principales componentes de la arquitectura empresarial y luego utiliza un enfoque de abajo hacia arriba para evaluar la arquitectura de acuerdo con la significancia de cada componente para la misión y metas organizacionales.

Lakhrouit & Baina (2016)establecen un método para evaluar el logro de agilidad de diferentes escenarios TO-BE y utilizar diferentes componentes de complejidad de la arquitectura empresarial a la arquitectura del higher institute of applied engineering IGA. La AE empieza a ser una ciencia muy interesante para dar instrumentos, herramientas y modelos conceptuales, de cómo diseñar una organización para que pueda sobrevivir en un futuro y permita su evolución.

Andersen, Carugati, & Sorensen (2015)proponen un enfoque para evaluar la eficacia de la AE. Mientras que en la mayoría de los autores previos se utiliza un enfoque de arriba hacia abajo (top-down approach) en el que se adaptan medidas derivadas de la teoría de otros campos, como por ejemplo indicadores financieros, este propone un enfoque de abajo hacia arriba (bottom-up approach) en donde las medidas se derivan de estudios empíricos de la práctica de la AE. Andersen pone en práctica su enfoque analizando el caso de una Universidad pública grande de Dinamarca, en la cual las facultades venían operando como Silos, cada una con un sistema de información y unos objetivos estratégicos independientes, y se procedió a buscar la unificación de estos departamentos.

Petrikina, Drews, Schirmer, & Zimmermann) (2014)buscan integrar los modelos de negocios y la AE, con el objetivo de lograr un mayor entendimiento entre el mundo del negocio y el de TI. En su artículo ellos presentan un enfoque para integrar los modelos de negocio y la arquitectura empresarial en dos dimensiones: componentes y procesos administrativos, y para ambas se definen los artefactos iniciales.

3.2. Arquitectura Empresarial en las Instituciones de Educación Superior (IES): Universidades

En las dos últimas décadas, la AE se ha convertido en un enfoque para la gestión de sistemas de información en las organizaciones. Sin embargo, el desarrollo del sector educativo en este campo no ha sido recíproco a los acontecimientos, por lo tanto, se está optimizando esta situación relacionada con la AE. La metodología utilizada por (Ramadhan & Arman, 2014) en Indonesia comprende: primero recolección de datos como observaciones y entrevistas; segundo análisis de los resultados de la implementación de la Arquitectura Empresarial en la universidad y evaluación de la arquitectura del negocio, garantizando la implementación de la AE con el modelo de madurez; tercero aplicación de encuestas para identificar los problemas relacionados con el proceso empresarial sin resolver en la universidad y por último un estudio de caso que proporcione al investigador la capacidad de realizar una disertación en profundidad de un fenómeno en un contexto de la vida real.

Olsen & Trelsgard (2016) han elaborado un estudio de la AE en el sector de la educación superior en Noruega, han elegido soluciones muy diferentes a sus necesidades de TI, esto ha llevado a sistemas de TI rígidos e incompatibles. Estas instituciones luchan con los mismos retos de AE, y hay grandes beneficios potenciales de cooperar y estandarizar en una AE común. Encontraron que se trata de una tarea muy difícil, y el progreso se ha

visto obstaculizado por la falta de políticas del Ministerio de Educación e Investigación de Noruega. Esto conllevó a una falta de visión y dirección en los esfuerzos locales de AE. También es un problema significativo que el concepto de AE sea poco comprendido entre los altos directivos, y a menudo se deje al personal de TI.

Según Carrillo, Cabrera, Román, Abad, & Jaramillo (2010) un aspecto clave para definir un Framework de AE para instituciones de educación superior es comenzar por especificar claramente el estado actual de la institución (AE as-is) y cuál es la meta a la que se quiere llegar (AE to-be). Otros factores claves son contar con el soporte de la administración de la Universidad y definir una estructura de gestión y control de la AE. La adecuada asignación de recursos y la implementación de un plan fuerte lograrán que el proceso de desarrollo de la AE sea más iterativo y simplificado. También se debe tener en cuenta que el ciclo de vida de la arquitectura debe ser llevado a cabo en un límite de tiempo medio, acorde con el crecimiento y evolución de la institución. Finalmente es necesario adoptar un Framework de AE, ya sea seleccionando alguno de los más utilizados en la actualidad (Zachman, FEA, TOGAF, Gartner o E2AF, entre otros), o juntando características de estos para formar un Framework Híbrido que mejor se adapte a las necesidades de las instituciones de educación superior.

Para mejorar la Enseñanza y aprendizaje en la educación superior, según Llamosa-Villalba et al. (2015) proponen un prototipo de AE de Educación Superior de Colombia -CHE2A- que se justifica en un framework para la gestión del Ciclo de Vida de Sistemas (SoSL). CHE2A es un sistema de gestión del conocimiento cuyo objetivo es enriquecerse de las experiencias de las organizaciones y apoyar la innovación de la educación superior. Además, crea un aprendizaje dinámico, ágil y colaborativo integrado con el medio ambiente, situación que motiva y crea un valor para todos los involucrados en el proceso educativo, siguiendo las directrices del Ministerio de Educación de Colombia u otros factores similares establecidos por entidades nacionales o internacionales.

4. Fase 1: Framework propuesto de Arquitectura Empresarial (SISG)

"Una arquitectura empresarial describe los componentes necesarios en que opera la empresa para ofrecer servicios o fabricar productos para los clientes."(Burégio, 2015) La figura 2, representa el Framework de Arquitectura Empresarial propuesto. Aunque este marco puede ser implementado en cualquier organización, está especialmente pensado para aplicarse en Instituciones públicas de Educación Superior. Este Framework está inspirado en algunos modelos de arquitectura empresarial populares en la actualidad; principalmente TOGAF.

Figura 2 – Esquema Framework SISG de Arquitectura Empresarial
Fuente propia

5. Fase 2: Descipción del framework de Arquitectura Empresarial (SISG) propuesto

La Tabla 2 describe de manera general las 6 dimensiones base del Framework propuesto.

Dimensión	Descripción
D1. Visión estratégica institucional	Establecer el alcance, las limitaciones y las expectativas de un proyecto, creación de la visión de la arquitectura, identificación de los interesados y sus preocupaciones, validar el contexto del negocio y crear la declaración de arquitectura. Obtener aprobaciones} Modelado: Business Process Workflow (BPW).
D2. Arquitectura de procesos	Describir los procesos operativos que deben ser soportados por la plataforma existente. Describir los procesos de negocio que no son soportads por ninguna plataforma existente. En cada caso, desarrollar la arquitectura de la línea base-destino y analizar las brechas entre ambas. Modelado: Business Process Workflow (BPW).
D3. Arquitectura de aplicaciones	Describir los sistemas de información (SI) existentes involucrados en los procesos de negocio, SI necesarios para soportar los objetivos de negocio; interfaces de comunicaciones necesarias y demás aplicaciones. Modelado: Modelo de componentes. Definir los datos que soportan el proyecto. Analizando las necesidades de las partes interesadas con relación a los datos que requieren la institución. Modelado: Modelo Entidad-Relación.
D4. Arquitectura de información	Describir la infraestructura que soporta la operación de los sistemas existentes, sus bases de datos y demás servicios que agregan valor a los objetivos del negocio. Modelado: Modelo de despliegue.
D5. Arquitecura de infraestructura	Describir la infraestructura que soporta la operación de los sistemas existentes, sus bases de datos y demás servicios que agregan valor a los objetivos del negocio. Modelado: Modelo de despliegue.
D6. Arquitectura de canales de entrega	Definir las interfaces de comunicación con el usuario final y demás canales de entrega. Se confirma y supervisa el alcance y las propiedades de los proyectos de implementación. Modelado: Modelos de Wireframe de la UI.

Fuente Elaboración propia

Tabla 2 – Componentes Base del Framework SISG

Además de las seis (6) dimensiones existen cuatro (4) componentes transversales que se describen a continuación:

1. **Innovación Organizacional:** este proceso transversal se refiere a la búsqueda constante de oportunidades de mejora, tanto a nivel de procesos, aplicaciones, datos, infraestructura o canales de entrega. Es un proceso de mejora continua en el que se toma como insumo los conocimientos y el aprendizaje de los miembros de la organización para crear una cultura de innovación constante con el fin de lograr el liderazgo de la industria en la que se encuentre la organización.

2. **Gestión del riesgo:** este es un proceso realizado desde la administración de la organización y consiste en el establecimiento de estrategias de toda la empresa para identificar eventos potenciales que puedan afectar a la entidad y proponer cómo deben manejarse éstos para proporcionar seguridad e integridad razonable sobre el logro de los objetivos organizacionales. La gestión del riesgo incluye actividades como evaluación del riesgo, el desarrollo de estrategias para manejarlo y mitigación del riesgo utilizando recursos gerenciales. Según Cano (2015) "la gestión de riesgo busca: reconocer las actividades o prácticas que no se están ejecutando adecuadamente; sensibilizar a la organización de los impactos que se pueden tener del evento no deseado y establecer una forma sencilla y expedita de valorar y priorizar los esfuerzos". En resumen, en este componente se previene y detectan intrusiones, evitan fugas de información para asegurarla.
3. **Gestión y Gobierno de TI:** la Gestión de TI es el proceso de supervisión de todos los asuntos relacionados con los recursos de tecnologías de la información dentro de una organización, con el fin de asegurar que todos los recursos tecnológicos son utilizados correctamente de una manera que aporta valor a la organización. Por otra parte, el Gobierno de TI es el conjunto de acciones que realiza el área de TI en coordinación con la alta dirección para movilizar sus recursos de la forma más eficiente en respuesta a requerimientos regulatorios, operativos o del negocio. El Gobierno de TI garantiza que las TI están alineadas con la estrategia del negocio y según Vanti *et al.* (2018) se concentra en la administración y utilización de las TI para concretar metas de desempeño corporativo, así como también mejorar las estrategias empresariales y aumentar la competitividad; los servicios y funciones de TI se proporcionan de manera que aporten el mayor valor posible y que todos los riesgos relacionados con TI sean conocidos y administrados.
4. **Sistema de ciberseguridad corporativa (SCC):** este proceso transversal implica el diseño, implantación, y mantenimiento de un conjunto de procesos para gestionar eficientemente la accesibilidad de la información, con el objetivo de asegurar la confidencialidad, integridad y disponibilidad de ésta y minimizar los riesgos de seguridad de la información. Por lo general, la implementación de un SCC implica una fase de planificación, en la que se realiza la evaluación de riesgos de seguridad y se seleccionan los controles adecuados, una fase de hacer, que involucra la implantación y operación de dichos controles, una fase de control, que tiene por objetivo evaluar el desempeño del SCC y fase de actuar, en la que se realizan los cambios necesarios con el fin de garantizar el máximo rendimiento del SCC. Estas 4 fases se repiten constantemente generando un proceso de mejora continua. Valencia & Orozco (2017), indican que las normas para implementar un Sistema como este, corresponden a la serie ISO/IEC 27000 publicadas por la ISO y la Comisión Electrotécnica Internacional (IEC)"

5.1. Metodología ABC para desarrollar el Framework SSCI propuesto

Para la construcción de la AE para las IES se ha diseñado una metodología con 8 requerimientos (filas) asociados a 5 fases (columnas), tal como se muestra en la Tabla 3:

Requerimientos	Inicio	Ejecución	Cierre	Entrega	Mejoramiento continuo
Visión estratégica institucional	Direccionamiento Estratégico Corporativo	Portafolio del Proyecto	Entregables	Despliegue de productos de proyectos	
Arquitecturas	De aplicaciones	Modelado Aplicaciones	De procesos	de Infraestructura	De Canales de entrega
	Diagramas de casos de uso; Diagrama de clase; Diagrama de compontes; Diagrama; Entidad relación; Diagrama de despliegue; Modelos basados en red; Business Process; Worflow (BPW); Business Process Model (BPM)				
Innovación Organizacional	Metodología Stage gate				
Gestión de riesgo	ISO 27005 ISO 31022				
Gestión y Gobierno de TI	COBIT 5				
Sistema de Ciberseguridad Corporativa	ISO 27032				
Implementación	Entornos de Desarrollo de Soluciones. Basadas en modelos ágiles tipo DevOps				
Prueba	Pruebas automáticas; Pruebas de sistemas o aplicación; Pruebas de integración				

Fuente Elaborado por los autores

Tabla 3 – Metodología de la Arquitectura empresarial propuesta para las IES

La metodología se puede aplicar desde la perspectiva de filas o columnas y representa como se debe abordar la construcción del modelo de AE para IES. Las fases van desde la perspectiva conceptual o inicial hasta el mejoramiento continuo; los requerimientos van desde la visión estratégica institucional hasta las pruebas. En la intercepción de cada fila y columna se observa la integración de las fases y requerimientos y el producto o artefacto que lo identifica.

6. Fase 3: Modelo de Madurez para Framework propuesto SSCI

Un modelo de madurez es una técnica ampliamente utilizada para evaluar procesos de negocio o ciertos aspectos de una organización, y representa un camino hacia una manera de hacer negocios incrementalmente más organizada y sistemática (Vieira, Cardoso, & Becker, 2014).

Dimesiones	V1	V2	V3	V4	V5	V6
1. Visión estratégica institucional	Misión	Visión	Objetivos	Estructura organizacional	Políticas	Estrategías
2. Arquitecturas de aplicaciones	Existencia de un modelo de arquietctura de aplicaciones	Número de aplicaciones propias	Número de aplicaciones libres	Número de sitios web		
3. Arquitectura de información	Porcentaje de integridad	Porcentaje de disponibilidad	Porcentaje de capacidad	Porcentaje de elasticidad	Porcentaje de interoperabilidad	
4. Arquitectura de infraestructura	Inventario de productos y servicios	políticas de mantenimiento	porcentaje de utilización	Número de intranets	Número de extranets	
5. Arquitectura de procesos	Alineación de procesos a la vision institcional	Documentación de Procesos	Estandarización de Procesos	Existencia de procesos	Número de procesos misionales	Número de procesos de apoyo
6. Arquietctura de Canales de entrega	Número de canales de entrega	Número de Portales web	Número de redes sociales			
7. Innovación Organizacional	conocimiento de procesos de negocio,	Conocimiento servicio de TI o los productos creados	Documentación de productos creados			
8. Gestión de riesgo	Existencia de politicas de gestión de riesgos	Existencia de Medidas de Rendimiento				
9. Gestión y Gobierno de TI	Existencia de politicas para la gestión y gobierno de TI					
10. Sistema de Ciberseguridad Corporativa	Políticas de seguridad	Auditorias	Balanceo de cargas	Control de acceso		

Tabla 4 – Modelo de Madurez para el Framework propuesto

Los modelos de madurez por lo general poseen niveles de madurez, generalmente cinco: Inicial, Administrado, Definido, Administrado Cuantitativamente y Optimizado. Para proponer un modelo de madurez para Instituciones de Educación Superior nos basamos en el modelo EAMMF- Enterprise Architecture of Maturity Model Framework (J., K., Lakhrouit, & Baina, 2013) el cual toma como centro para evaluar la madurez que tan definido y administrado se encuentra el tema de la AE en la organización.

Por otra parte, el nivel de madurez en una organización se puede medir con respecto a los componentes clave de la arquitectura: estrategia, procesos, aplicaciones, datos,

infraestructura y canales de entrega, midiendo el nivel de desarrollo o alcance de la arquitectura empresarial para cada aspecto.

La medición del modelo se plantea de la siguiente manera:

$$\textit{Modelo de medición} = \sum_{d=1}^{\mathbb{0}} \sum_{j=1}^{n} Valor_{jd} . W_{jd}$$

d: controla el número de dimensiones
j: controla el número de variables
W: peso asociado a la variable

describe el modelo de madurez correspondiente al framework propuesto El cual se compone de un conjunto de dimensiones y cada dimensión tiene un conjunto de variables.

Una vez se tiene evaluado el modelo, la medición nos arroja una valoración como la que se muestra en la Tabla 5.

D(*) de AE propuesta	Nivel 1-Inicial	Nivel 2-Desarrollando	Nivel 3-Definido	Nivel 4-Administrado	Nivel 5-Optimizando
1. *Visión estratégica institucional*	Ad hoc o sin establecer	Visión estratégica Básica. Conocimiento básico.	Identificado y estándarizado	Gestionado y utilizado por todos	Mejoramiento continuo (métricas) u optimizado.
2. *Arquitecturas de aplicaciones*	Ad hoc o sin establecer	Aplicaciones Básicas. Conocimiento básico.	Identificado y estándarizado	Gestionado y utilizado por todos	Mejoramiento continuo (métricas) u optimizado.
3. *Arquitectura de información*	Ad hoc o sin establecer	Información Básica. Conocimiento básico.	Identificado y estándarizado	Gestionado y utilizado por todos	Mejoramiento continuo (métricas) u optimizado.
4. *Arquitectura de infraestructura*	Ad hoc o sin establecer	Infraestructura Básica. Conocimiento básico.	Identificado y estándarizado	Gestionado y utilizado por todos	Mejoramiento continuo (métricas) u optimizado.
5. *Arquitectura de procesos*	Ad hoc o sin establecer	Proceso Básico. Conocimiento básico.	Identificado y estándarizado	Gestionado y utilizado por todos	Mejoramiento continuo (métricas) u optimizado.
6. *Arquietctura de Canales de entrega*	Ad hoc o sin establecer	Canales entrega Básico. Conocimiento básico.	Identificado y estándarizado	Gestionado y utilizado por todos	Mejoramiento continuo (métricas) u optimizado.
7. *Innovación Organizacional*	Ad hoc o sin establecer	Innovación Básica. Conocimiento básico.	Identificado y estándarizado	Gestionado y utilizado por todos	Mejoramiento continuo (métricas) u optimizado.
8. *Gestión de riesgo*	Ad hoc o sin establecer	Gestión riesgo Básica. Conocimiento básico.	Identificado y estándarizado	Gestionado y utilizado por todos	Mejoramiento continuo (métricas) u optimizado.

D(*) de AE propuesta	Nivel 1-Inicial	Nivel 2-Desarrollando	Nivel 3-Definido	Nivel 4-Administrado	Nivel 5-Optimizando
9. *Gestión y Gobierno de TI*	Ad hoc o sin establecer	Gestión y gobierno de TI Básico. Conocimiento básico.	Identificado y estándarizado	Gestionado y utilizado por todos	Mejoramiento continuo (métricas) u optimizado.
10. *Sistema de Ciberseguridad Corporativa*	Ad hoc o sin establecer	Ciberseguridad Básica. Conocimiento básico.	Identificado y estándarizado	Gestionado y utilizado por todos	Mejoramiento continuo (métricas) u optimizado.

(*) D=Dimensions. Fuente: elaboración propia

Tabla 5 – Valoración del Modelo de Madurez para el Framework propuesto

6. Conclusiones

Se ha presentado la propuesta de un Framework de Arquitectura Empresarial construido a patrir de una revisión sistemática de literatura, análisis e integración de estándares, análisis de procesos misionales en IES, finalizando con un modelo de madurez que permite su medición. El framework permite describir aspectos claves de una organización como Visión estratégica institucional, aplicaciones, información, infraestructura, procesos, Canales de entrega, Innovación Organizacional, Gestión de riesgo, Gestión y Gobierno de TI, Sistema de Ciberseguridad Corporativa. Este Framework puede utilizarse para modelar la AE de Instituciones de Educación Superior Públicas. En estas instituciones se encuentran procesos académico-administrativos, de investigación, de extensión y procesos de soporte. Estos macroprocesos son la base para definir la AE ya que impactan en los Sistemas de Información, datos y plataformas tecnológicas que son necesarias para la puesta en marcha y eficiencia de procesos en este tipo de Instituciones. El *Framework* propuesto fue construido pensando en Universidades e Instituciones Educativas, pero fácilmente puede aplicarse para alinear los objetivos estratégicos con la plataforma de TI de cualquier organización. Para trabajos futuros, se recomienda instanciar el modelo y analizar el nivel alcanzado por la institución, de acuerdo a las variables establecidas y el nivel de cumplimiento de las mismas.

Referencias

Andersen, P., Carugati, A., & Sorensen, M. G. (2015). Exploring Enterprise Architecture Evaluation Practices: The Case of a Large University. *2015 48th Hawaii International Conference on System Sciences*, 4089–4098. https://doi.org/10.1109/HICSS.2015.491

Burégio, V. (2015). An Architecture and Guiding Framework for the Social Enterprise. *IEEE Internet Computing*, *19*(1), 64–68. https://doi.org/10.1109/MIC.2014.85

Cabrera, A., Carrillo, J., Abad, M. , Jaramillo, D. & Romero, F.(2015). Diseño y Validación de Arquitecturas de Aplicaciones Empresariales. Revista lbérica de Sistemas y Tecnologías de Información. ISSN: 1646-9895. Nro. EP4. Pag 79-91. http://www.scielo.mec.pt/pdf/rist/nspe4/nspe4a07.pdf

Cano, J. (2015). Repensando los fundamentos de la gestión de riesgos. Una propuesta conceptual desde la incertidumbre y la complejidad. Revista lbérica de Sistemas y Tecnologías de Información. ISSN: 1646-9895. Nro. E15 Pag 76-87. http://www.risti.xyz/issues/ristie15.pdf

Carrillo, J., Cabrera, A., Román, C., Abad, M., & Jaramillo, D. (2010). Roadmap for the implementation of an enterprise architecture framework oriented to institutions of higher education in Ecuador. *ICSTE 2010 - 2010 2nd International Conference on Software Technology and Engineering, Proceedings*, *2*, 7–11. https://doi.org/10.1109/ICSTE.2010.5608752

Chen, S., Tang, Y., & Li, Z. (2016). UNITA: A Reference Model of University IT Architecture. In *Proceedings of the 2016 International Conference on Communication and Information Systems* (pp. 73–77). New York, NY, USA: ACM. https://doi.org/10.1145/3023924.3023949

Gao. (2001). A Practical Guide to Federal Enterprise Architecture. *Public Law*, *1*(February 2001), 112.

J., L., K., B., Lakhrouit, J., & Baina, K. (2013). State of the art of the maturity models to an evaluation of the enterprise architecture. *2013 3rd International Symposium ISKO-Maghreb*, 1–8. https://doi.org/10.1109/ISKO-Maghreb.2013.6728119

Javanbakht, M., Pourkamali, M., & Feizi, M. (2009). A new method for enterprise architecture assessment and decision-making about improvement or redesign. *4th International Multi-Conference on Computing in the Global Information Technology, ICCGI 2009*, 69–76. https://doi.org/10.1109/ICCGI.2009.18

Lakhrouit, J., & Baina, K. (2016). Evaluating enterprise architecture complexity using fuzzy AHP approach: Application to university information system. *Proceedings of IEEE/ACS International Conference on Computer Systems and Applications, AICCSA*, *2016–July*.https://doi.org/10.1109/AICCSA.2015.7507211

Lapalme, J. (2008). Three Schools of Thought on Enterprise Architecture. *IT Professional*, *10*(2), 37–43.

Llamosa-Villalba, R., Carreño, L. T., Paéz, Q. A. M., Delgado, Q. D. J., Barajas, A. B., & Sneyder, E. G. (2015). Enterprise architecture of Colombian Higher Education. *Proceedings - Frontiers in Education Conference, FIE*, *2014*. https://doi.org/10.1109/FIE.2015.7344353

Olsen, D. H., & Trelsg??rd, K. (2016). Enterprise Architecture Adoption Challenges: An exploratory Case Study of the Norwegian Higher Education Sector. *Procedia Computer Science*, *100*(1877), 804–811. https://doi.org/10.1016/j.procs.2016.09.228

Petrikina, J., Drews, P., Schirmer, I., & Zimmermann, K. (2014). Integrating business models and enterprise architecture. *Proceedings - IEEE International Enterprise Distributed Object Computing Workshop, EDOCW*, 47–56. https://doi.org/10.1109/EDOCW.2014.16

Ramadhan, A., & Arman, A. A. (2014). Enterprise Architecture in University: Analyzing of Implementaion Using Business Process Management. *ICT For Smart Society (ICISS), 2014 International Conference*, 250–255. https://doi.org/10.1109/ICTSS.2014.7013182

Riascos, S., Aguilera,A. & Achicanoy, H. (2016). Inversión en Tecnologías de la Información y las Comunicaciones y su relación con en el direccionamiento estratégico de las PYMES de Santiago de Cali – Colombia. Revista lbérica de Sistemas y Tecnologías de Información. ISSN: 1646-9895. Nro. 18. Pag 1-17. In: http://www.scielo.mec.pt/pdf/rist/n18/n18a02.pdf

Shah, H., & Kourdi, M. El. (2007). Frameworks for Enterprise Architecture. *IT Professional*, *9*(5), 36–41.

Valencia, F. & Alzate, F. Metodología para la implementación de un Sistema de Gestión de Seguridad de la Información basado en la familia de normas ISO/IEC 27000. Revista lbérica de Sistemas y Tecnologías de Información. ISSN: 1646-9895. Nro. 22. Pag 73-88. In: http://www.scielo.mec.pt/pdf/rist/n22/n22a06.pdf

Vanti, A, Solana, P., & Seibert, R. (2018). Gobernanza Corporativa y Gobernanza Corporativa de TI utilizando Analytic Hierarchy Process en la creación de valor. Revista lbérica de Sistemas y Tecnologías de Información. Nro.27. ISSN: 1646-9895. Pag 86-108. In: http://www.risti.xyz/issues/risti27.pdf

Vieira, R., Cardoso, E., & Becker, C. (2014). A Traceable Maturity Assessment Method Based on Enterprise Architecture Modelling. *2014 IEEE 18th International Enterprise Distributed Object Computing Conference Workshops and Demonstrations*, 245–253. https://doi.org/10.1109/EDOCW.2014.44

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 08/09/2018
Aceitação/Acceptance: 16/12/2018

Análisis del valor potencial de la tecnología: Caso Universidad Francisco de Paula Santander Ocaña

Dewar Rico-Bautista[1], Cesar D. Guerrero[2], Yurley Medina-Cárdenas[3], Antón García-Barreto[4]

dwricob@ufpso.edu.co, cguerrer@unab.edu.co, ycmedinac@ufpso.edu.co, anton@ufpso.edu.co

[1] Departamento Sistemas e Informática. Universidad Francisco de Paula Santander Ocaña, Sede Algodonal Vía Acolsure, 546551, Ocaña, Colombia.

[2] Centro de Excelencia y Apropiación en IoT - Nodo Oriente, Grupo de Investigación en Tecnologías de Información. Universidad Autónoma de Bucaramanga, Avenida 42 No. 48 – 11, 680003, Bucaramanga, Colombia.

[3] Coordinadora SIG. Universidad Francisco de Paula Santander Ocaña, Sede Algodonal Vía Acolsure, 546551, Ocaña, Colombia.

[4] Jefe División de Sistemas. Universidad Francisco de Paula Santander Ocaña, Sede Algodonal Vía Acolsure, 546551, Ocaña, Colombia.

Pages: 756–774

Resumen: La tecnología es la aplicación intencionada de la información en el diseño, la producción y la utilización de bienes y servicios. De igual forma participa en la organización de las actividades humanas. Se contextualiza desde las arquitecturas, tecnologías y su vital importancia para el concepto emergente de universidad inteligente. El propósito de este documento es dar a conocer la situación actual y el progreso de mejora continua. Se dan a conocer los resultados hasta el momento generados en la universidad Francisco de Paula Santander Ocaña, Colombia, desde el proceso de apoyo sistema de información, telecomunicaciones y tecnología.

Palabras-clave:Impacto de la tecnología, campus inteligente, universidad inteligente.

Analysis of the potential value of technology: Case of Universidad Francisco de Paula Santander Ocaña

Abstract: Technology is the intentional application of information in the design, production and use of goods and services. It also participates in the organization of human activities. It is contextualized from the architectures, technologies and their vital importance for the emerging concept of intelligent university. The purpose of this document is to make known the current situation and the progress of continuous improvement. The results so far generated in the University Francisco de Paula Santander Ocaña, Colombia,from the process of support Information System, Telecommunications and Technology.

Keywords*:* Impact of Technology, smart campus, smart university

1. Introducción

En Colombia, el Plan Nacional de Desarrollo (Colombia, 2015) se basa en tres pilares, entre ellos la educación. Asume la educación como el más poderoso instrumento de igualdad social y crecimiento económico en el largo plazo, con una visión orientada a cerrar brechas en acceso y calidad al sistema educativo. Dentro de las estrategias transversales y regionales para la consolidación de los pilares se encuentran: *Competitividad e infraestructura estratégicas*; *movilidad social; transformación del campo; seguridad, justicia y democracia para la construcción de paz; buen gobierno y crecimiento verde*.

De igual forma, el consejo privado de competitividad lanza su undécimo Informe Nacional de Competitividad (Consejo Privado de Competitividad, 2017), donde da a conocer ciertas cifras de Colombia. En educación, la calidad de la educación superior en Colombia tiene una brecha importante que cerrar. Apenas el 14,9 % de los programas de pregrado de educación superior y el 14,5 % de las instituciones de educación superior cuentan con acreditación de alta calidad. La calidad en la educación superior, tomado del modelo del CNA (Generales & Calidad, 2006), supone el esfuerzo continuo de las instituciones por cumplir en forma responsable con las exigencias propias de cada una de sus funciones misionales: la investigación, la docencia y la proyección social; comprometiéndolas a un mejoramiento continuo (Enrique & Silva, 1997).Para el cumplimiento y gestión de dichos procesos misionales, la universidad cuenta con insumos y recursos: humanos (docentes, directivos y administrativos), físicos (infraestructura física), financieros (transferencias de la nación y de las entidades territoriales, ingresos propios por matrículas y derechos académicos y la tecnología (infraestructura TI y sistemas de información), (Ministerio de Educación Nacional, 2013; Roa, 2003). Dichos insumos de cada universidad son medidos desde el ministerio de educación a través de diferentes modelos e indicadores, que al final representan su capacidad total de gestión, constituyéndose así en una medida de eficiencia.

La tecnología hace parte de la vida cotidiana, ésta es usada para fines tan variados como la comunicación, el trabajo, entretenimiento, enseñanza, aprendizaje, entre otros. La comunicación entre el hombre y la máquina, para hacer posible su finalidad, ha sido motivo de estudio en los años recientes, en los cuales la mediación entre el hombre y la máquina aumenta a causa de las aplicaciones computacionales que requieren cada vez más interacción para lograr un aprovechamiento eficiente de su potencial (Rautaray & Agrawal, 2012). La tecnología se convierte en un elemento fundamental de la universidades, sin la cual no se podría llevar a cabo ninguna de sus funciones esenciales (Gómez, Jimenez, Gumbau, & Llorens, 2017). Para mejorar la gestión de sus recursos las universidades diseñan su direccionamiento estratégico en función de unas necesidades propias, pero no se tienen en cuenta los requerimientos de orden nacional que demandan su innovación y competitividad en áreas como la infraestructura tecnológica. Sus planes de desarrollo consideran la tecnología como soporte, pero no como una necesidad estratégica para su funcionamiento. Se delega su implementación para suplir necesidades específicas de software y hardware, pero no se evalúa la capacidad actual para aportar a los objetivos estratégicos de las universidades.

Al ser el pilar básico del desarrollo de las comunidades del futuro les permiten dotar de inteligencia a todos sus ámbitos y generar servicios que proporcionen una mejora calidad de vida a los ciudadanos y una mayor sostenibilidad a las ciudades y universidades (MACIÁ, 2017). Sin embargo, no convendría concebirla sólo como un elemento táctico,

no deberían planificarse de manera aislada, sino que tendrían que formar parte de la planificación global de la universidad (Fernández Martínez & Llorens Largo, 2016).

Este artículo se encuentra dividido principalmente en (i) introducción, (ii) Arquitecturas y tecnologías, (iii) Universidad inteligente, (iv) Aplicación de metodología de Impacto de la tecnología en la UFPS Ocaña (UFPSO), (v) Resultados parciales y por último conclusiones.

2. Arquitecturas y tecnologías

El mundo ha cambiado, la tecnología ha cambiado el mundo. Tecnología y sociedad son ahora compañeros inseparables, se han fusionado para ofrecer un nuevo nivel de servicios(Y. C. Medina-Cárdenas, Areniz-Arévalo, & Rico-Bautista, 2016; Medina Cárdenas, Rico Bautista, Constanza, & Cárdenas, 2009; Medina & Rico, 2008). TI no es sólo hablar de computadores, dispositivos móviles, sensores, redes, es hablar de nuevos conceptos y paradigmas socio-tecnológicos (Mineraud, Mazhelis, Su, & Tarkoma, 2016), (Cárdenas & Bautista, 2012)(Y. Medina-Cárdenas & Rico- Bautista, 2012).

Para analizar, proponer y desarrollar infraestructura en una organización, por ejemplo en las universidades, se debe conocer las arquitecturas y sus tecnologías (Lin et al., 2017), (Marotta, Both, Rochol, Granville, & Tarouco, 2015). Con este fin, se ofrece un panorama general con respecto a la arquitectura de los sistemas, las tecnologías, la implementación en campus inteligentes y universidades inteligentes (Zapata-Ros, 2018) , como se observa en la Figura 1. En el caso de la arquitectura en capas, el número de capas decide la complejidad de la arquitectura. La arquitectura de cinco capas es la arquitectura ideal desde la perspectiva de la seguridad y la compatibilidad (Ara, Gajkumar Shah, & Prabhakar, 2016). Diversas tecnologías soportan una infraestructura de red generalizada, destacándose el desarrollo de IoT y sus dispositivos (Khan, Khan, Zaheer, & Khan, 2012), (Rodríguez Molano et al., 2015). Estos se clasifican en dispositivos con recursos limitados y dispositivos ricos en recursos. Los dispositivos ricos en recursos como un teléfono inteligente, computadora personal estándar o un servidor tienen suficiente hardware, software y memoria que admiten el protocolo TCP / IP. Los dispositivos con limitaciones de recursos como dispositivos basados en microcontroladores, sensores y actuadores no tienen suficientes capacidades de hardware / software que admitir el protocolo TCP / IP (Ara et al., 2016).

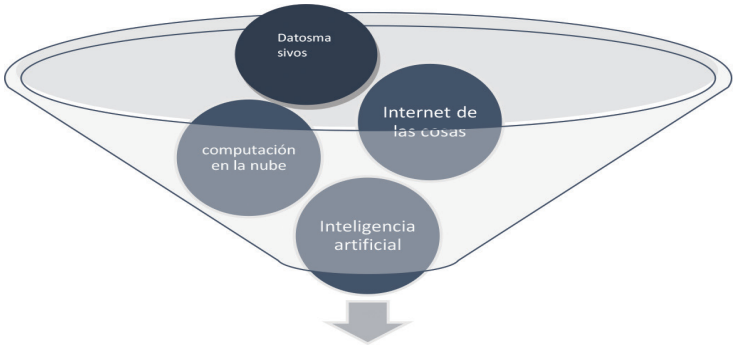


Figura 1 – Tecnologías que se implementan en campus inteligentes y universidades inteligentes.

En los últimos años, varios estudios publicados examinaron IoT desde diferentes aspectos (Dar, Taherkordi, Baraki, Eliassen, & Geihs, 2015). Por ejemplo, el trabajo realizado por Atzori(Atzori, Iera, & Morabito, 2017; Nitti, Atzori, & Cvijikj, 2015) presentó las diferentes visiones de IoT, generando un primer acercamiento a su comprensión. En la revisión realizada por Al-Fuqaha y otros (Al-Fuqaha, Guizani, Mohammadi, Aledhari, & Ayyash, 2015), se presentaron los protocolos y las posibles aplicaciones de la IoT, se proporcionó una visión general y se presentaron los retos clave para trabajos futuros donde se incluyen temas de seguridad y privacidad (Tan, Lu Tan, & Neng Wang, 2010), (Liu & Wang, 2017), (W. jun Lee, 2016), (Sánchez-Torres, Rodríguez-Rodríguez, Rico-Bautista, & Guerrero, 2018). Así también, el trabajo realizado por Andrea y otros (Chatzigiannakis, Vitaletti, & Pyrgelis, 2016) presentó las vulnerabilidades de seguridad y los desafíos de IoT desde el punto de vista de las aplicaciones, las redes y los sistemas físicos. El trabajo de Sha y otros (Tandon, Shah, & Gajjar, 2016), planteó problemas y oportunidades IoT.

Además de los documentos antes mencionados, Botta y otros (Botta, De Donato, Persico, & Pescap, 2016)consideraron la integración del cloudcomputing con IoT. Asimismo, Wu y Zhao (Xiong et al., 2017), propusieron una nueva infraestructura de IoT, a saber, WInternet, que cumple varios requisitos de IoT. Los trabajos de Sheng y otros (Ngu, Gutierrez, Metsis, Nepal, & Sheng, 2017); Palade(Razzaque, Milojevic-Jevric, Palade, & Clarke, 2016); Rodrigues y otros (Cruz, Rodrigues, Al-Muhtadi, Korotaev, & Albuquerque, 2018),presentan el middleware como una tecnología clave en la realización de los sistemas IoT, que suele describirse como un sistema informático diseñado para ser el intermediario entre los dispositivos y las aplicaciones IoT. En el documento de Lee, Bae y Kim (S. K. Lee, Bae, & Kim, 2017), se centran en la capa de red, que para ellos es la parte más importante de la realización del entorno de IoT. La idea de IoT, por lo tanto, ha evolucionado con el tiempo y los cambios se seguirán produciéndose en los próximos años con la llegada e incorporación de nuevas tecnologías, como la computación en nube, ICN, Big data, las redes sociales, entre otras (Akyildiz, Nie, Lin, & Chandrasekaran, 2016; Al-Fuqaha et al., 2015; European Technology Platform on Smart Systems Integration, 2008; Gagliardi et al., 2017; Tan et al., 2010; Weyrich & Ebert, 2016).

3. Universidad inteligente

En la actualidad, se incorpora el término de universidad inteligente "Smart university"(Rico-Bautista, Parra-Valencia, & Guerrero, 2017). Las universidades pueden considerarse universidades inteligentes, ya que utilizan de manera rentable las tecnologías disponibles para mejorar su desempeño y mejorar la calidad de sus graduados (Aldowah, Ul Rehman, Ghazal, & Naufal Umar, 2017), (Bueno-Delgado, Pavón-Marino, De-Gea-García, & Dolón-García, 2012). A pesar de esto, todavía hay espacio para mejorar y las universidades deberían convertirse en universidades más inteligentes (Shvetsova, 2017). En una universidad más inteligente, las mejores soluciones tecnológicas fomentan la colaboración y la cooperación entre las personas (Coccoli, Guercio, Maresca, & Stanganelli, 2014). Las universidades inteligentes y la educación inteligente están surgiendo y están creciendo rápidamente en áreas que representan una integración de 1) sistemas inteligentes e inteligentes, objetos inteligentes y entornos inteligentes, 2) tecnologías inteligentes, varias ramas de la informática y la ingeniería informática, 3)

sistemas, agentes y herramientas de software y/o hardware educativos inteligentes de última generación, y 4) pedagogía innovadora y estrategias de enseñanza y metodologías de aprendizaje basadas en tecnologías avanzadas (Savov, Terzieva, Todorova, & Kademova-Katzarova, 2017), (Rueda-Rueda, Manrique, & Cabrera Cruz, 2017).

Maciá, Berná, Sánchez, Lozano, & Fuster (MACIÁ, 2017), expresa que el concepto de Smart University. Mantiene el mismo objetivo principal de mejorar la calidad de vida de su comunidad aplicando de forma global, intensiva y sostenible las TI bajo el principio de servicio a la ciudadanía. Es por eso por lo que se trata de un entorno controlado, en el cual las políticas económicas, de sostenibilidad, construcción o desarrollo deciden den forma local y no dependen de factores externos, aunque siguen criterios similares a las políticas y estrategias nacionales o internacionales (Banica, Burtescu, & Enescu, 2017; Coccoli et al., 2014; Heinemann & Uskov, 2018; Staskeviciute & Neverauskas, 2008). La infraestructura actual permite disponer de sistemas digitales conectados de manera global, capaces de soportar aplicaciones, servicios, plataformas y sensores, todos ellos funcionando en conjunto bajo un ecosistema o marco estandarizado de normas y protocolos.

4. Impacto de la tecnología

Determinar el valor potencial de la tecnología con la que cuenta una organización, por ejemplo, la universidad, es un tema que no es fácil de responder. Y la razón es que sólo tenemos datos sobre el pasado, pero gustaría saber qué ocurrirá en el futuro. Para citar a Clayton Christensen(C M Christensen, 1997; Clayton M. Christensen, 2006), la única manera de mirar hacia el futuro, ya que no hay datos, es tener una muy buena teoría o un modelo. Intentar dibujar un mapa del futuro. Cómo será el futuro, cómo puede ser el futuro si se tiene en cuenta todas estas cosas. Pero mira si se utiliza todo el potencial de la tecnología que se tiene hoy. ¿Cómo puede ser ese sistema, industria o en particular la universidad en el futuro?. Se divide el desarrollo en varias fases. La última fase es esa perspectiva de 10 a 20 años, y la primera es hoy. Y luego se podría dividir entre el escenario actual y el futuro en varias fases. Por lo general, sólo se inicia con uno porque es más simple. Así que, al tener el presente, tenemos algún tipo de fase intermedia, y luego se tiene el escenario futuro allí. Entonces normalmente se miran los detonantes. ¿Qué habrá que hacer para pasar de la primera a la segunda fase, o de la segunda a la fase final? Y eso puede ser, por ejemplo, que se necesita reglamentación, o estandarización.

Se necesita desde luego, un detalle de nivel de costos. Debe ser lo suficientemente barato, se necesita confiar desde allí, en los clientes potenciales, etc. Hay una serie de cosas que necesitan estar en su lugar para que esto suceda. Luego, se tiene que ver cómo se puede desarrollar ese papel a lo largo de las diferentes fases. Y lo que puede traer éxito en la última fase puede no ser lo mismo que lo que se necesita para construir esa posición en la fase intermedia. Así que también, se incluye un problema de tiempo. Así que cuando se construye el mapa, se de miran los diferentes desencadenantes y entonces se visualiza cuánto tiempo tomará esto. Por lo tanto, un indicador de que ahora se encuentra listo para entrar en una nueva fase, es que todos los factores desencadenantes están ahí.Una vez que se tenga ese mapa se verá lo que pueden ser destinos interesantes en ese mapa. Destinos en este sentido es otra palabra para los roles de negocios en una red de valor. Una vez que se ha dibujado el mapa, visto los destinos y encontrado uno que puede ser

interesante, también se puede tratar de averiguar cómo se muevedesde el estado inicial a hoy a ese destino deseado en el futuro.

4.1. Plan Estratégico TIC, el mapa de ruta del futuro.

El enfoque por procesos, como se presenta en la figura 2, permite mejorar la satisfacción de los clientes y el desempeño de la gestión de la Universidad en su misión de formar profesionales idóneos; la implementación de la NTCGP 1000:2009 permite el cumplimiento de la norma internacional ISO 9001:2008 que brinda como beneficios herramientas para el uso eficiente de los recursos, la toma de decisiones basada en evidencias objetivas y hacia el logro del plan de desarrollo y el cumplimiento de los requisitos y las necesidades en materia de formación profesional.La universidad Francisco de Paula Santander Ocaña (UFPS Ocaña), en búsqueda de su mejoramiento continuo, rediseñó varios de sus procesos, dentro de los cuales se encuentran los procesos administrativos, ampliación de la cobertura, implementación de indicadores de gestión, modernización de los sistemas de información y la acreditación de los programas; dichos mejoramientos van dirigidos a la efectividad de los procesos de la institución. Sin dejar de lado el cumplimiento de normas, estrategias y directrices gubernamentales como lo es Gobierno en Línea.

La institución ha implementado Tecnologías de Información y Comunicación (TIC´s) como herramientas que permiten un desempeño efectivo en la gestión de la misma, transversales a los diferentes procesos administrativos y de gestión académica.La elaboración del Plan Estratégico TIC´s en la UFPS Ocaña desde al año 2015 (Arrieta, Sanguino, & Lobo, 2015), se convirtió en los lineamientos del desarrollo tecnológico en los procesos estratégicos (dirección y planeación), misionales (gestión académica, investigación y extensión), de apoyo (gestión bienestar universitario, gestión humana, sistemas de información, telecomunicaciones y tecnología, gestión administrativa y financiera, gestión de comunicaciones, secretaría general, infraestructura y mantenimiento y admisiones, registro y control) y de evaluación y mejora (control interno, sistema integrado de gestión). Con base en el diagnóstico, se definieron los objetivos estratégicos que orientan el desarrollo del plan a través de proyectos, desglosados en actividades y responsables.La visión se concibe en un horizonte de (5) años, y se concreta un primer momento o espacio de (3) años, el cual se considera apropiado para implementar y hacer seguimiento a las acciones definidas, de acuerdo con la dinámica que vive la institución.

Han pasado cinco siglos y la revolución tecnológica se encuentra inmersa en la sociedad; donde esta ha logrado ingresar fácilmente. Las nuevas tecnologías han cambiado el modo en que las personas trabajan, se relacionan y aprenden; por tal razón esta se ha integrado en la enseñanza; últimamente ha crecido su popularidad en el contexto educativo y por consiguiente los gobiernos trabajan para reducir la brecha digital.Este proceso de evolución tecnológica ha transformado el proceso de enseñanza-aprendizaje a través de la difusión de hardware y software educativo, generando un entorno de trabajo dinámico donde el profesor y el alumno disponen de un sinfín de herramientas y posibilidades educativas.

Por esta razón la UFPS Ocaña en su apuesta a la mejora continua y en la búsqueda de la calidad, en el sistema integrado de gestión se encuentra el proceso de apoyo

Sistema de Información, Telecomunicaciones y Tecnología, cuyo objetivo es diseñar, desarrollar, implementar, administrar y mantener los sistemas de información, las telecomunicaciones y la infraestructura tecnológica; asesorar y gestionar la adquisición e implementación de nuevas tecnologías de información y comunicaciones, que brinden soluciones eficaces, efectivas y oportunas a las necesidades del cliente y partes interesadas, teniendo en cuenta el uso eficiente de los recursos tecnológicos, minimizando el impacto ambiental y bajo un ambiente laboral propicio para los trabajadores.

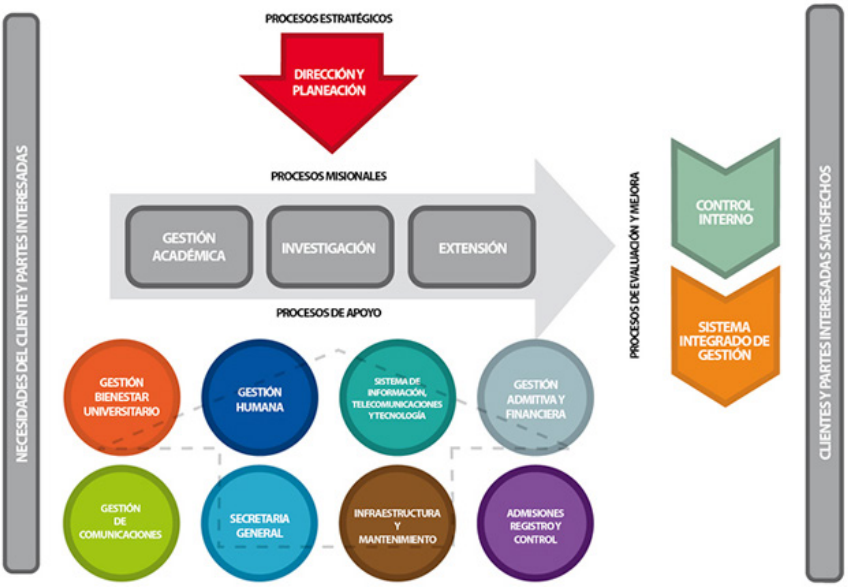


Figura 2 – Mapa de procesos UFPS Ocaña. Fuente: (Medina, 2014)

Al incluir tecnología en la organización, no solo se mejora la conectividad, el acceso a la información y el tiempo de respuesta en los procesos, también la formación académica logra enmarcase en un contexto dinámico, adaptado a los cambios que trae consigo la globalización.

4.2. Prospectiva plan estratégico TIC

¿Por qué se hace planificación y previsión de escenarios? Bueno, eso se debe a que el futuro no está predeterminado ni es predecible. Y, por supuesto, si lo fuera, no tendría sentido actuar hoy porque no tendría ningún efecto en el futuro. Y también la razón es que no se tiene toda la información sobre el futuro, así que se necesita tener alguna ayuda para hacer estas conjeturas.Cuando se construye un modelo o una teoría, se usan datos históricos y con el uso de la inducción se produce.

Se tienen los datos históricos y también se conoce unas actividades sobre el futuro, hay algunas preguntas que hacer, pero la cuestión principal es que las cosas que no sabemos que no sabemos. La UFPS Ocaña, cuenta actualmente con un Plan de Desarrollo Institucional “Hacia la Excelencia Institucional” formulado para la vigencia 2014 –

Factoresclavedeéxito – FCE		Descripciónde l FCE	Objetivogeneral	PotencialcontribuciónlasTI
FCE 1	*Investigación y Formación Académica*	Implica la necesidad de considerar a la investigación como eje transversal de la formación académica.	Incorporar e implementar las TIC en los procesos académicos, la cualificación docente, la calidad y pertinencia de la oferta, la cobertura y el desarrollo estudiantil.	Mejorar los procesos académicos a través de la apropiación de las TI.
FCE 2	*Desarrollo físico y tecnológico*	Comprende la modernización de la infraestructura tecnológica y la adecuación de los espacios para el desarrollo de las actividades académicas y administrativas.	Fortalecer la gestión tecnológica, modernizar los recursos y adecuación de espacios físicos suficientes para el desarrollo de las funciones institucionales.	Mejorar la infraestructura física y tecnológica.
FCE 3	*Sostenibilidad administrativa y financiera*	Comprende el mejoramiento de los procesos de planeación, ejecución y evaluación administrativa y financiera.	Implementar y mantener procesos eficientes en la planeación, ejecución y evaluación administrativa y financiera.	Optimizar los servicios institucionales para lograr la satisfacción de sus clientes y terceros.
FCE 4	*Bienestar Institucional*	Tiene en cuenta la creación de programas para la formación integral y el desarrollo de los miembros de la comunidad universitaria.	Generar programas para la formación integral, el desarrollo humano y el acompañamiento institucional que permitan el mejoramiento de las condiciones de vida de la comunidad universitaria.	Brindar soporte para el mejoramiento de los servicios de seguimiento a egresados y acompañamiento a estudiantes.
FCE 5	*Visibilidad nacional e internacional*	Comprende las acciones necesarias para fortalecer la investigación, docencia y extensión en el ámbito internacional.	Integrar, transformar y fortalecer las funciones de investigación, docencia y extensión para su articulación en un ambiente globalizado de excelencia y competitividad.	Brindar soporte para la creación de espacios para trabajo colaborativo con grupos de investigación a nivel internacional.
FCE 6	*Impacto y proyección social*	Tiene en cuenta la proyección social de la Universidad en la región.	Desarrollar las capacidades institucionales promoviendo impactos positivos en la región, el medio ambiente y la comunidad.	Brindar soporte para las diferentes actividades de extensión, a través de la implementación de TI.

Tabla 1 – Factores clave de éxitoUFPS Ocaña. Fuente: (Arrieta et al., 2015)

2019, con el cual reafirma su voluntad y compromiso de hacer de la educación superior, un factor dinamizador del desarrollo regional y nacional. Así mismo plantea, que el posicionamiento de la universidad en el siglo XXI, implica incorporar las Tecnologías de Información y Comunicación a todos los procesos institucionales, administrativos y académicos, además de fortalecer los procesos de conectividad que le permitan ser un referente para las demás instituciones educativas y administraciones municipales

de la región. Según Rockart(1979), los Factores Clave de Éxito - FCE se definen como el número limitado de áreas, en las cuales los resultados, si son satisfactorios, asegurarán un funcionamiento competitivo y exitoso para la organización. Para el caso de la UFPS Ocaña por ser una institución pública, los FCE se constituyen en las variables que pueden afectar su posición comparativa y de cumplimiento de su orientación estratégica. En la tabla 1, se muestra en color verde aquellos FCE que son más sensibles o en los que podría existir una mayor contribución potencial de las Tecnologías de Información TI, en color amarillo los medianamente sensibles, y finalmente en color rojo los que de manera comparativa con los anteriores serían los menos sensibles y/o sólo con impacto indirecto de las TI en los FCE. Los FCE son tomados de los ejes estratégicos del Plan de Desarrollo Institucional vigencia 2014 – 2019.

4.3. Red de Valor

Figura 3 – Red de valor UFPS Ocaña. Fuente: Fuente: (Arrieta et al., 2015)

Para que la tecnología tenga un impacto, es esencial que se cree un ecosistema de negocios que funcione, un sistema donde todos los roles necesarios estén en su lugar y en equilibrio. Son estos papeles los que asumen las empresas y los que también se dirigen a los inversores y empresarios, así que ¿qué es entonces una red de valor? El concepto de la cadena de valor fue derivado por primera vez por Michael Porter en su libro, TheCompetitiveAdvantage,(1985). Y aborda la creación de valor a nivel de empresa, es decir, a nivel de empresa o de organización. La red de valor representa un sistema de roles con modelos de negocio alineados, observados en la figura 3. Sin una red de valor que funcione, la industria no funcionaría. Sin la creación de valor en un rol que cree valor para aquellos a los que el rol se relaciona, el rol no sobrevivirá. Para entender la red de valor, es fundamental ser capaz de predecir el impacto de la tecnología.

4.4. Estrategia

Ahora se necesita pensar en cómo se puede manejar la transformación desde la situación actual hasta donde se quiere estar en el futuro. Alcanzar el estado final deseado y de eso

se trata la estrategia. Para hacer estrategia correctamente usted tiene que entender la naturaleza de su desafío. No se puede resolver un problema que no se ha identificado. En esencia, es necesario ser capaz de entender cuál es el problema más crítico y encontrar una solución que aborde ese problema específico. En términos simples, la estrategia se refiere a lo que se necesita hacer para lograr una meta y tácticas particulares, cómo llevar a cabo estas acciones, observado en la tabla 2.

Actividades	**Indicador**
Gestionar el análisis de Hacking Ético, con prácticas de hardening y/o pentesting a los servicios de la Universidad	Informe de Auditoria
Socialización del Plan Estratégico de TIC a la comunidad universitaria, las Políticas de Seguridad de la Información y las Políticas de Tratamiento de Datos Personales	PETIC, Políticas de Seguridad y Políticas de Tratamiento de Datos Personales socializadas a la comunidad universitaria.
Elaboración de controles dados en la guía de buenas prácticas de TI basado en ISO 27002.	Controlesestablecidos.
Formulación del Plan de Continuidad de TI.	Plan de continuidad de TI
Desarrollo, Implementación y Mantenimiento de los Sistemas de Información de la intranet de la Universidad.	Modulo del Sistema de Información Total / Modulo del Sistema de Información Implementados
Gestionar la adquisición de una planta de emergencia para la subestación eléctrica de la Universidad.	Respuesta a la solicitud hecha a Planeación
Gestionar un sistema de video vigilancia para el campus universitario.	Porcentaje aumentado de cámaras de video vigilancia
Gestionar el sistema de acceso digital para el portal Institucional.	Porcentaje de adecuación del Data Center para el centro de computo
Diseño e implementación de un sistema de acceso biométrico para las salas de computo del Bloque de aulas de la Universidad.	Actividad de control de Acceso
Gestionar la adquisición de una nueva torre de comunicaciones para la Sede de Pueblo Nuevo.	Requerimientospropuestos / Requerimientosatendidos
Gestionar con la Oficina de Planeación el inicio de la construcción del Data Center y centro de cómputo para la UFPS Ocaña (Incluyendo un sistema de seguridad para la granja de servidores, detectores de humo, lectores de proximidad y/o biométricos, cámaras de video actividad por movimiento, además la ubicación de una bodega para el almacenamiento de los elementos ajenos a un centro de datos.	Requerimientospropuestos / Requerimientosatendidos
Gestionar y Administrar el Servicio de Data Center para almacenamiento externo de los sistemas de información de la Universidad.	Contratación de alquiler mensual del servicio de información en la nube.
Diseño e implementación del cableado estructurado de las diferentes sedes y dependencias de la Universidad; estas incluyen: Edificio de postgrados, edificio de ingenierías, Granja Experimental (Con sus proyectos).	Porcentaje de Implementación
Actualización y mejoramiento del cableado estructurado de las diferentes sedes y dependencias de la Universidad; estas incluyen: Centro de Datos de la Universidad (División de Sistemas) y el Centro de datos en la Sede Bellas Artes como respaldo al centro de datos de la UFPS Ocaña.	Porcentaje de Actualización y Mejoramiento

Actividades	Indicador
Ampliación en la cobertura del servicio de voz IP en el campus Universitario.	Puntos voz IP propuestos / Puntos voz IP implementados
Implementación del protocolo IPv6 y enlaces inalámbricos de respaldo para comunicación y servicios en todas las sedes de la Universidad.	Porcentaje de protocolos IPv6 implementados en la sede
Ampliación del canal de datos e internet hacia el proyecto RENATA y para el área académico-administrativa, e implementar el servicio de movilidad eduroam en las sedes Algodonal, Bellas Artes e Invías.	Ampliación del canal / Ampliación Establecida
Actualización del servidor para la consola del antivirus e instalación del agente para cada uno de los equipos de la Universidad;en su totalidad para el año 2016, 1040 equipos; para el año 2017, 1072 equipo y para el año 2018, 1163 equipos actualizados; configuración de servidores virtualizados para respaldo de los servicios Web que brinda la Universidad y brindar mantenimiento y soporte a la granja de servidores, equipos tecnológicos y sistemas de información de la Universidad.	Requerimientospropuestos / Requerimientosatendidos

Tabla 2 – Estrategias e indicadores de desarrollo tecnológico

5. Resultados parciales: caso UFPS Ocaña

A continuación, se da a conocer el último informe de gestión procesosistema de información, telecomunicacionesy tecnología (SITT)a cargo de la dependencia de división de sistemas correspondiente al I semestre del año 2018 de los últimos avances realizados.En la figura 4 se muestran las actividades desarrolladas por el proceso SITT, las cuales obtuvieron el mayor cumplimiento en la implementación de la meta propuesta, y se nombran a continuación;

- Desarrollo, implementación y mantenimiento de los sistemas de información de la intranet de la universidad, izquierda. Red de datos y de voz, derecha.
- Gestión un sistema de video vigilancia para el campus universitario.
- Gestionar con la oficina de Planeación el inicio de la construcción del data center y centro de cómputo para la UFPS Ocaña.
- Gestionar y administrar el servicio de data center para almacenamiento externo de los sistemas de información de la universidad.
- Actualización del servidor para la consola del antivirus e instalación del agente para cada uno de los equipos de la universidad; en su totalidad para el año 2016, 1040 equipos; para el año 2017, 1072 equipo y para el año 2018, 1163 equipos actualizados; configuración de servidores virtualizados para respaldo de los servicios web.
- Ampliación en la cobertura del servicio de voz IP en el campus universitario.
- Socialización del plan estratégico de TIC a la comunidad universitaria, las políticas de seguridad de la información y las políticas de tratamiento de datos personales.

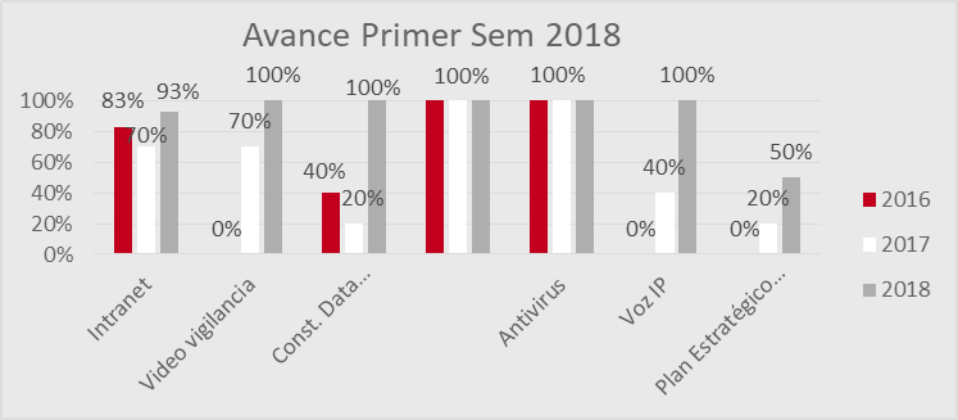


Figura 4 – Actividades desarrolladas por el proceso SITT

Para este año 2018, se estima alcanzar un 10 % adicional en la implementación de los controles dados, de los cuales se le ha dado cumplimiento a los siguientes requisitos de la norma por parte de los profesionales encargados de la administración de servidores, observado en la figura 5:

- Copias de Seguridad de la Información, se realizan diariamente y al realizar estas copias se envía un correo electrónico con el resultado del proceso ejecutado al administrador de servidores.
- Controles de Red, Se realiza el Rack cerrado con llave
- Mecanismos de Seguridad Asociados a Servicios en Red, Instalación de Firewall, de Antivirus.
- Segmentación de la red por medio de Vlan.

Se cuenta con un plan de contingencia de TI el cual ya está en revisión y en espera de aprobación con un documento borrador del Plan de continuidad de TI que está en desarrollo y revisión para su respectiva aprobación, observado en la figura 6.

Se cuenta con un análisis y diseño de los equipos que se van a adecuar e instalar para el desarrollo de esta actividad. Y se quiere llegar a la conformación de un registro deingreso y salida tanto para administrativos como estudiantes y visitante avanzando gradualmente en la implementación, observado en la figura 7.Se gestiona un (1) sistema biométrico de control de acceso para entrada y salida vehicular, cuatro (4) sistemas biométricos de entrada y salida peatonal y un (1) sistema biométricode entrada y salida para discapacitados.El proceso sigue atento a una buena prestación de servicio, cumpliendo con los tiempos instaurados y los seguimientos frecuentes a las actividades establecidas en el plan de acción, generando un nivel de percepción cada vez mejor, observado en la figura 8.

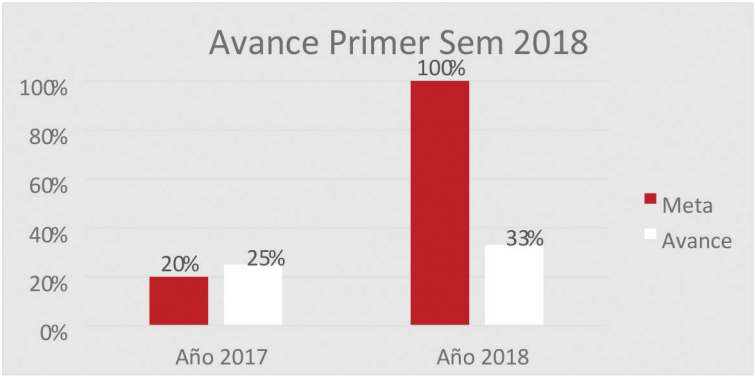


Figura 5 – Elaboración de controles dados en la guía de buenas prácticas de TI basada en ISO 27002

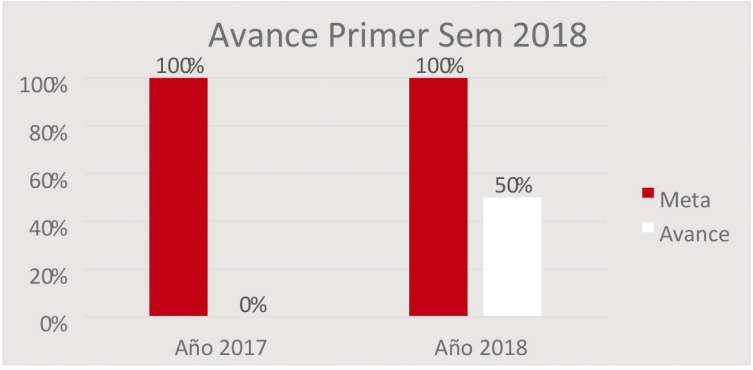


Figura 6 – Formulación del plan de continuidad de TI

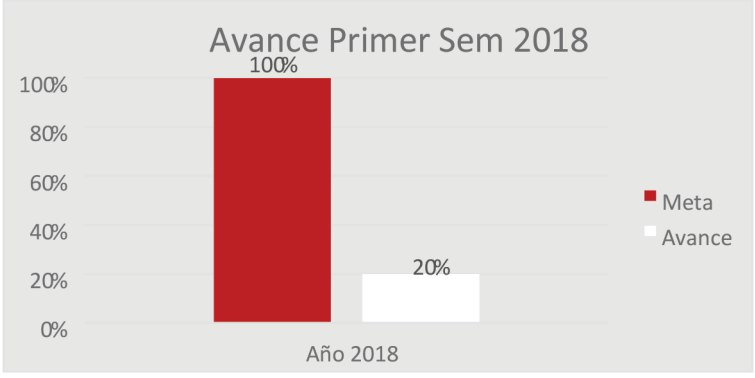


Figura 7 – Gestionar el sistema de acceso digital para el portal Institucional.

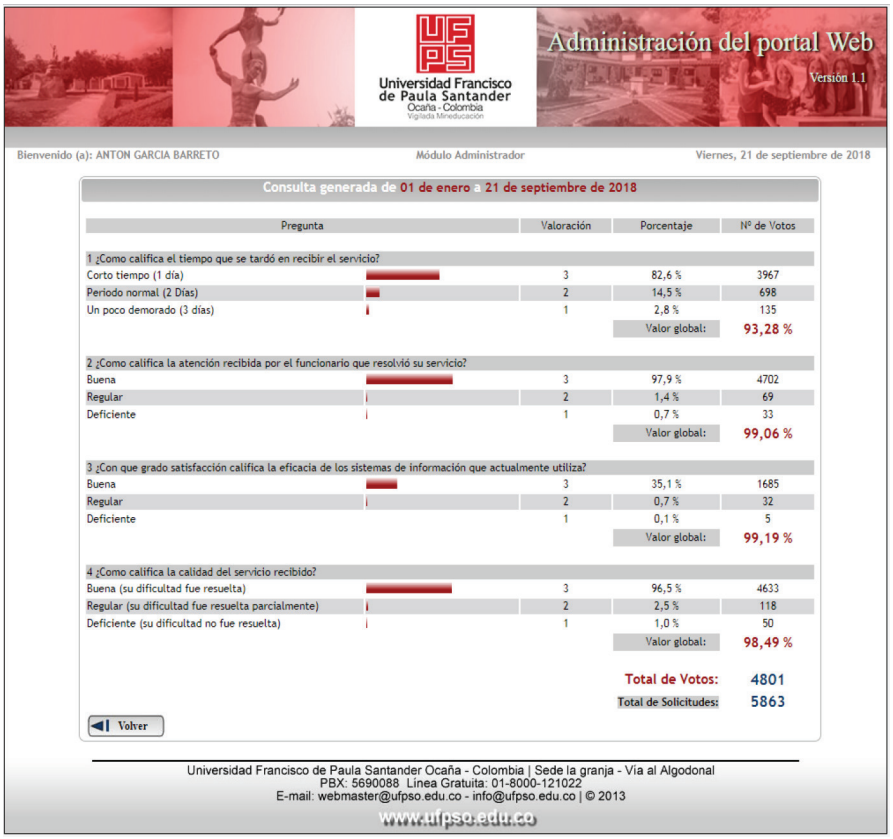


Figura 8 – Percepción de los usuarios del proceso de SITT de los últimos 3 años

6. Conclusiones

El proceso SITT, en busca de la mejora continua y hacia el camino a la excelencia académica, siguiendo los lineamientos de los procedimientos seguirá trabajando en el fortalecimiento de su infraestructura permitiendo un mejor desempeño de las actividades.

Se examinó cómo se relaciona el delta tecnológico (la nueva tecnología es mucho mejor en comparación con la tecnología que se está utilizando actualmente) con el impacto potencial y cómo se relaciona con las barreras para la implementación. Aunque no se puede saber exactamente cómo una nueva tecnología puede influir en un sistema. Se puede asumir que cuanto mayor es el delta tecnológico, mayor es el potencial para impulsar el cambio.

Se encuentran todavía barreras que superar. Cuanto mayor sea el delta y el potencial de creación de valor, mayor será la probabilidad de que se gestionen estas barreras. Pero dependiendo de cuán establecida esté la tecnología existente, la necesidad del delta

tecnológico puede variar. El momento oportuno, para estimar el momento oportuno, en que una nueva tecnología tendrá un impacto depende del tamaño del delta tecnológico, pero también de las barreras.

Referencias

Akyildiz, I. F., Nie, S., Lin, S. C., & Chandrasekaran, M. (2016). 5G roadmap: 10 key enabling technologies. *Computer Networks*, *106*, 17–48. https://doi.org/10.1016/j.comnet.2016.06.010

Al-Fuqaha, A., Guizani, M., Mohammadi, M., Aledhari, M., & Ayyash, M. (2015). Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications. *IEEE Communications Surveys and Tutorials*, *17*(4). https://doi.org/10.1109/COMST.2015.2444095

Aldowah, H., Ul Rehman, S., Ghazal, S., & Naufal Umar, I. (2017). Internet of Things in Higher Education: A Study on Future Learning. *Journal of Physics: Conference Series*, *892*, 012017. https://doi.org/10.1088/1742-6596/892/1/012017

Ara, T., Gajkumar Shah, P., & Prabhakar, M. (2016). Internet of Things Architecture and Applications: A Survey. *Indian Journal of Science and Technology; Volume 9, Issue 45, December 2016*. Retrieved from http://www.indjst.org/index.php/indjst/article/view/106507

Arrieta, M., Sanguino, M., & Lobo, C. (2015). *Diseño de un plan estratégico de tecnologías de información para la Universidad Francisco de Paula Santander Ocaña*. Retrieved from http://repositorio.ufpso.edu.co:8080/dspaceufpso/bitstream/123456789/814/1/27740.pdf

Atzori, L., Iera, A., & Morabito, G. (2017). Understanding the Internet of Things: definition, potentials, and societal role of a fast evolving paradigm. *Ad Hoc Networks*, *56*, 122–140. https://doi.org/10.1016/j.adhoc.2016.12.004

Banica, L., Burtescu, E., & Enescu, F. (2017). The impact of internet-of-things in higher education. *Scientific Bulletin-Economic Sciences*, *16*(1), 53–59.

Botta, A., De Donato, W., Persico, V., & Pescap, A. (2016). Integration of Cloud computing and Internet of Things: A survey. *Future Generation Computer Systems*, *56*, 684–700. https://doi.org/10.1016/j.future.2015.09.021

Bueno-Delgado, M. V, Pavón-Marino, P., De-Gea-García, A., & Dolón-García, A. (2012). The Smart University Experience: An NFC-Based Ubiquitous Environment. In *2012 Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing* (pp. 799–804). https://doi.org/10.1109/IMIS.2012.110

Cárdenas, Y. C. M., & Bautista, D. W. R. (2012). Mejores prácticas de gestión para la calidad de los servicios en tecnologías de información. *Gerencia Tecnológica Informática*.

Chatzigiannakis, I., Vitaletti, A., & Pyrgelis, A. (2016). A privacy-preserving smart parking system using an IoT elliptic curve based security platform. *Computer Communications*, *89*, 165–177. https://doi.org/10.1016/j.comcom.2016.03.014

Christensen, C. M. (1997). The Innovator's Dilemma. *Business*. https://doi.org/10.1515/9783110215519.82

Christensen, C. M. (2006). The ongoing process of building a theory of disruption. *Journal of Product Innovation Management*. https://doi.org/10.1111/j.1540-5885.2005.00180.x

Coccoli, M., Guercio, A., Maresca, P., & Stanganelli, L. (2014). Smarter universities: A vision for the fast changing digital era. *Journal of Visual Languages and Computing*, *25*(6), 1003–1011. https://doi.org/10.1016/j.jvlc.2014.09.007

Colombia, C. de la república de. (2015). *Plan nacional de desarrollo 2014-2018*.

Consejo Privado de Competitividad. (2017). Informe nacional de competitividad 2017-2018, 271. https://doi.org/ISSN 2016-1430

Cruz, M. A. A. da, Rodrigues, J. J. P. C., Al-Muhtadi, J., Korotaev, V., & Albuquerque, V. H. C. (2018). A Reference Model for Internet of Things Middleware. *IEEE Internet of Things Journal*, *PP*(99), 1. https://doi.org/10.1109/JIOT.2018.2796561

Dar, K., Taherkordi, A., Baraki, H., Eliassen, F., & Geihs, K. (2015). A resource oriented integration architecture for the Internet of Things: A business process perspective. *Pervasive and Mobile Computing*. https://doi.org/10.1016/j.pmcj.2014.11.005

Enrique, L., & Silva, O. (1997). La calidad de la universidad. Más allá de toda ambigüedad, 1–14.

European Technology Platform on Smart Systems Integration. (2008). *Internet of Things in 2020*. *Internet of Things in 2020*. https://doi.org/10.1007/978-3-319-49736-5_2

Fernández Martínez, A., & Llorens Largo, F. (2016). *Gobierno de las TI para universidades*. Retrieved from http://tic.crue.org/wp-content/uploads/2016/04/gobierno_de_las_TI_para_universidades.pdf

Gagliardi, D., Schina, L., Sarcinella, M. L., Mangialardi, G., Niglia, F., & Corallo, A. (2017). Information and communication technologies and public participation: interactive maps and value added for citizens. *Government Information Quarterly*, *34*(1), 153–166. https://doi.org/10.1016/j.giq.2016.09.002

Generales, P., & Calidad, C. De. (2006). Modelo de acreditación CNA.

Gómez, J., Jimenez, T., Gumbau, J., & Llorens, F. (2017). *UNIVERSITIC 2017 Análisis de las TIC en las Universidades Españolas*.

Heinemann, C., & Uskov, V. L. (2018). *Smart Universities* (Vol. 70). https://doi.org/10.1007/978-3-319-59454-5

Khan, R., Khan, S. U., Zaheer, R., & Khan, S. (2012). Future internet: The internet of things architecture, possible applications and key challenges. In *Proceedings - 10th International Conference on Frontiers of Information Technology, FIT 2012* (pp. 257–260). https://doi.org/10.1109/FIT.2012.53

Lee, S. K., Bae, M., & Kim, H. (2017). Future of IoT Networks: A Survey. *Applied Sciences*, *7*(10), 1072. https://doi.org/10.3390/app7101072

Lee, W. jun. (2016). When the future technology is now: Paradoxical attitudes of consumer and evaluation of iot service. *International Journal of Smart Home*, *10*(6), 115–126. https://doi.org/10.14257/ijsh.2016.10.6.13

Lin, J., Yu, W., Zhang, N., Yang, X., Zhang, H., & Zhao, W. (2017). A Survey on Internet of Things: Architecture, Enabling Technologies, Security and Privacy, and Applications. *IEEE Internet of Things Journal*, *4*(5), 1125–1142. https://doi.org/10.1109/JIOT.2017.2683200

Liu, R., & Wang, J. (2017). Internet of Things: Application and Prospect. *MATEC Web Conf.*, *100*. Retrieved from https://doi.org/10.1051/matecconf/201710002034

MACIÁ, F. (2017). *SMART UNIVERSITY. Hacia una universidad más abierta*. (Alfaomega & Marcombo, Eds.) (Primera).

Marotta, M. A., Both, C. B., Rochol, J., Granville, L. Z., & Tarouco, L. M. R. (2015). Evaluating management architectures for Internet of Things devices. *IFIP Wireless Days*, *2015–Janua*(January). https://doi.org/10.1109/WD.2014.7020811

Medina-Cárdenas, Y. C., Areniz-Arévalo, Y., & Rico-Bautista, D. (2016). *Modelo estratégico para la gestión tecnológica en la organización: plan táctico de la calidad (ITIL & ISO 20000)*. (Instituto Tecnológico Metropolitano, Ed.). Fondo Editorial ITM. https://doi.org/https://doi.org/10.22430/9789585414006

Medina-Cárdenas, Y., & Rico- Bautista, D. (2012). Mejores prácticas de gestión para la calidad de los servicios en tecnologías de información. *Gerencia Tecnológica Informática*, *11*(29), 47–57. Retrieved from http://revistas.uis.edu.co/index.php/revistagti/article/view/2816

Medina Cárdenas, Y. C., Rico Bautista, D. W., Constanza, Y., & Cárdenas, M. (2009). Modelo de gestión basado en el ciclo de vida del servicio de la Biblioteca de Infraestructura de Tecnologías de Información (ITIL). *Revista Virtual Universidad Católica Del Norte*, (27), 1–21. Retrieved from 148.215.2.11/redalyc-seam/articulo.oa?id=194215432010

Medina, Y. C., & Rico, D. W. (2008). Modelo De Gestión De Servicios Para La Universidad De Pamplona: Itil. *Scientia Et Technica*, *XIV*(39), 314–319. Retrieved from http://www.redalyc.org/articulo.oa?id=84920503056

Mineraud, J., Mazhelis, O., Su, X., & Tarkoma, S. (2016). A gap analysis of Internet-of-Things platforms. *Computer Communications*, *89–90*. https://doi.org/10.1016/j.comcom.2016.03.015

Ministerio de Educación Nacional. (2013). Propuesta metodológica para la distribución de recursos Artículo 87 de la Ley 30 de 1992 Vigencia 2013, 6. Retrieved from http://www.mineducacion.gov.co/sistemasdeinformacion/1735/articles-212353_modelo1.pdf

Ngu, A. H., Gutierrez, M., Metsis, V., Nepal, S., & Sheng, Q. Z. (2017). IoT Middleware: A Survey on Issues and Enabling Technologies. *IEEE Internet of Things Journal*, *4*(1), 1–20. https://doi.org/10.1109/JIOT.2016.2615180

Nitti, M., Atzori, L., & Cvijikj, I. P. (2015). Friendship selection in the social internet of things: Challenges and possible strategies. *IEEE Internet of Things Journal*, *2*(3), 240–247. https://doi.org/10.1109/JIOT.2014.2384734

Porter, M. E. (1985). Competitive Advantage. *Competitive Advantage: Creating and Sustaining Superior Performance*. https://doi.org/10.1182/blood-2005-11-4354

Rautaray, S. S., & Agrawal, A. (2012). Vision based hand gesture recognition for human computer interaction: a survey. *Artificial Intelligence Review*, *43*(1), 1–54. https://doi.org/10.1007/s10462-012-9356-9

Razzaque, M. A., Milojevic-Jevric, M., Palade, A., & Clarke, S. (2016). Middleware for Internet of Things: A Survey. *IEEE Internet of Things Journal*, *3*(1), 70–95. https://doi.org/10.1109/JIOT.2015.2498900

Rico-Bautista, D., Parra-Valencia, J. A., & Guerrero, C. D. (2017). IOT: UNA APROXIMACION DESDE CIUDAD INTELIGENTE A UNIVERSIDAD INTELIGENTE. *Revista Ingenio UFPSO*, *13*(1), 9–20.

Roa, A. (2003). Hacia un modelo de aseguramiento de la calidad en la educación superior en Colombia: estándares básicos y acreditación de excelencia. *Educación Superior, Calidad y Acreditación. Alfa Omega Colombiana, Bogotá*, 101–107.

Rockart, J. F. (1979). Chief executives define their own data needs. *Harvard Business Review*. https://doi.org/Article

Rodríguez Molano, J. I., Montenegro marín, C. E., Cueva Lovelle, J. M., Molano, J., Marin, C., & Cueva, J. (2015). Introducción al Internet de las Cosas. *Redes de Ingeniería*, *6*(7), 53–59. https://doi.org/10.14483/udistrital.jour.redes.2016.1.a04

Rueda-Rueda, J., Manrique, J., & Cabrera Cruz, J. (2017). *Internet de las Cosas en las Instituciones de Educación Superior*.

Sánchez-Torres, B., Rodríguez-Rodríguez, J. A., Rico-Bautista, D. W., & Guerrero, C. D. (2018). Smart Campus: Trends in cybersecurity and future development. *Revista Facultad de Ingeniería*, *27*(47). https://doi.org/10.19053/01211129.v27.n47.2018.7808

Savov, T., Terzieva, V., Todorova, K., & Kademova-Katzarova, P. (2017). Contemporary technology support for education. *CBU International Conference Proceedings*, *5*, 802.

Shvetsova, O. A. (2017). Smart education in high school: New perspectives in global world. In *2017 International Conference "Quality Management,Transport and Information Security, Information Technologies" (IT&QM&IS)* (pp. 688–691). https://doi.org/10.1109/ITMQIS.2017.8085917

Staskeviciute, I., & Neverauskas, B. (2008). The Intelligent University's Conceptual Model. *Inzinerine Ekonomika-Engineering Economics*, (4), 53–58.

Tan, L., Lu Tan, & Neng Wang. (2010). Future internet: The Internet of Things. *2010 3rd InternationalConferenceonAdvancedComputerTheoryandEngineering(ICACTE)*, V5-376-V5-380. https://doi.org/10.1109/ICACTE.2010.5579543

Tandon, A., Shah, V., & Gajjar, S. (2016). Emerging Applications Perspective for Internet of Things. In *Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies - ICTCS '16* (Vol. 04–05–Marc, pp. 1–7). https://doi.org/10.1145/2905055.2905331

Weyrich, M., & Ebert, C. (2016). Reference architectures for the internet of things. *IEEE Software*, *33*(1), 112–116. https://doi.org/10.1109/MS.2016.20

Xiong, N., Liu, R. W., Liang, M., Wu, D., Liu, Z., & Wu, H. (2017). Effective alternating direction optimization methods for sparsity-constrained blind image deblurring. *Sensors (Switzerland)*, *17*(1). https://doi.org/10.3390/s17010174

Zapata-Ros, M. (2018). La universidad inteligente La transición de los LMS a los Sistemas Inteligentes de Aprendizaje en Educación Superior The transition from Learning Management Systems (LMS) to Smart Learning Systems (SLS) in Higher Education. *RED. Revista de Educación a Distancia. Núm*, *57*(10), 31–1. https://doi.org/10.6018/red/57/10

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 27/09/2018
Aceitação/Acceptance: 13/11/2018

Introducción de metodologías participativas en la modernización de la administración de tierras en Colombia acercando la tecnología a la comunidad

Luz Angela Rocha Salamanca[1], Alvaro Enrique Ortiz Dávila [2], Edilberto Suárez [3]

lrocha@udistrital.edu.co, aeortizd@udistrital.edu.co, esuarez@udistrital.edu.co

[1] Universidad Distrital Francisco José de Caldas, Facultad de Ingeniería, Carrera 7 No. 40B-53, 11021, Bogotá, Colombia.

[2] Universidad Distrital Francisco José de Caldas, Facultad de Ingeniería, Carrera 7 No. 40B-53, 11021, Bogotá, Colombia.

[3] Universidad Distrital Francisco José de Caldas, Facultad de Ingeniería, Carrera 7 No. 40B-53, 11021, Bogotá, Colombia.

Pages: 775–786

Resumen: La modernización de la administración de tierras en Colombia es relevante debido a que el gobierno nacional adquirió el compromiso, con la firma del proceso de paz, de formalizar la tenencia de la tierra a corto plazo. Este artículo presenta los resultados del proyecto de investigación cuyo objetivo es probar desde el punto de vista tecnológico la metodología participativa e innovadora denominada "*Fit for Purpose*", para la titulación de tierras, tomando como prototipo los municipios de postconflicto Vista Hermosa y Apartado y mostrar las implicaciones que tiene acercar la tecnología a las áreas rurales en el país. En ese sentido se espera que con la implementación de esta metodología se pueda cumplir la meta del gobierno nacional de titular alrededor de cinco millones de predios en siete años y mostrar como la participación de la comunidad, genera confianza en el proceso, utilizando tecnología en tiempo real.

Palabras-clave: Administración de tierras, Geo-tecnologías, Fit For Purpose, Levantamiento predial participativo

Introducing participatory methodologies in the modernization of Land Administration in Colombia bringing technology to the community

Abstract: The modernization of land administration in Colombia is relevant because the national government acquired the commitment to formalize land tenure in the short term with the signing of the peace process. This article presents the results of the research project, which objective is to prove since the technological point of view the participatory and innovative methodology called "Fit for Purpose", for land titling, using as prototype the post-conflict municipalities Vista Hermosa and Apartado, in addition presents the implications of bringing technology to rural areas in the country. In this sense, it is expected that with the implementation

of this methodology the aim of the national government of entitling around five million properties in seven years can be fulfilled and show how the participation of the community, generates confidence in the process, using technology in real time.

***Keywords**: Land Administration, Fit for Purpose, geo-information tools, parcel data acquisition*

1. Introducción

Tradicionalmente el catastro colombiano se ha desarrollado bajo los procesos de formación catastral, actualización de la Formación y conservación catastral, los cuales enmarcan todas las actividades requeridas para determinar las características físicas, jurídicas, fiscales y económicas de cada predio (IGAC, 2011). Según datos del Instituto Geográfico Agustín Codazzi, entidad encargada del catastro en el país, en el año 2017 estaban registrados 16.531.108 predios de los cuales 4.321.524 son rurales y 12.209584 son urbanos de los 995 municipios administrados por esta Institución; el problema radica en que en las áreas rurales el 7.1 % de municipios no tienen formación catastral y el 67% que lo tienen está desactualizado, lo cual representa un atraso representativo en desarrollo del catastro rural colombiano (Matiz I., 2017). En ese sentido, el Departamento Nacional de Planeación de Colombia a través del documento CONPES 3859, señala que el país debe trabajar para tener un sistema de administración de tierras actualizado, confiable, conectado con el sistema de registro de bienes raíces e integrado con otros sistemas de información (Consejo Nacional de Política Económica y Social, 2015).

De la misma forma el proceso de paz en Colombia desencadenó y evidenció la necesidad de generar proyectos indispensables para el desarrollo del país, especialmente aquellos relacionados con la tenencia de la tierra, donde necesariamente es importante contar con una administración efectiva de la misma, para resolver el problema del porcentaje de personas que no tienen documentada ni poseen ningún documento que permita determinar su relación con la tierra en ámbito rural en el país. Por esta razón el gobierno colombiano determinó como uno de sus objetivos primordiales minimizar el porcentaje de área rural sin títulos de propiedad, especialmente en aquellas zonas que fueron parte del conflicto armado. Para el país cumplir este objetivo es un reto, ya que las metodologías utilizadas en la actualidad para realizar el barrido predial rural son costosas, toman demasiado tiempo y se ajustan a estándares altamente técnicos, debido a los procedimientos definidos para la realidad colombiana, siendo este un país con gran diversidad topográfica. En ese sentido con las metodologías actuales se necesitarán de muchas décadas para llegar a un registro completo de las tierras (Molendijk, 2018).

Igualmente, el programa de las Naciones Unidas para el desarrollo determino 17 objetivos para el desarrollo sostenible (ODS) los cuales son un llamado universal a la adopción de medidas ante el cambio climático, la desigualdad económica, la innovación, el consumo sostenible, la paz y la justicia, entre otras prioridades, temas en los que Colombia al igual que otros países deben trabajar. En ese sentido los gobiernos internacionales especialmente europeos han venido al país para colaborar con el gobierno nacional en el cumplimiento de los ODS, específicamente: el número 1: fin de la pobreza; el 2: cero hambre, el 8: trabajo decente y crecimiento económico y el 17: alianzas (United Nations, 2015).

Es así como actualmente en Colombia se están desarrollando proyectos que pretenden actualizar el catastro colombiano especialmente en las áreas rurales, uno de ellos se conoce como el proyecto de Modernización de la Administración de Tierras en Colombia, el cual está siendo liderado principalmente por el Departamento Federal de Economía y la Secretaria de Estado para Asuntos Económicos del Gobierno Suizo y diferentes entidades nacionales entre las cuales la principal es el IGAC, que busca la implementación de una Infraestructura de Datos para la administración de tierras (Agencia Implementación, 2017). De la misma forma la Agencia internacional de Catastro de Holanda conocida como "*Kadaster*", está trabajando en el proyecto "*Land in peace*" el cual es un proyecto de colaboración entre el gobierno de Colombia y el de los Países Bajos para contribuir a la creación de una Administración de Tierras efectiva en el país a través de la introducción de metodologías innovadoras y participativas (Kadaster, 2018).

En todos estos procesos de introducción de nuevas metodologías, la academia juega un papel relevante, pues se requiere desarrollar investigación para la aplicabilidad de las mismas en el país. Es por esta razón el grupo de investigación NIDE (núcleo de investigación en datos espaciales) de la Universidad Distrital Francisco José de Caldas, se ha unido al proyecto liderado por la Agencia internacional de Catastro de Holanda "Kadaster" para la prueba e implementación de la Metodología "*Fit for Purpose*" (FFP) (traducido como ajustado al propósito) para la modernización de la administración de tierras en Colombia, utilizando herramientas geo-tecnológicas, en el marco de un Memorando de Entendimiento (MoU) entre las dos instituciones. Para tal fin se creó el proyecto de investigación "Implementación de la metodología de Administración de tierras "*Fit for Purpose*" (FFP) en la realidad colombiana, desde la teoría hasta la práctica" cuyo objetivo en su primera fase, es mostrar la viabilidad técnica de la implementación de la metodología FFP en el país. La metodología utilizada se basó en la participación en los pre-pilotos desarrollados en los municipios de posconflicto: Vista Hermosa (vereda Termales) en el departamento del Meta y Apartadó (vereda los Mandarinos) en el Departamento de Antioquia, donde profesores y estudiantes fueron los encargados de la recopilación de los datos en campo, posteriormente se realizó el análisis detallado de los resultados de los datos y de la tecnología geomática utilizada para tal fin y como ésta brinda el soporte para la realización de todas las actividades en tiempo real.

La investigación realizada tiene gran relevancia ya que a partir de los resultados se determina la viabilidad de la implementación de esta metodología en el contexto colombiano, teniendo en cuenta que FFP está siendo probada en diferentes países en el mundo, sin que se presenten casos en los cuales se haya implementado masivamente en un territorio (Lemmen, 2015).

Todas las actividades desarrolladas en esta primera fase del proyecto se presentan en los siguientes apartados de este artículo, comenzado por los antecedentes, una descripción detallada de la Metodología y las conclusiones del proyecto de investigación.

2. Antedecentes

La Administración de tierras se conoce principalmente como el proceso de recolectar, almacenar y difundir información relacionada con la tenencia, el valor y el uso de la tierra

utilizando políticas de gestión en el territorio, cuyos procesos son administrados por el gobierno y desarrollados por agencias públicas o privadas (Williamson, Enermark, & Wallace, 2010). La administración de la tierra es el proceso de regulación de la tierra, el derecho de propiedad, el uso, conservación, valor de la tierra (venta, arrendamiento, etc.) y la resolución de conflictos sobre la propiedad y el uso de la tierra (Dale & Mclaughlin, 1999).

Por otro lado, y como apoyo a los procesos de administración de tierras el concepto de "*Fit For Purpose*" (FFP) fue desarrollado para la construcción de sistemas de administración de tierras sostenibles, ajustadas a las condiciones de los países en vía de desarrollo. Esta propuesta surge como resultado de la cooperación entre el Banco Mundial y la Federación Internacional de Geómetras (FIG) como respuesta a los Objetivos de Desarrollo Sostenible (ODS) planteados por el Programa de las Naciones Unidas para el Desarrollo (PNUD) en 2016. La metodología "*Fit For Purpose"* se encuentra en una etapa de implementación y evaluación para estimar su efectividad y comprobar las bondades para las que fue concebido. FFP está siendo implementado en países como Rwanda, Kenya y Ethiopia, países que también están enfrentando la construcción de sistemas de administración de tierras (Lemmen, 2015).

Los componentes básicos del concepto FFP son: (Enemark, Clifford, Lemmen, & McLaren, 2014).

a. El uso de tecnologías modernas asequibles para construir un marco espacial, que permitan mostrar cómo es la ocupación y el uso de la tierra. La escala y la precisión pueden variar de acuerdo a la topografía y a otras variables dependiendo del terreno.
b. La utilización de un enfoque participativo, teniendo como base el marco espacial, para Identificar y registrar los diferentes derechos legales y sociales de tenencia, relacionados con la ocupación y el uso de la tierra.
c. La adopción de un marco legal que se adapte a la flexibilidad necesaria para implementar FFP. Este marco se puede establecer con anterioridad o se puede desarrollar incrementalmente.

Igualmente, el concepto FFP involucra tres componentes centrales: el marco espacial, el marco legal y el marco institucional. Cada uno de estos componentes posee la flexibilidad pertinente para satisfacer las necesidades reales de la región donde se puede implementar, como respuesta a las necesidades sociales y a los recursos financieros disponibles (Enemark, 2016). En este concepto se plantea una implementación flexible y rápida de la captura de datos del predio, utilizando tecnologías de la información geográfica, con la participación activa de la comunidad y del propietario (Enemark, Clifford, Lemmen, & McLaren, 2014).

El marco espacial incluye la información espacial a escala grande, que muestra la forma en que se divide el terreno en unidades espaciales (como parcelas o predios) para uso específico y ocupación; proporciona la información básica para realizar las funciones de la administración de tierras tales como: registro y gestión de tenencia legal y social; evaluación de la tierra y el valor de la propiedad, identificación y gestión del uso actual de la tierra; planificación para el uso futuro de la tierra y administración y protección de recursos naturales. El marco legal señala que existe un rango de posibles formas de tenencia de la tierra (continuum), que es continuo, por lo cual implica diferentes tipos

de derechos, grados de seguridad y responsabilidad y en cada uno de ellos se puede dar igualmente, diferentes grados de cumplimiento. De esta forma en cada paso del proceso se puede formalizar la tenencia de la tierra. Lo esencial es que en el marco legal se reconozcan los derechos sobre la tierra de una forma equitativa, con base en los derechos humanos. El marco institucional se refiere a la creación de sistemas de administración de tierras flexibles y de la capacidad de los gobiernos para lograr este objetivo de una forma eficiente que beneficie a la sociedad. (Enemark, Clifford, Lemmen, & McLaren, 2015)

El modelo utilizado para la captura de los datos se basó en el modelo LADMcol, que es el perfil para Colombia de la norma ISO 19152:2012, (Land Administration Domain Model) la cual define el modelo de administración de tierras en seis clases: (Lemmen, 2015).

- Party (Interesados)
- RRR (derechos, restricciones y responsabilidades)
- BA Unit (Objetos administrativos)
- Spatial Unit (Objetos espaciales)
- Surveying (levantamientos de campo)
- Source (Fuentes de datos)

La parte fundamental del modelo LADMcol consiste en definir un lenguaje unificado y una estructura de relaciones entre los diferentes elementos que componen los objetos considerados en la administración de tierras, ajustadas a las necesidades propias de cada institución (Agencia Implementación bsf swissphoto, 2016).

2.1.Área de estudio

Las zonas escogidas por Kadaster para la realización de los pre-pilotos fueron los Municipios de postconflicto Vista Hermosa en el Departamento del Meta y Apartadó en el Departamento de Antioquía y de cada uno de ellos se seleccionó una vereda, por ser representativos en cuanto a la problemática de la restitución de Tierras en Colombia (Figura 1).

Figura 1 – Localización geográfica de los Municipios escogidos para la realización de los pre-pilotos FFP

El Municipio de Vista Hermosa se encuentra ubicado a 235 Km. aproximadamente, al sur oriente de Bogotá D. C. y al sur occidente del departamento del Meta, a una distancia de la capital Villavicencio de 145 Km. Tiene una superficie de 4.084 Km2, de los cuales la extensión de área urbana es de 1,023 Km^2 y la extensión del área rural de 3,726 Km^2. Está situado a una altura sobre el nivel del mar de 460 metros y su temperatura media es de 28°C. Cuenta con una población de 26347 habitantes y una densidad de población de 5,36 hab/km^2. El Municipio cuenta con una zona plana perteneciente a los llanos orientales de la Orinoquia y otra zona Montañosa perteneciente a la serranía de la Macarena y dentro de su jurisdicción se encuentra en el Parque Nacional Natural Sierra de La Macarena. Según la Unidad para la Atención y Reparación Integral a las Victimas se reportan 32.660 personas que salieron por desplazamiento y 9.534 que llegaron provenientes de otros lugares. Para realizar el pre-piloto en el Municipio de Vista Hermosa se escogió la Vereda Termales la cual tiene un área aproximada de 1660 hectáreas (Alcaldía de Vista Hermosa, 2018).

El Municipio de Apartadó está ubicado al noreste del Departamento de Antioquia, hace parte del denominado Urabá Antioqueño. Tiene una extensión de 600 Km^2, de los cuales 592.35 km^2 pertenecen al área rural. Para realizar el pre-piloto en Apartadó se escogió la Vereda Los Mandarinos, la cual tiene un área aproximada de 237 hectáreas (Alcaldía de Apartadó, 2016).

3. Metodología

La metodología del proyecto se basó básicamente en el desarrollo de las siguientes actividades: Adquisición de datos, post-procesamiento de los datos e inspección pública, las cuales ayudan responder a la pregunta de investigación ¿Cómo implementar la metodología "*Fit for Purpose*", en las condiciones propias de los Municipios de Colombia, para garantizar la tenencia de la tierra en los en el país en los próximos 10 años, utilizando para tal propósito herramientas geo-tecnológicas?

Los pre-pilotos para probar la metodología fueron liderados por profesores y estudiantes de Ingeniería Catastral y Geodesia y la Maestría de Ciencias de la Información y las comunicaciones con énfasis en Geomática de la Universidad Distrital Francisco José de Caldas de Bogotá (Colombia) y expertos en el tema de FFP del Catastro internacional de Holanda y de la Universidad de Twente (Holanda).

3.1. Adquisición de datos

El proceso de adquisición de datos prediales es la base de la administración de tierras, por esta razón es la más costosa y la que consume más tiempo. En ese sentido la metodología "*Fit For Purpose"*, propone que esta se realice de una forma integrada, colaborativa y asequible apoyándose en la tecnología existente para tal fin. Los métodos de adquisición de datos utilizando FFP se pueden definir de acuerdo al propósito tal y como se presenta en la figura 2.

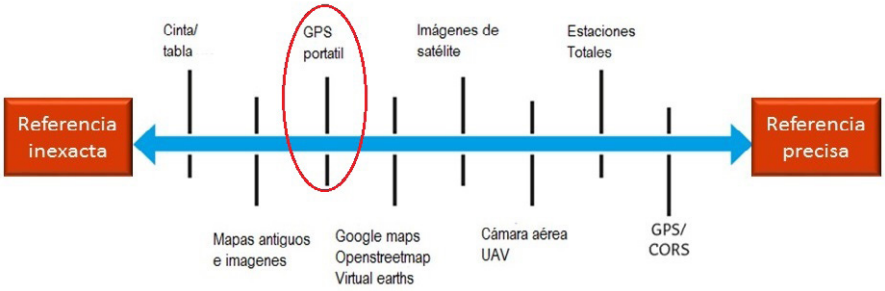


Figura 2 – Métodos de adquisición de datos en FFP. Tomado y ajustado de (Lemmen, 2015)

En este proyecto se aplicó la metodología basada en la utilización de GPS portátil (Trimble R1 y R2), un dispositivo móvil (teléfono inteligente y tableta) y conexión a internet; se contó con el apoyo logístico y económico de Kadaster. Este método fue escogido teniendo en cuenta la falta de información de los municipios en cuanto a datos catastrales y a imágenes de satélite de buena resolución de la zona.

Previamente a la visita al predio, se realizó una reunión con la comunidad, el cual es el primer paso para obtener buenos resultados en la labor a desarrollar. Allí se informó a la comunidad de manera detallada cuál es objetivo del trabajo a realizar y cuál es el papel de los diferentes agentes en cada paso del proceso utilizando la metodología FFP. Es importante realizar esta reunión, pues permite el acercamiento entre las diferentes instituciones del Estado y la comunidad generando, así un carácter de confiabilidad al proceso.

El flujo de trabajo para la captura de los datos se fundamenta en que es integrado y colaborativo y se realiza todo en tiempo real, tal y como se muestra en la figura 3.

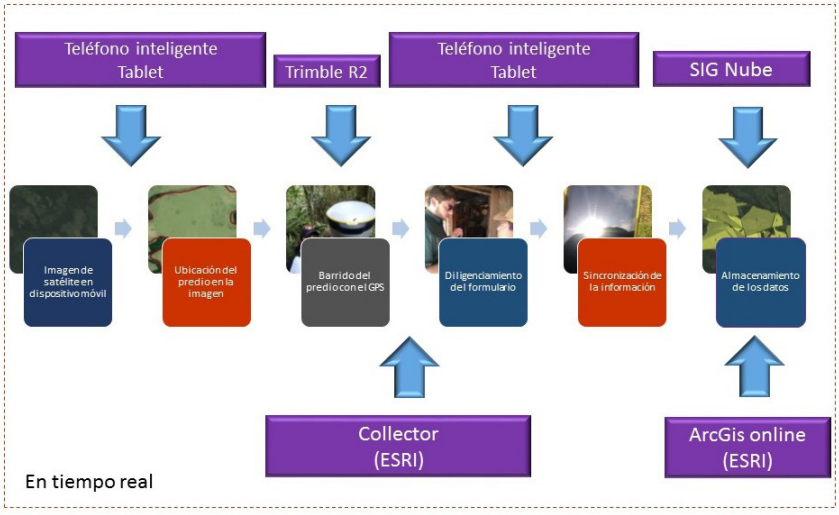


Figura 3 – Flujo de trabajo para la adquisición de datos en campo mediante FFP

En la visita al predio, se le da una charla inicial al propietario o encargado del mismo, explicando que se realizará un levantamiento de los linderos de su terreno, de tal forma que él indique, en forma ordenada, los puntos que permitan delimitar la parcela o predio, se define el recorrido y el punto para iniciar el levantamiento.

El procedimiento consiste entonces, en que el propietario o poseedor coloca instrumento receptor GPS en los puntos que él considera definen los linderos del predio, se almacenan las coordenadas que el GPS reporta en un dispositivo móvil, el cual cuenta con la aplicación Collector (de la firma ESRI), sobre una imagen base georreferenciada. De esta forma se hace el recorrido del predio, tomando los puntos que señala el propietario bajo la coordinación y asesoría de un profesional con conocimiento en catastro (para este proyecto un profesor). En esta parte del proceso se evidencia como la tecnología puede ser utilizada de una forma sencilla por el propietario y que todos los elementos técnicos que se involucran en este procedimiento son transparentes para ellos.

Para el almacenamiento de los datos, es un estudiante de ingeniería catastral y Geodesia quien maneja el aplicativo de recolección de datos, bajo la coordinación del profesor encargado del grupo. Al finalizar el proceso de la captura de los puntos que conforman los linderos, el propietario debe aprobar la forma del polígono capturado para su terreno, esto garantiza confiabilidad en el levantamiento ya que es el mismo propietario quien ha determinado sus propios linderos (Grupo de trabajo FFP UD, 2018).

Posteriormente el estudiante hace el levantamiento de la información jurídica del predio, a partir de una encuesta al propietario, donde se le pregunta acerca de si posee algún documento que identifique el derecho sobre la tierra, si existe algún otro propietario, el uso que tiene la tierra y los datos de la cédula de ciudadanía; en algunos casos se puede anotar la información de la tradición del inmueble, si los actuales poseedores la tienen. Toda esta información es almacenada en el dispositivo móvil, teniendo en cuenta el Modelo LADMcol. Es así como se toman los datos importantes referentes a la relación entre los predios y sus propietarios o poseedores, a quienes se les solicita además una foto de la cédula y se toma una fotografía del rostro de cada persona. En este proceso no sólo se recopila la información relacionada con la parcela y su propietario, sino que también se identificaban puntos de interés y objetos de referencia como: vías, cuerpos de agua, accidentes geográficos y sitios de interés, entre otros, que servirán para determinar algunos puntos de los linderos. Finalmente, la información se almacena en la nube, en ArcGis online (Grupo de trabajo FFP- UD, 2018).

En el grupo de trabajo en campo igualmente participan un profesional especialista en la metodología FFP de Kadaster y de la Universidad de Twente de Holanda. También se tuvo el acompañamiento de un representante de las organizaciones gubernamentales involucradas en el proceso de titulación de tierras, como son el Instituto Geográfico Agustín Codazzi, la Agencia Nacional de Tierras y la Superintendencia de Notariado y Registro.

En el primer levantamiento en la vereda Termales de Vista Hermosa se probaron los equipos R1 y R2 (Trimble), mostrando un mejor desempeño en la zona montañosa el equipo Trimble R2, ya que no presentó problemas en su funcionamiento y la precisión obtenida se mantuvo dentro del rango requerido en el levantamiento, que oscilaba en 50 cm (en zonas despejadas) y valores entre 90 cm a 1.5 m. en zonas muy cubiertas

por vegetación, dependiendo de la densidad de la misma. En cuanto a los dispositivos móviles, la ventaja de las tabletas fue la visualización de los datos, pero los teléfonos inteligentes mostraron mejores resultados ya que son más portables y tienen mejor alcance del bluetooh.

3.2. Post procesamiento de los datos

Una vez se tienen los datos capturados en campo se realiza un post- procesamiento totalmente automático, el cual fue realizado por un profesor experto en Bases de Datos espaciales. Los puntos almacenados son de tres clases: puntos de terreno que son aquellos que define el propietario en la delimitación del lindero; los puntos de colindancia o de anclaje, que son aquellos que son compartidos por dos o más predios y los puntos de referencia, los cuales señalan algún elemento del paisaje que hace parte del lindero como por ejemplo una corriente de agua (figura 4).

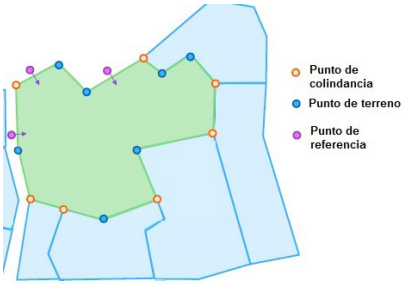


Figura 4 – Tipos de puntos recolectados en campo. Fuente (University of Twente, 2018)

De los datos se verifica la consistencia topológica y se hacen los ajustes necesarios para luego ser almacenados en la Base de Datos espacial creada en PostgreSQL-PosGIS. (Figura 5). Para la edición y procesamiento de los datos se crearon algoritmos y scripts con el fin de realizar todo el proceso de corrección de la topología automáticamente y para la generación de los mapas prediales, se utilizó Quantum GIS (Figura 5). Esta información luego se presenta a los propietarios, quienes validarán y aprobarán los linderos de su predio en la actividad denominada inspección pública.

3.3. Inspección Pública

La inspección pública es el proceso por el cual se reúne a la comunidad y es allí donde el propietario aprueba sus linderos y los de sus colindantes, a partir de la información recolectada por el mismo, en el trabajo de campo, la cual está almacenada en la Base de Datos espacial (United Nations Human Settlements Programme UN-Habitat, 2016). Es así como una vez analizada la información por cada uno de los propietarios, se coloca el lindero en color verde si está totalmente aprobado por todas las partes, en color azul si es aprobado solo por una de las partes, en color naranja si está parcialmente aprobado y color rojo cuando no hay acuerdo por ninguna de las partes, todo directamente en la Base de Datos. Cuando los linderos son aprobados por todas las partes los propietarios firman digitalmente, lo que se conoce como el acta de colindancia, utilizando una tableta.

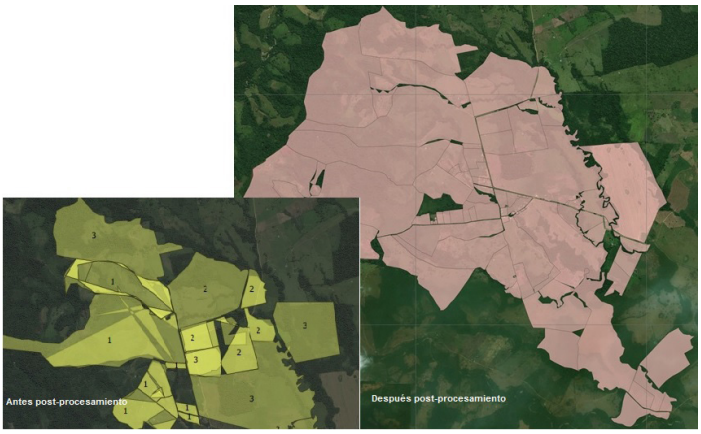


Figura 5 – Datos de los predios de Vista Hermosa antes y después del procesamiento automático

Posteriormente con las actas de colindancia firmadas se procede a la generación de los títulos de propiedad, función que es desarrollada misionalmente por la Agencia Nacional de Tierras en Colombia, este proceso no hace parte del proyecto de investigación.

Como resultado del proyecto se pudo determinar que utilizando esta metodología se levantaron un total de 982 hectáreas en 10 días en la vereda Termales de vista Hermosa, lo cual indica lo eficiente de la metodología. Igualmente se pudo evidenciar que el número de predios registrados en el catastro nacional es menor al número de predios levantados en la vereda Termales como resultado de los pre-pilotos. Esto representa que los cálculos del gobierno nacional en cuanto al número de predios a titular son sustancialmente menores a los que existen realmente (University of Twente, 2018).

4. Trabajos futuros

El proyecto de implementación de la metodología FFP en Colombia está en su etapa preliminar, por esta razón se proyecta realizar una segunda fase en la cual se determine el costo real que tiene el proyecto, para poder determinar si realmente cumple con el objetivo de ser asequible a bajo costo.

Por otro lado, se explorarán otras metodologías como son el uso de imágenes de radar que permitan hacer una descripción de la zona de trabajo y determinar elementos relevantes como los drenajes y las vías que muchas veces hacen parte de los linderos de los predios y tomar esta información para ser complementada en los levantamientos en campo.

Finalmente se propone aplicar la metodología FFP en zonas del país que tengan características diferentes a las de los municipios de postconflicto, es decir que tengan una economía más consolidada, para verificar el funcionamiento de dicha metodología en estas áreas del país.

5. Conclusiones

Como resultado del proyecto de investigación desde el punto de vista técnico, si es posible implementar la metodología FFP para la administración de tierras en el país y se cumple con el principio de ser colaborativa, participativa e innovadora, debido a que se tiene una participación activa de la comunidad y especialmente de los propietarios quienes son los que definen sus propios linderos. Adicionalmente permite agilizar los procesos de recolección de información de los predios rurales y por ende del proceso de restitución de tierras, el cual beneficia grandemente a los propietarios y poseedores quienes con los títulos podrán acceder a diferentes beneficios por parte del gobierno como préstamos para mejorar las condiciones del terreno.

El uso de la tecnología GPS y de dispositivos móviles por la comunidad de los municipios prototipos, permitió evidenciar que es posible utilizar metodologías participativas, desde que las aplicaciones sean sencillas, lo que genera gran confiabilidad del proceso de recolección de los datos de los linderos de los predios. Sin embargo, se debe tener cuidado en la recolección de los datos jurídicos, con el fin de que la Base de Datos sea consistente y para que la manipulación en el post-proceso sea mínima.

Teniendo en cuenta que el proceso de adquisición de datos se realiza en tiempo real, se minimizan costos ya que los datos pasan directamente de la aplicación en el dispositivo móvil a la Base de Datos espacial, evitando así manipulaciones y reduciendo el tiempo en los levantamientos.

Por otro lado, el desarrollo del enfoque FFP, impacta positivamente a los municipios prototipos, atendiendo así a una necesidad actual del país por conocer los recursos y sobre todo para ayudar a las partes interesadas en garantizar un derecho sobre la tierra, proponiendo otros métodos para la construcción de un mejor sistema de administración de tierras que se adapte a las necesidades presentes del país apoyados en el uso de las geo-tecnologías.

Finalmente, desde el punto de vista académico es una gran oportunidad para la Universidad Distrital Francisco José de Caldas participar en este proyecto conjuntamente con la Agencia de Catastro internacional de Holanda, ya que permitió aprender partiendo del concepto hasta la práctica la metodología FFP en la realidad colombiana, siendo así pioneros en el desarrollo de esta metodología y en general en los procesos de cambio de la administración de tierras en el país. Adicionalmente tanto los estudiantes como profesores que hacen parte del proyecto tuvieron la posibilidad de visitar 2 veredas de los Municipios escogidos para este proyecto (Vista Hermosa y Apartadó), una por cada Municipio y realizar el trabajo directamente en campo utilizando diferentes equipos tecnológicos para este fin y tratando directamente con los problemas de la comunidad, situación que enriquece el conocimiento y el trabajo en grupo. En un futuro próximo los profesores y estudiantes serán los multiplicadores de este conocimiento en el país.

Referencias

Agencia Implementación bsf swissphoto. (2016, marzo 23). Versión 0.9 del Perfil Colombiano de la ISO19152 - Modelo de Datos para el Dominio de la Administración de Tierras

Alcaldía de Apartadó. (2016). *Municipio de Apartadó*. Obtenido de http://www.apartado-antioquia.gov.co/MiMunicipio/Paginas/Informacion-del-Municipio.aspx

Alcaldía de Vista Hermosa. (13 de Marzo de 2018). *Municipio de Vista Hermosa*. Recuperado el 30 de Marzo de 2018, de http://www.vistahermosa-meta.gov.co/

Consejo Nacional de Política Económica y Social. (2015). *Politica para la adopción e implementación de un catastro multipróposito rural-urbano*. Obtenido de https://colaboracion.dnp.gov.co/CDT/Conpes/Econ%C3%B3micos/3859.pdf

Dale, & Mclaughlin. (1999). Land Administration.

Enemark, S., Clifford, K., Lemmen, C., & McLaren, R. (2014). Fit-for-purpose land administration. Fig pub 60

Enemark, S., McLaren, R., Lemmen, C., Antonio, D., & Gitau, J. (2016, marzo). Scaling up responsible land governance: guiding principles for building fit-for-purpose land administration systems in developing countries. Recuperado el 16 de enero de 2018

Grupo de trabajo FFP UD. (2018). *Informe implementación metodología "FIT FOR PURPOSE" municipios Vista Hermosa y Apartadó*. Bogotá.

IGAC. (2011). Resolución 070. Por la cual se reglamenta técnicamente la formación catastral, la actualización de la formación catastral y la conservación catastral

Kadaster. (2018). *Land in Peace*. Obtenido de https://tierraenpaz.net/home

Lemmen, C. (Mayo de 2015). New Practises in Land Administration. Bogotá, Colombia.

Matiz I. (2017). El catastro multipropósito en Colombia. *FIG Commission 7 and 9. Annual Conference and meeting*. Cartagena.

Molendijk, M. (2018). Land and Peace in Colombia: FFP Methodology for Field Data Collection and Data Handling. *2018 World Bank Conference On Land And Poverty*. Washington.

United Nations. (2015). *Sustainable development goals*. Obtenido de http://www.undp.org/content/undp/es/home/sustainable-development-goals/resources.html

United Nations Human Settlements Programme UN-Habitat. (2016). *Fit-For-Purpose Land Administration Guiding Guiding Principles for Country Implementation*. Netherlands.

University of Twente. (Mayo de 2018). Workshop FFP prepilots 2018 Fit For Purpose Land Administration. Bogotá, Colombia.

Williamson, Enermark, & Wallace. (2010). Land Administration for sustainable development

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 21/10/2018
Aceitação/Acceptance: 11/12/2018

Pensamiento Computacional para la mejora de las capacidades en Geometría Espacial – una experiencia constructivista con Sphero

Jason Cahuana[1], Elizabeth Vidal[1]

jcahuana@ulasalle.edu.pe, evidal@ulasalle.edu.pe

[1] Universidad La Salle, Arequipa, Perú.

Pages: 787–794

Resumen: El Programa "Desarrollando el Pensamiento Computacional" es una iniciativa de la Universidad La Salle, que busca acercar la Computación a niñas y niños en edad escolar como una herramienta para fortalecer otras áreas de estudio como son las matemáticas. En este artículo presentamos la experiencia de un taller, en el que los alumnos aprenden una forma distinta de solucionar problemas de geometría mediante el uso del robot Sphero. Como primer resultado se ha visto un acercamiento no sólo a la Computación sino a la resolución de problemas en Geometría de manera lúdica e interactiva, a través de la exploración y la resolución de problemas haciendo uso de los fundamentos del pensamiento computacional.

Palabras-clave: pensamiento computacional, constructivismo, resolución de problemas, Sphero.

Computational thinking for the improvement of capacities in Spatial Geometry - a constructivist experience with Sphero

Abstract: The Program "Developing Computational Thinking" is an initiative of the Universidad La Salle, which seeks to bring computing to school as a tool to strengthen other areas of study such as mathematics. In this article we present the experience of a workshop, in which students learn a different way of solving geometry problems by using the Sphero robot. As a first result we have seen an approach not only to Computing but to the resolution of problems in Geometry in a ludic and interactive way, through the exploration and the resolution of problems making use of the fundamentals of computational thinking.

Keywords*:* computational thinking, constructivism, problem solving, Sphero

1. Introducción

Las habilidades espaciales son un componente importante en la inteligencia humana (Gardner, 2011). El término espacial cubre cinco componentes: percepción espacial, visualización espacial, rotaciones mentales, relaciones espaciales y orientación espacial

(Maier, 1996). El objetivo del estudio de la geometría es mejorar esas habilidades espaciales. El estudio de conceptos de carácter espacial (3D) lleva en sí misma a la exploración de problemas espaciales que son difíciles de captar cuando se expresan en 2D (Woods et al., 2004). El no poder ver objetos en 3D hace que los estudiantes tengan que calcular y construir dichos objetos con métodos tradicionales: en papel (2D). Esto hace que los problemas sean poco comprendidos y que se perciba la geometría como una materia complicada y hasta muchas veces aburrida produciendo un clima de desmotivación que se ve reflejado en los resultados de aprendizaje.

El Pensamiento Computacional toma un enfoque para resolver problemas basado en los principios de la computación (Wing, 2006). Según lo expuesto en el trabajo de Gadanidis (Gadanidis et al. , 2017) el Pensamiento Computacional tiende a ser visto como una materia en si misma, y no como una herramienta para integrar e enriquecer otras áreas, como por ejemplo las matemáticas. Investigaciones recientes han demostrado la posibilidad de la integración del Pensamiento Computacional y la enseñanza de Matemáticas (Gadanidis et al., 2017) , (Yaday et al., 2014) .Si bien existen varios entornos de programación que permiten el desarrollo de pensamiento computacional como son Scratch (Scratch, 2018), Alice (Alice, 2018) o la inicitiava Code.org (Code.org, 2018), nuestro trabajo en busca de soportar las características 3D que son necesarias para el desarrollo de las habilidades espaciales hace uso del robot Sphero (Sphero, 2018).

El resto del artículo está organizado de la siguiente manera. En la sección II se presenta las principales características del Pensamiento Computacional. En la Sección III se explican las características de Sphero. En la Sección IV se presenta el desarrollo de la Experiencia. Finalmente en la sección V se presentan nuestras conclusiones.

2. Pensamiento Computacional

2.1. Definición

El Pensamiento Computacional (PC) es un método para resolver problemas usando tecnología y está inspirado en el conjunto de competencias y habilidades que un profesional utiliza cuando crea una solución o aplicación computacional. PC una forma de resolver problemas integrando las tecnologías digitales con las ideas de las personas (Wing, 2006). Notemos que esto no reemplaza la creatividad, el razonamiento y el pensamiento crítico, sino que refuerza estas competencias dándole a las personas nuevas formas de organizar un problema con la ayuda de la computadora o de forma más general con algún tipo de tecnología.

2.2. Resolución de Problemas

Para formular problemas de manera que puedan ser resueltos usando computadoras se consideran los siguientes pasos (Lee, 2014): generar una abstracción del problema a resolver, organizar los datos de manera lógica, identificar y analizar posibles soluciones, especificar una solución mediante una serie de pasos ordenados (algoritmo), codificar la solución en un lenguaje que entienda la computadora (programa), generalizar y transferir esta solución a otros problemas donde sea aplicable.

3. Sphero

3.1. Descripción

Sphero (Figura 1) es un dispositivo robótico creado por la compañía Orbotix en 2012. Sphero posee un hardware controlable mediante software (modificación de parámetros en los sensores, límite de velocidad del motor, frecuencia de recepción bluetooth, etc.), además de controlar sus movimientos (Sphero, 2018).

Figura 1 – Sphero

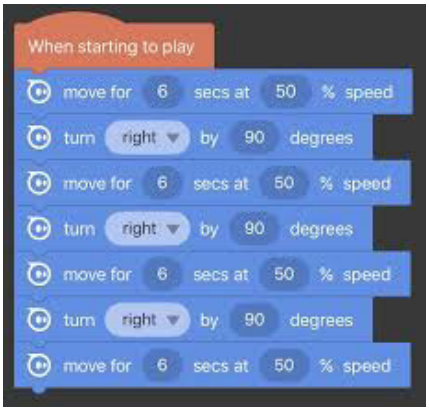


Figura 2 –Programación en Bloques

Sphero se controla mediante un dispositivo Tablet o un teléfono inteligente, vía Bluetooth a un radio de hasta treinta metros. Tanto la Tablet como el teléfono inteligente pueden tener sistemas operativos Android o iOS. La programación en Sphero se hace mediante bloques (Figura 2). El estilo de programación por bloques no es nuevo, este es utilizado en programas como Alice, Scratch y Code.org

3.2. Características Físicas

Entre las principales características de Sphero que lo hacen atractivo para su uso en el enseñanza se destaca: tiene un acelerómetro similar al que pueda tener cualquier teléfono inteligente, apoyándose en sus tres ejes de coordenadas: X, Y, Z; dispone de un giroscopio; posee un LED el cual dará información al usuario acerca del punto en el que se encuentra el giroscopio para realizar movimientos; incluye un sensor anticolisiones y tiene un sensor de localización, mediante el cual, se puede conocer su posición en todo momento, enviando estos datos al dispositivo que lo está controlando.

3.3. Elementos

Sphero trae consigo una cinta métrica que permite realizar mediciones en el suelo o terreno de trabajo que luego pueden ser programadas para comprobar resultados (Figura 3) o realizar circuitos que Sphero podría recorrer (Figura 4).

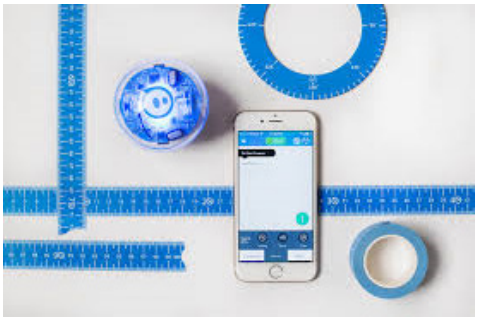


Figura 3 – Cinta Métrica

Figura 4 – Circuito

4. La Experiencia

4.1. Fundamentación Pedagógica

El constructivismo propone un paradigma donde el proceso de enseñanza se percibe y se lleva a cabo como un proceso dinámico, participativo e interactivo del estudiante. El constructivismo en pedagogía se aplica como concepto didáctico en la enseñanza orientada a la acción (Schunk, 2012). De acuerdo a lo expresado por Jean Piaget, represente del Constructivismo (Carretero, 1997) es importante mantener activos a los estudiantes. Los estudiantes necesitan ambientes estimulantes que les permitan explorar de forma activa y que incluyan actividades prácticas. Este tipo de enseñanza facilita la construcción activa del conocimiento Tomando también la teoría de Vygotsky, otro representante del Constructivismo (Schunk, 2012), se destaca la interacción de los factores interpersonales. Al interactuar con las personas en el entorno, como cuando se trabaja en grupos de aprendizaje o en colaboración, se estimulan procesos del desarrollo y se fomenta el crecimiento cognoscitivo.

En resumen podemos resumir nuestra propuesta fundamentada en un enfoque constructivista orientada a la acción, a la construcción activa del conocimiento y con interacción en el entorno.

4.2. Metodología

Los estudios exploratorios sirven para familiarizarnos con fenómenos relativamente desconocidos, obtener información sobre la posibilidad de llevar a cabo una investigación más completa respecto de un contexto particular, indagar nuevos problemas, identificar conceptos o variables promisorias, establecer prioridades para investigaciones futuras (Hernández , Fernández & Baptista, 2006). Dentro de este marco podemos decir que nuestro trabajo es exploratorio y se trata de una primera aproximación para poder identificar variables a estudiar a futuro sobre el impacto del uso de Sphero y su aplicación bajo el enfoque Constructivista.

Para llevar a cabo esta primera aproximación se llevó a cabo un Taller, el cual fue programado en una sesión de dos horas. La estructura del taller se dividió en cinco actividades considerando el brindar a los estudiantes un ambiente estimulante para

el aprendizaje que les permita explorar en forma activa mediante el uso del robot Sphero los nuevos conceptos y poder realizar discusiones en grupo para poder diseñar diferentes soluciones a un problema planteado. Las etapas del Taller fueron divididas en cinco actividades en un formato Secuencial pero con un componente iterativo tal como se muestra en la Figura 5.

Figura 5 – Actividades del Taller

Actividad 1: Explicación del problema a resolver: a los estudiantes se les da un problema de acuerdo a su grado de estudio. El problema debe ser un concepto nuevo programado en su plan de estudios. El profesor explica el problema.

Actividad 2: Explicación del uso de: se da una breve explicación sobre como usar el robot Sphero con el celular y el cómo se realiza la programación en bloques indicando los comandos básicos. Dependiendo del año de estudios y el plan de estudios se pueden seleccionar temas más avanzados (como por ejemplo el uso del acelerador para enseñar Física)

Actividad 3: Resolución del Problema: En esta actividad se organiza a los estudiantes en equipos y se les pide: (i) generar una abstracción del problema a resolver, (ii) organizar los datos de manera lógica, (iii) identificar y analizar posibles soluciones y (iv) especificar una solución mediante una serie de pasos ordenados (algoritmo) , tal como se ha descrito en la Sección 2

Actividad 4: Programación del Problema con Sphero: los estudiantes programan su solución diseñada en la actividad (3). Si el comportamiento de Sphero no refleja el diseño propuesto, los estudiantes vuelven a la actividad (3) para discutir sobre el porque falló la primera solución y proponer un nuevo diseño. Las actividades (3) y (4) podrían darse tantas veces como fuera necesario hasta lograr la solución esperada.

Actividad 5: Comentarios y Preguntas: esta actividad es de cierre para obtener información sobre la experiencia desde la perspectiva de los estudiantes en el proceso de enseñanza/aprendizaje.

Es importante resaltar que las etapas (3) y (4) están enmarcadas dentro de los pasos de resolución de problemas planteados por el Pensamiento Computacional expuesto en la Sección II.

4.3 Contexto

La experiencia se llevó a cabo entre Octubre y Diciembre del 2017 en dos colegios de la ciudad de Arequipa. El primero es el Colegio Particular Tersa de Calcuta con alumnos del sexto de primaria. El segundo es el Colegio Particular San Gerónimo con alumnos de cuarto de secundaria.

En las Figuras 6 , 7, 8, y 9 se muestran las primeras cuatro etapas en el Colegio Particular Teresa de Calcula. A ellos se les dejó como problema resolver el Teorema de Pitágoras.

Figura 6 – Actividad 1 - Explicación del Problema a Resolver

Figura 7 – Actividad 2- Explicación del Uso Sphero

Figura 8 – Actividad 3 - Resolución del Problema

Figura 9 – Actividad 4 - Programación del Problema: Demostración de Pitágoras con Sphero

5. Conclusiones

Para esta primera experiencia se ha realizado solamente un tipo de investigación exploratoria para medir el impacto del uso de Sphero en la enseñanza de un tema dentro de la Geometría Espacial. De acuerdo a las observaciones realizadas por el equipo participante (docente universitario y profesor en aula) los alumnos mostraron un alto grado de atención, participación y motivación durante el desarrollo del taller. Sphero, de acuerdo al enfoque constructivista permite que el proceso de enseñanza se lleve a cabo como un proceso dinámico, participativo e interactivo del estudiante. Sphero permite potenciar el desarrollo del trabajo en equipo al generar discusiones para poder lograr la solución. Así mismo se pudo observar en la experiencia la factibilidad de aplicar los pasos propuestos por Wing (Wing, 2006) para el desarrollo del pensamiento computacional para la resolución de problema.

En este trabajo hemos presentado nuestra experiencia en la implementación del uso del Pensamiento Computacional para la mejora de las capacidades en Geometría Espacial utilizando el robot Sphero. Se ha demostrado la factibilidad de la implementación de un primer taller de dos horas basado en cinco etapas. Las dos primeras son explicativas tanto del problema a resolver como del uso y programación de Sphero. Las siguientes etapas: resolución del problema y programación del problema se enmarcan en la teoría del pensamiento computacional.

Sphero presenta un potencial no sólo para desarrollar habilidades espaciales a través de la enseñanza de la geometría, sino que es posible ser utilizado para reforzar la enseñanza de conceptos de física. Las características de Sphero permiten por ejemplo mostrar la relación linear entre tiempo, velocidad y distancia. El uso de Sphero agrega un componente novedoso y lúdico que ha captado la atención de los alumnos y llevado a resolver el problema en una forma distinta. Así mismo las características de Sphero hacen que sea posible la enseñanza en 3D de un concepto que tradicionalmente se enseña en 2D: el teorema de Pitágoras. Como trabajo futuro se busca hacer una medición cuantitativa.

Referencias

Gardner, H. (2011). Frames of mind: The theory of multiple intelligences. Hachette UK.

Maier, P. H. (1996, March). Spatial geometry and spatial ability–How to make solid geometry solid. In Selected papers from the Annual Conference of Didactics of Mathematics (pp. 63-75).

Woods, E., Billinghurst, M., Looser, J., Aldridge, G., Brown, D., Garrie, B., & Nelles, C. (2004, June). Augmenting the science centre and museum experience. In Proceedings of the 2nd international conference on Computer graphics and interactive techniques in Australasia and South East Asia (pp. 230-236). ACM.

Wing, J. M. (2006). Computational thinking. Communications of the ACM, 49(3), 33-35.

Gadanidis, G., Cendros, R., Floyd, L., & Namukasa, I. (2017). Computational thinking in mathematics teacher education. Contemporary Issues in Technology and Teacher Education, 17(4), 458-477

Gadanidis, G., Hughes, J. M., Minniti, L., & White, B. J. (2017). Computational thinking, grade 1 students and the binomial theorem. *Digital Experiences in Mathematics Education*, *3*(2), 77-96.

Yadav, A., Mayfield, C., Zhou, N., Hambrusch, S., & Korb, J. T. (2014). Computational thinking in elementary and secondary teacher education. ACM Transactions on Computing Education (TOCE), 14(1), 5.

Scratch (2018). Recuperado de http://scratch.mit.edu/.

Alice (2018). Recuperado de http://www.alice.org/.

Code.org (2018. Recuperado de http:/code.org/.

Sphero (2018). Recuperaod de https://www.sphero.com/education.

Lee, I. (2014). CSTA Computational Thinking Task Force.

Carretero, M. (1997). ¿ Qué es el constructivismo. Progreso. Recuperado de: http://www.micentroeducativo.pe/docente/fileproject/file_docentes/ 549bi_2c5224.pdf

Schunk, D. H. (2012). Learning theories an educational perspective sixth Edition. Pearson.

Hernández , R., Fernández, C., & Baptista, P. (2006). Metodología de la investigación (Vol. 3). México: McGraw-Hill.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 22/10/2018
Aceitação/Acceptance: 10/12/2018

Educação a Distância: as características do líder aplicada ao papel do tutor no processo de ensino aprendizagem

Arcângelo dos Santos Safanelli[1], Dalton Francisco de Andrade[1], Juliana Brito[2], Luiz Salgado Klaes[3], Léia Mayer Eyng[4], Vânia Ribas Ulbricht[4]

safanelli.arcangelo@gmail.com[1], dalton.andrade@ufsc.br[2], juligbrito@hotmail.com[3], l.klaes@ufsc.br[4], leiamayer@gmail.com[5], vrulbricht@gmail.com[6]

[1] Universidade Federal de Santa Catarina, Programa de Pós-Graduação em Engenharia de Produção, Florianópolis, Brasil

[2] Instituto Federal de Pernambuco, Tecnologia em Gestão Ambiental, Recife, Brasil

[3] Universidade Federal de Santa Catarina, Departamento de Administração, Florianópolis, Brasil

[4] Universidade Federal de Santa Catarina, Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento, Florianópolis, Brasil

Pages: 795–803

Resumo: Este artigo contextualiza a Educação a Distância (EaD) no Brasil, destacando a sua evolução histórica o processo pedagógico nas distintas fases e cenário que a EaD se insere. Com o objetivo analisar quais as características necessárias de liderança para um tutor de polo EaD, o estudo utilizou da pesquisa bibliográfica e a análise qualitativa acerca do tema proposto. Percebeu-se como resultado, que quando mencionamos os diferentes elementos envolvidos na avaliação do processo de EaD assim destacados: professor-tutor, supervisor na comunicação e o auxílio da aprendizagem dos alunos, e trajetória desses atores junto a EaD, ratificam a necessidade e preferência por estabelecer vínculos que demonstram comprometimento entre os pares, em prol de um objetivo comum. Desse modo, o resultado obtido projeta que o professor-tutor da EaD bem como os demais atores, devem contemplar competências, habilidades e atitudes no processo de mediação de aprendizagem coerente e que valorize as potencialidades dos discentes da EaD de modo a integrá-los nas novas plataformas proporcionadas pela EaD atualmente, visando ampliar a interação dos discentes em relação às múltiplas ferramentas disponibilizadas pelos ambientes acadêmicos virtuais de aprendizagem.

Palavras-chave: Educação a Distância (EaD); Tutor; Liderança.

Distance Education: the characteristics of the leader applied to the role of the tutor in the process of teaching learning

Abstract: This article contextualizes Distance Education (DE) in Brazil, highlighting its historical evolution the pedagogical process in the different phases

and scenario that the DE is inserted. With the objective of analyzing the necessary leadership characteristics a tutor of DE pole, the study used bibliographic research and qualitative analysis about the proposed theme. As a result, when we mentioned the different elements involved in evaluating the DE, the following are highlighted: teacher-tutor, supervisor in communication and the aid of student learning, and the trajectory of these actors with DE, ratify the need and preference by establishing links that demonstrate commitment among peers, for the sake of a common goal. In this way, the obtained result projects that the profesor-tutor of the DE as well as the other actors, must contemplate competences, abilities and attitudes in the process of mediation of coherent learning and that it values the potentialities of the DE students in order to integrate them in the new platforms offered by the DE currently, aiming to broaden the interaction of the students in relation to the multiple tools made available by the academic virtual learning environments.

***Keywords**:* Distance Education (DE). Tutor. Leadership.

1. Introdução

O desenvolvimento da Educação a Distância (EaD) no Brasil, segue inicialmente uma experiência mundial, ressaltadas algumas diferenças. Na EaD brasileira, há a oferta de cursos pelo modelo de correspondência num primeiro momento, todavia, o uso de mídias - rádio e a televisão - , serão incorporadas a EaD e utilizadas em larga escala, obtendo eficácia e efetividade, tornando os conteúdos criativos, bem antes da introdução da Internet.

Historicamente a EaD teve várias denominações como educação por correspondência (Reino Unido); nos Estados Unidos como estudo em casa ou estudo independente; já na Austrália foi chamada de estudos externos. Ainda na França, foi conhecido como telensino ou ensino a distância e teleducação em Portugal. Por fim, Alemanha denominou de estudo ou ensino a distância, e educação a distância na Espanha.

Independente das múltiplas nomenclaturas relatadas pelo mundo, a educação a distância é também um espaço de pessoas. Isso remonta a idéia de que o sucesso da EaD, perpassa a necessidade de criação de instituições e pessoas capacitadas que estabeleçam uma relação de diálogo entre professor e aluno, mediados por materiais didáticos adequados.

Assim, o artigo tem por objetivo analisar quais as características necessárias de liderança para um tutor de polo EaD.

Refletir sobre este papel delegado ao tutor é compreender como este influencia o comportamento de outras pessoas no que refere ao processo de ensino-aprendizagem, sem especificações e detalhes do motivo ou dos processos de liderança. O tutor possibilita a cooperação, promove a interação, auxilia na minimização dos conflitos, mas ao mesmo tempo busca maximizar a criatividade integrativa legítima dos discentes envolvidos.

Neste cenário, o papel do tutor é direcionado a estimular a motivação das pessoas, fazendo com que elas, neste ambiente EaD, tenham objetivos comuns e os percebam como significativos quando menciona-se o processo de formação, mediação e transmissão de conhecimento.

2. Método

A presente pesquisa foi realizada a partir da revisão bibliográfica de uma abordagem qualitativa, por meio da análise de artigos científicos, livros e teses com aprofundamento nos temas relacionados às características no processo pedagógico de ensino-aprendizagem na modalidade EaD e sua evolução ao longo das décadas nos seus diferentes cenários e fases, notadamente pela avaliação dos diferentes elementos envolvidos nesse processo: o professor-tutor, supervisor na comunicação e auxílio da aprendizagem dos alunos, a sua forma de interação ao longo da trajetória em EaD, as formas de estudos e elaboração de ferramentas utilizadas no processo de ensino e aprendizagem, tendo-se em vista uma análise focada no comprometimento entre as equipes de tutores e os discentes na construção pedagógica enfatizando a valorização da figura humana. Por fim, a divulgação dos resultados alcançados por esse estudo através da elaboração de trabalhos científicos resultantes da participação direta dos autores em diferentes instituições de ensino tecnológico à distância e posterior avaliação da interação dos alunos em relação às diferentes ferramentas utilizadas nos ambientes acadêmicos virtuais, tais como os fóruns de discussões e dúvidas, fóruns avaliativos, envios de arquivos, chats de dúvidas e sugestões e as aulas presenciais como suporte ao encontro presencial entre tutores e alunos.

3. Referencial Teórico

3.1. Educação a Distância

A EaD quando se refere a uma padronização conceitual não é unânime. Em 1967, Dohmen, apontava que a EaD, como algo sistemático e organizado, em que o auto-estudo é destacado, mediante o uso de um material de apoio, acompanhamento e supervisão (BERNARDO, 2009).

De forma polêmica, "o conceito de Peters em 1973, dá ênfase a metodologia da EaD, afirmando que "a Educação a Distância é uma forma industrializada de ensinar e aprender" (BERNARDO, 2009, p.85).

> *Educação/ensino a distância é um método racional de partilhar conhecimento, habilidades e atitudes, através da aplicação da divisão do trabalho e de princípios organizacionais, tanto quanto pelo uso extensivo de meios de comunicação, especialmente para o propósito de reproduzir materiais técnicos de alta qualidade, os quais tornam possível instruir um grande número de estudantes ao mesmo tempo, enquanto esses materiais durarem. É uma forma industrializada de ensinar e aprender (MOORE, 1973 apud VASCONCELOS, 2010).*

Com esta fala, Moore (1973 apud VASCONCELOS, 2010), esclarece que as ações do professor e o desenvolvimento da comunicação com os alunos deverá ser facilitada, com o uso de meios e métodos instrucionais. Já observava o autor, que a intervenção do professor deveria ocorrer a partir da ação do aluno sempre mediada por meios impressos, eletrônicos, mecânicos entre outros.

O conceito de Holmberg em (1977 apud VASCONCELLOS, 2010), aponta para as múltiplas formas de estudos em que a EaD apresenta. Na sua conceitualização, Holmberg

ratifica a contínua e imediata supervisão com tutores, porém em sala de leituras e em um espaço físico local.

Em contraponto a questão da presença física entre tutor e aluno, 14 anos depois, em 1991, Keegan ressalta que a EaD faz uso do planejamento, direção e instrução que uma Keegan organização de ensino possui. Conceitualmente, sua definição passar a ser compreendida como:

> *a separação física entre professor e aluno, que a distingue do ensino presencial, comunicação de mão dupla, onde o estudante beneficia-se de um diálogo e da possibilidade de iniciativas de dupla via com possibilidade de encontros ocasionais com propósitos didáticos e de socialização. (VASCONCELOS, 2010).*

De encontro ao proposto por Keegan em 1991, Chaves ratifica (1999 apud VASCONCELOS, 2010), a separação física e o uso de tecnologias de telecomunicação. Para o autor o "ensinante e o aprendente estão separados no tempo e no espaço". Esta separação deveria ser preenchida pela telecomunicação por meio da transmissão de dados, voz e imagens.

Paralelo a esta conceituação, o mercado de trabalho tem solicitado dos indivíduos, uma aquisição constante de conhecimento, ampliando habilidades e competências. Estas novas demandas ratificam a necessidade da modalidade EaD, de modo a explorar este nicho e atender os anseios da comunidade.

Todavia, requer apontar que esta nova demanda, aponta para uma forma diferencial de construção do conhecimento, de saberes. Na EaD, o *modus operanti* baseia-se na ampliação do conhecimento, na troca de informações, no rearranjo dos materiais instrucionais, de modo a extrapolar a estrutura física de uma sala de aula e o sinal do relógio.

Se antes a educação a distância, era reconhecida em uma época em que o correio e o rádio se destacam como suportes e veículos para os cursos das mais diversas áreas, o avanço dos meios de comunicação e da informática computacional, a ampliação e o desenvolvimento de *softwares* educativos, apresentam uma forma diferenciada para a modalidade, ampliando a possibilidade de ensino e a aprendizagem.

Para Salvucci, Lisboa e Mendes (2012, p.51):

> *a didática do ensino virtual deve contemplar a multidimensionalidade dos recursos e ferramentas tecnológicas e os procedimentos adequados prevêem: - Ênfase na autonomia do aluno, Exploração das possibilidades do material didático, Domínio das ferramentas, Conhecimento prévio dos processos de interação e mediação, Disponibilidade e interesse para a comunicação diferenciada das fontes de informação.*

Assim, Pallo e Pratt (2002), acenaram em seus estudos acerca aprendizagem a distância, mediada por mídia computacional, que são fundamentais:

> *[...]as interações entre os próprios estudantes, as interações entre professores e os estudantes e a colaboração na aprendizagem que resulta de tais interações. Em outras palavras, a formação de uma comunidade de alunos por meio da qual o conhecimento é transmitido e os significados sejam criados*

> *conjuntamente, prepara o terreno para bons resultados na aprendizagem. (PALLO; PRATT, 2002, p. 27.)*
>
> *Por fim, conceituar a EaD nos requer compreender com detalhismo o processo de formação, os limites e possibilidades das Tecnologias da Informação e Comunicação (TICs), porém não perdendo de vista o quanto é possível configurar e reconfigurar de modo dinâmico a rede.*

3.2. A Educação a Distância, a Tutoria e a Liderança

Muitos autores vêm criticando a expressão EaD, pois ela apontaria para uma distância que, do ponto de vista educacional ou psicológico, não existe necessariamente quando fazemos EaD. Tori (2010), por exemplo, fala em uma Educação sem Distância. Uma expressão alternativa, cada vez mais utilizada, é Educação Aberta e a Distância. Outros autores, como Silva e Santos (2006) e Silva, Pence e Zuim (2010), preferem a expressão Educação On-line, que sinalizaria a interatividade possibilitada pelas novas TICs. Algumas variações em inglês são online learning, virtual learning, networked learning e web-based learning. A expressão e-learning é também bastante utilizada, em geral para representar a EaD corporativa, que se faz em instituições que não têm o ensino como missão primária.

Contudo, a educação a distância, deve considerar, conforme Faria (2010), que o estudante é sujeito ativo do processo tornando-se protagonista na construção do próprio conhecimento e para isso, este precisa estar disposto a buscar as informações, com estímulo e motivação.

Um elemento importante neste processo de construção do conhecimento e da ratificação da cultura EaD, cabe a presença do tutor, neste processo de mediação. Bruno e Lemgruber (2010) falam de docência coletiva, e de possibilidades que vão da fragmentação neotecnicista à parceria cooperativa, e que tudo depende da concepção de educação que se tem. Lembram que

> *o tutor a distância é também um docente e não simplesmente um animador ou monitor neste processo, e muito menos um repassador de pacotes instrucionais. Este profissional, como mediador pedagógico do processo de ensino e de aprendizagem, é aquele que também assume a docência e, portanto, deve ter plenas condições de mediar conteúdos e intervir para a aprendizagem. Por isso, na prática, o professor-tutor é um docente que deve possuir domínio, tanto tecnológico quanto didático (BRUNO; LEMGRUBER, 2010, p. 75).*

É visível que a equipe de tutores possui um papel fundamental neste processo de educacional, pois é visto como um sujeito protagonista que participa de modo ativo na prática e na construção pedagógica. Independente de estar presencial ou a distância, sua contribuição se dá de modo efetivo na construção do processo de ensino aprendizagem e por conseqüência no processo de acompanhamento e avaliação do projeto pedagógico dos cursos.

Assim, segundo Shulman (1995 apud LITWIN, 2001), é ressaltado o domínio do conteúdo como algo imprescindível e elemento essencial para o exercício da função.

Paralelo a isso, o domínio do conteúdo está aliado à necessidade de dinamismo, visão crítica e global.

Neste cenário, o tutor deve considerar aspectos relacionados a liderança como elementos indispensáveis à sua conduta no processo de ensino e aprendizagem. As transformações que ocorrem no segmento educacional, impõe uma nova realidade para as instituições, exigindo das mesmas, inovadoras formas de agir, estruturas mais flexíveis e valorização do capital humano. No entanto, para se conseguir efetivar o processo de mudança nas instituições de ensino, é necessário que as pessoas que delas fazem parte estejam comprometidas com a mudança. E, neste caso, o papel do tutor que possui liderança é fundamental, uma vez que ele é co-responsável pela valorização da participação efetiva dos discentes, motivando-as e inspirando-as do processo de aprendizagem.

"A ligação aluno-professor ainda é, no imaginário pedagógico, uma dominante, o que torna a tutoria um ponto-chave em um sistema de ensino a distância (MAIA, 1998 apud NISKIER, 1999, p.391). Em função disto, é indispensável que as instituições desenvolvam planos de capacitação de seu corpo de tutores. Um programa de capacitação de tutores deve, no mínimo, prever três dimensões como apontado por Shulman (1995 apud LITWIN, 2001):

- capacitação no domínio específico do conteúdo;
- capacitação em mídias de comunicação; e
- capacitação em fundamentos da EaD e no modelo de tutoria.

Para Niskier (1999, p.320), o papel e a capacitação do tutor deve extrapolar as competências e habilidades, visando a inclusão de atitudes. Deve-se considerar os aspectos didático-pedagógico voltados para acompanhar, motivar, orientar e estimular o aprendizado do discente. "Seu trabalho torna-se mais qualificado com o conhecimento, com profundidade, do material didático produzido pelo professor, muitas vezes, podendo até participar, com sugestões, de sua produção e/ou de sua reformulação, sobremodo se estiver inteirado do projeto político-pedagógico do curso".

Assim, os tutores qualificados, são para Bennis (1996), os líderes que desenvolvem estratégias em aprender, em adquirir conhecimentos e experiências. Assim, seu mundo vai tornando-se mais amplo e complexo, sendo que a complementaridade entre educação formal e informal é extremamente importante para a formação do líder.

O verdadeiro líder tutor, forma-se de maneira espontânea, sem se propor racionalmente para se tornar um líder. E, de acordo com Bennis (1996), a questão não é tornar-se líder, é tornar-se você mesmo, usar-se completamente – todas as suas capacidades, talento e energia – para que sua visão se manifeste.

Desse modo, o tutor atual deve ampliar as dimensões apontadas por Litwin (2001) e para ratificar este apontamento, é fundamental a colaboração de Covey (1994, p.48), ao apontar que valores, idéias, normas e ensinamentos "enobrecem, elevam, realizam, legitimam e inspiram as pessoas".

Para Covey (1994) os princípios/características que devem nortear o exercício da tutoria no atual cenário pode ser assim apontados:

a. os tutores estão **continuamente aprendendo**. Neste sentido, são pessoas que desenvolvem novas habilidades e interesses a partir da curiosidade e estão abertas a adquirir novos conhecimentos, pois são humildes o suficiente para admitir opiniões diferentes e divergentes das suas;
b. os tutores **constantemente voltados para o serviço do processo educacional**. Compreendem a vida como uma missão e não como uma carreira. O sucesso desses líderes está diretamente relacionado com o sucesso das pessoas que estão ao seu redor, com o todo, havendo uma atuação harmônica entre as pessoas;
c. os tutores possuem atitude otimista, e seu espírito é entusiasta. Essas pessoas **irradiam energia positiva** e contagiam todos a sua volta. Assim, percebe-se que o ambiente é influenciado pela postura que as pessoas assumem perante os fatos e situações da vida;
d. os tutores **acreditam nas outras pessoas**, mesmo que o potencial delas ainda esteja latente. Como conseqüência, essa postura cria um ambiente favorável ao crescimento e à oportunidade, podendo despontar potencialidades que não haviam se configurado ainda;
e. os tutores baseados em princípios conciliam aspectos físicos e mentais, desenvolvendo ambas as habilidades físicas e cognitivas. Por isso, **suas vidas são equilibradas**. Este equilíbrio refere-se também ao reconhecimento de seu próprio valor, porém, sem a necessidade de mostrar-se aos outros. Pessoas baseadas em princípios, "vivem sensatamente no presente, planejam cuidadosamente o futuro, e adaptam-se com flexibilidade às circunstâncias mutáveis". (Covey, 1994, p. 11);
f. os tutores sabem saborear a vida, pois possuem segurança, ao invés de dependerem de fatores externos. Portanto, **encaram a vida como uma aventura** e fazem ricas as suas vidas, pois não lhes faltam vontade de aprender;
g. a **sinergia** é uma outra característica dos tutores, eles devem ser catalisadoras de mudanças, melhoram qualquer situação em que se envolvem e trabalham de forma tão árdua quanto inteligente. Além disso, sabem trabalhar em equipe, relacionam-se de forma sincera, contribuindo para o processo criativo na solução de problemas.
h. complementando uma vida equilibrada, os tutores orientadas por princípios **exercitam-se pela auto-renovação**, representada pela congruência das quatro dimensões da personalidade humana: física, mental, emocional e espiritual.

Com a observância dos princípios acima descritos, a função essencial do tutor voltada para orientação didática tornar se menos árdua, pois grande parte dos alunos que não possui hábito de autoestudo, nem autodisciplina, característicos dessa modalidade de educação serão acolhidos de modo coerente, orientados e estimulados por meio do eles tem de potencial, e não pelo que apresentam como distorções diante do modelo EaD. É preciso que o tutor seja flexível, perseverante, persistente e sensível para lidar com eventuais dificuldades encontradas por parte dos alunos (ALVES; NOVA, 2003).

Por fim, o quadro de tutores previstos para o processo de mediação pedagógica deve especificar a relação numérica estudantes/tutor capaz de permitir interação no processo de aprendizagem. Como Santos e Silva (2009) apontava a construção do conhecimento

deve considerar que a interação é um processo importante seja no modelo presencial ou na EaD. Cada pessoa é um ser social, relacional e participante de um processo único e histórico. Contudo, a este processo de ensino-aprendizagem é requisitado que aquele que ensina e o que aprende, estabeleçam uma relação entre eles.

4. Conclusões

Com as alterações ocorridas no cenário educacional nas últimas décadas, e a ampliação do modelo de ensino a distância no segmento das instituições de ensino superior, homem e neste caso, especificamente o tutor, passou a assumir uma posição primordial e elementar dentro da estrutura educacional da EaD. Não resta dúvida que no processo de ensino aprendizagem, a mediação realizada pelos tutores deve ser ratificada, qualificada e legitimada, pelas instituições da EaD.

Este processo ocorre conforme o modelo de gestão das IES envolvidas, e desse modo algumas estão preocupadas com a qualificação do seu capital humano e outras nem tanto. Todavia, este tema relacionado especificamente a análise da gestão, cabe um olhar detalhado, em outro momento.

Cabe ressaltar com o estudo ora apresentado, observou que ainda de modo incremental está acontecendo no cerne das organizações um processo global de desenvolvimento da EaD, o qual propicia e resulta no desenvolvimento das pessoas, do chamado "capital intelectual", neste caso os tutores, foco do estudo.

Percebeu-se que aos tutores foi direcionado uma formação de competências e técnicas que significativamente interferem no modelo de ensino proposto. Porém, o estudo, apresentou que o desenvolvimento de atitudes juntos aos tutores, e estes focados na perspectiva da liderança, legítima facilitadores integrados aos objetivos organizacionais, bem como ao modelo de EaD hoje ratificado.

Por conseqüência, a organização que ampliará o olhar sobre o tutor da EaD, como líder no seu segmento, e nas atribuições de suas funções, tenderá a se tornar uma organização com maior competitividade frente a outras organizações, o que torna isso um diferencial no atual cenário de crise.

Assim, a liderança vem recebendo, cada vez mais, um papel de destaque nas organizações, que no decorrer do século XX privilegiaram o gerenciamento. Da mesma forma, uma nova era no desenvolvimento das organizações vem surgindo, é a era do conhecimento, onde a liderança não se restringe aos cargos ocupados pela gestão, mas ampliou-se às pessoas capazes orientar, de disseminar o conhecimento, o espírito de liderança e de equipe.

Hoje, entretanto, espera-se das organizações da EaD o respeito e a valorização da figura humana, tutor ou professor. E, neste contexto, cabe a gestão, o desenvolvimento do profissional, respeitando-o enquanto indivíduo, o seu espaço substancial dentro da racionalidade instrumental da organização.

Referências

ALVES, L.; NOVA, C. (2003). *Educação a Distância:* Uma Nova Concepção de Aprendizagem e Interatividade. São Paulo: Futura.

BENNIS, Warren. (1996). *A Formação do Líder*. São Paulo: Atlas.

BERNARDO, V. (2015). *Educação a distância*: fundamentos. Universidade Federal de São Paulo UNIFESP. Disponível em: http://www.virtual.epm.br/material/tis/enf.htm Acesso em: 14 set.

BRUNO, A. R.; LEMGRUBER, M. S. (2010). Docência na educação online: professorar e (ou) tutorar? In: BRUNO, A. R. et al. *Tem professor na rede*. Juiz de Fora: UFJF.

COVEY, Stephen R. (1994). *Liderança baseada em princípios*. Rio de Janeiro: Campus.

FARIA, J. G. (2015). *A construção do conhecimento em cursos a distância*. 2010. Disponível em: <http:// pt.scribd.com/doc/26856323/A-CONSTRUCAO-DO-CONHECIMENTO-EM-CURSOS-DE-EAD>. Acesso em: 10 out.

LITWIN, E. (Org.). (2001). *Educação a Distância:* Temas para Debate de uma Nova Agenda Educativa. Porto Alegre: Artmed.

NISKIER, A. (1999). *Educação a Distância:* A Tecnologia da Esperança. São Paulo: Loyola.

PALLO, Rena M.; PRATT, Keith. (2002). *O aluno virtual:* um guia para trabalhar com estudantes on- -line Porto Alegre: Artmed.

SALVUCCI, M.; LISBOA, M.; MENDES, Nelson C. (2015). Educação a Distância no Brasil: Fundamentos legais e implementação. *ASSOCIAÇÃO BRASILEIRA DE EDUCAÇÃO A DISTÂNCIA*, s.n. , vol. 11, 2012. Disponível na internet: < http://www.abed.org.br/revistacientifica/Revista_PDF_Doc/2012/artigo_04_v112012.pdf>. Acesso em: 10 out.

SANTOS, R. M.; SILVA, P. (2009). A didática da EAD virtual. In: *Anais do II Seminário Educação, Comunicação, Inclusão e Interculturalidade*. EDaPECI.

SILVA, M. A.; SANTOS, E. (Org.). (2006). *Avaliação da aprendizagem em educação on-line:* fundamentos, interfaces e dispositivos, relatos de experiências. Loyola.

SILVA, M.; PESCE, L.; ZUIN, A. (Org.). (2010). *Educação on-line:* cenário, formação e questões didático metodológicas. Rio de Janeiro: Wak.

TORI, R. (2010). *Educação sem distância:* as tecnologias interativas. São Paulo: Senac SP.

VASCONCELOS, S. P. G. (2010). *Educação a Distância*: histórico e perspectivas. Universidade do Estado do Rio de Janeiro (UERJ). Disponível em:http://www.filosofia.org.br/viiifelin/19htm . Acesso em: 08 jan.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: xx/09/2018
Aceitação/Acceptance: xx/11/2018

Ingeniería de Requerimientos: Análisis ontológico de su desarrollo en la literatura

Jonathan Vásquez[1], Ariel La Paz[2]

Jonathan.vasquez@uv.cl, lapaz@fen.uchile.cl

[1] Universidad de Valparaíso, 2340000, Valparaíso, Chile.

[2] Universidad de Chile, 8330015, Santiago, Chile.

Pages: 804–815

Resumen: Las actividades de Ingeniería de Requerimientos (IR) han sido consideradas como claves de éxito de los proyectos de Sistemas de Información (SI). Una revisión literaria de la IR permitiría identificar qué áreas de la disciplina han sido cubiertas (*birght side*), pero queda al debe para identificar aquellas que no (*blind/blink side)*. En este artículo proponemos una ontología que permite (1) describir el cuerpo del conocimiento de la IR, (2) identificar los *blind/blink* y *bright side* por medio del mapeo de artículos en esta ontología. Los resultados del mapeo muestran que los artículos se han concentrado principalmente en las áreas relacionadas a los *stakeholders* y proceso de negocios de la IR, específicamente en la actividad de elicitación con el propósito de analizar los requerimientos, dejando de lado aquellas interesantes tales como simulación, negociación, re-uso de requerimientos y el entorno/ambiente empresarial de la disciplina.

Palabras-clave: requerimientos; ingeniería de requerimientos; ontología.

Requirements Engineering: Ontological analysis of its development in the literature

Absract: The Requirements Engineering activities have been considered as keys to the success of the Information Systems (IS) projects. A literary review of the IR would identify which areas of the discipline have been covered (birght side), but it is left to identify those that do not (blind / blink side). In this article we propose an ontology that allows (1) to describe the body of knowledge of the IR, (2) identify the blind / blink and bright side through the mapping of articles in this ontology. The results of the mapping show that the articles have been concentrated mainly in the areas related to the stakeholders and business process of the IR, specifically in the elicitation activity with the purpose of analyzing the requirements, leaving aside interesting ones such as simulation, negotiation, re-use of requirements and the environment / business environment of the discipline.

Keywords: Requirements; Requirements Engineering; Ontology.

1. Introducción

La Ingeniería de Requerimientos, cuyo proceso contempla las actividades de captura, análisis, negociación, documentación y validación de los requerimientos, ha sido considerada en la literatura como uno de los factores claves para el éxito de los proyectos de desarrollo de Tecnologías y Sistema de Información (Daneva et al, 2014; Cheng & Atlee, 2007). De este modo, una correcta ejecución del proceso de IR generaría una especificación de requerimientos como reflejo de las necesidades de los *stakeholders,* y por consiguiente, bajo el uso de este documento, accedería al equipo de desarrollo diseñar e implementar una solución tecnológica que mejor se ajuste a estas necesidades. La correcta ejecución permitiría también una mejor comunicación entre desarrolladores y *stakeholders*, como también un óptimo acuerdo en el proceso de negociación sobre la caracterización de las tecnologías y sistemas de información (TSI) que este último necesita (Parnas, 1999; Nuseibeh & Easterbrook, 2000; Cheng & Atlee, 2007). Desde el punto de vista académico, desde el surgimiento de la disciplina, se observa un creciente interés por la IR, mostrado en el número de citas y publicaciones. Más en detalle, el número de artículos y citas de trabajos publicados entre los años 1978 y 2017 en revistas Web of Science (WoS) refleja un importante crecimiento desde finales de la década de los '80 (ver Gráfico 1). Del mismo modo, la disciplina ha llegado a una madurez tal que ya se han establecido conferencias de alto nivel de la disciplina (RE, REFSQ y EmpiRE) y surgido revistas especializadas en IR, como por ejemplo, *Requirements Engineering Journal*. Este progreso académico ha producido y validado herramientas para las aplicaciones profesionales, las cuales pueden ser utilizadas en la obtención y análisis de requerimientos de los proyectos de software. Alexander (2016) plantea que entre los años 2009 y 2016, la variedad de herramientas y tecnologías de IR es bastante amplia y que el dinamismo respecto al uso de estas es bastante alto. Algunos ejemplos de herramientas IR son el *template* Volere y el EARS (Robertson, 1999; Ahrens et al, 2016), las herramientas de modelamiento de lenguaje unificado (UML), tales como los diagramas de casos de uso y diagramas de secuencias, y los softwares que apoyan en las distintas etapas del proceso de IR.

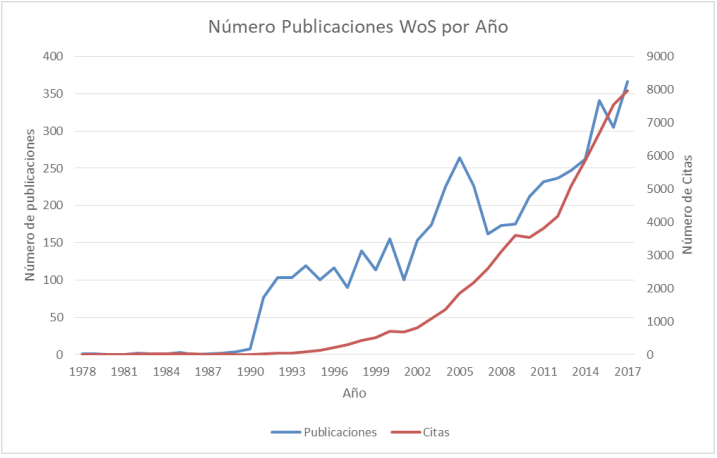


Gráfico 1 – Tendencia de publicaciones y citas WoS entre los años 1978 y 2017.

Aun con el gran interés y dinamismo logrado en el desarrollo de la IR por parte de investigadores y profesionales, la eficiencia y efectividad en el desarrollo de los proyectos de TSI siguen siendo deficientes. De este modo, surge la interrogante si el estado del desarrollo logrado en IR, principalmente en las últimas cuatro décadas, ha sido sistemático y comprehensivo en cuanto a la identificación de áreas que han sido ampliamente cubiertas por la academia. En cierto modo, las revisiones literarias (LR) han respondido a esta necesidad; es decir, las LR de la IR han permitido a los investigadores y profesionales conocer hasta dónde se ha llegado, qué caminos han sido ampliamente recorridos y, en cierto grado, cuáles faltarían por recorrer. En términos prácticos, el resultado obtenido de una RL es una descripción, resumen, evaluación y clarificación de lo desarrollado en la literatura de la disciplina, como también, por medio del análisis del revisor, la realización de una propuesta de posibles directrices futuras (Jesson, Matheson, & Lacey, 2011). Sin embargo, las RLs adolecen de sesgo en el sentido que solamente muestran las áreas que han sido cubiertas (*bright side*), siendo ciegas en identificar de manera sistemática y sistémica aquellas que no se han desarrolladas (*blind/blink side*).

El desarrollo de una disciplina es un proceso complejo y difuso, originalmente sin directrices o agenda establecidas. En efecto, sin una imagen clara del cuerpo de la disciplina, se puede correr el riesgo de solamente observar una parte del todo, lo que en consecuencia, generaría una RL sin abordar por completo el problema que se busca solucionar (Ramaprasad & Syn, 2013; Ramaprasad & Syn, 2015). En este sentido, se ofrece como una atractiva alternativa la identificación y el uso de una metodología que permita visualizar el todo para luego identificar los énfasis principales ya documentados y sus brechas. Bajo este contexto, la Ontología permite, primero, identificar y describir todas las partes que componen el cuerpo de la disciplina, y segundo, descubrir, por medio del mapeo de investigaciones publicadas, las áreas del cuerpo del conocimiento que han sido ampliamente cubiertas y, más interesante, las que han sido a la fecha ignoradas (Ramaprasad & Papagari, 2009; Ramaprasad & Syn; 2013).

En este artículo, con el objetivo de describir el cuerpo de la IR, desarrollamos una ontología, la cual permite estructurar y deconstruir el complejo proceso de desarrollo de la disciplina. Posteriormente, mapeamos los artículos que fueron utilizados en revisión literaria por los autores en Nuseibeh & Easterbrook (2000), Cheng & Atlee (2007) y Matyokurehwa et al (2017) pues abarcan tres ventanas temporales complementarias entre los años 1987 y 2016. La siguiente sección describe los resultados obtenidos en las tres RLs nombradas, mientras que las dos siguientes describen la construcción Ontológica y el mapeo de artículos realizados sobre esta ontología. Finalmente, en la última sección se plasman las conclusiones y discusiones de los resultados obtenidos de la sección 4.

2. Trabajos Relacionados

En la literatura es posible identificar RLs de la IR publicadas en épocas que abarca una ventana de tiempo que van desde los 6 a 16 años, clarificando así las áreas que han sido principalmente tratadas e investigadas en el rango de años de la investigación. Uno de la primeras revisiones fue realizada por Nuseibeh et al (2000), considerando una venta temporal de 10 años aproximadamente y que finalizaba en el 2000. En ella, los

autores establecieron que la IR experimentó un gran desarrollo en la década de los 90', pues esta pasa de ser una disciplina emergente a lograr un desarrollo suficiente para establecer una comunidad académica capaz de sustentar la realización de encuentros y *workshops*. Adicionalmente, establecen la necesidad de la creación de técnicas en las distintas etapas del proceso de la IR.

Siete años más tarde, Cheng & Atlee (2007) realizan una actualización del trabajo de Nuseibeh et al (2000), proponiendo una matriz que estructura y organiza las investigaciones realizadas en la IR hasta entonces. Los resultados indicaron la existencia de un gran número de herramientas generadas posterior al período considerado por Nuseibeh et al (2000), identificando un sostenido crecimiento en el desarrollo de la disciplina. Adicionalmente, los autores plantean la necesidad de fortalecer la relación investigador-*practitioner*, donde el investigador debe incorporar a los profesionales en sus proyectos y estos últimos retroalimentar y proveer de datos a los investigadores.

Más recientemente, Matyokurehwa, Mavetera y Jokonya (2017) realizaron una RL de las técnicas de IR para el período 2000-2016. Los autores identificaron que, aun existiendo una gran variedad de técnicas, ninguna de estas lograba cubrir todo el proceso de la IR, teniéndose que escoger más de una en los proyectos de esta índole. Como conclusión, los autores apelaron por la creación de técnicas que faciliten la integración de todas las etapas de IR, en especial, en la gestión del cambio de requerimientos. Cabe destacar que dada la complejidad y madurez de la disciplina, investigadores del área han realizado RL focalizadas, considerando sólo ciertas áreas de especialización del cuerpo del conocimiento de la IR. Algunas de estas son el uso de técnicas en metodologías de desarrollo ágiles (Zamudio et al, 2017) o la consideración solo de estudios empíricos de aplicaciones de técnicas de IR (Ambreen et al, 2018).

Una alternativa a las RL para la descripción del desarrollo de una disciplina es el uso de las Ontologías. La versatilidad de este instrumento se muestra por su uso en variadas disciplinas, tales como, sistemas de información (Manzano, 2018; Sastry et al, 2017; Ramaprasad & Syn, 2015), salud (Nuñez et al, 2018; Sastry et al, 2017; Ramaprasad, 2016; Ramaprasad & Syn, 2015), cultura (Yaco & Ramaprasad, 2018), educación (Coronado et al, 2015) y ecología (Cancino et al, 2018). Los ventajas de esta metodología varían en su uso y disciplina, pudiéndose obtener resultados que van más allá a los comparables con una RL. Por ejemplo, en el trabajo realizado por La Paz et al (2015) describieron sistemática y sistémicamente el cuerpo del conocimiento de la disciplina *e-commerce* por medio de la construcción de una ontología compuesta por dos subontologías y 6 dimensiones; además, mapearon los artículos publicados en la revista Journal of Theoretical and Applied Electronic Commerce Research (JTAER) para identificar las áreas que la revista más y menos ha abarcado en sus distintos volúmenes, pudiendo describir brechas respecto a lo publicado y la declaración de propósito de la revista; los resultados de este estudio permitirían a los investigadores y editores plantear enfoques de las publicaciones, dando directrices para posibles ediciones especiales de la revista. Por otro lado, los autores Cancino et al (2018) identificaron los énfasis y brechas de las innovaciones tecnológicas en un contexto de crecimiento sostenible de la organización por medio de (1) la construcción de un marco ontológico del cuerpo del conocimiento de la disciplina y (2) el mapeo y análisis de 375 artículos publicados en las revistas más

prestigiosas de la disciplina. Los resultados obtenidos les permitieron identificar aquellas áreas menos tratadas en la disciplina, como también aquellos temas que son más y menos estudiados en conjunto, como por ejemplo la consideración de universidades y gobiernos locales en los estudios de la innovación tecnológica sostenible; sus resultados ofrecen una suerte de mapa para los investigadores interesados en temas que todavía no han sido abarcados, considerando más aún el creciente interés por los temas ecológicos en las producciones limpieas.

3. Construyendo y mapeando el cuerpo del conocimiento de IR

3.1. Construyendo la Ontología de IR

La ontología se centra en la naturaleza y la estructura de las cosas, independientemente de cualquier otra consideración e incluso independientemente de su existencia real (Guarino et al, 2009). En este artículo definimos la ontología de la IR por medio de combinaciones ordenadas de dimensiones, cada una representada por taxonomías. Por conveniencia lingüística y correspondencia con los artículos descritos por la RL en la ventana en estudio, hemos definido cada elemento de la ontología en inglés.

Definimos la disciplina IR como la combinación ordenada de dos sub-ontologías, siendo estas el Proceso de Ingeniería y la Gestión de Requerimientos. Para el caso del Proceso de Ingeniería identificamos dos dimensiones: (1) Activity (actividad) y (2) Entity (entidad); mientras que para el caso de la Gestión del Requerimiento, las dimensiones fueron tres: (1) Stage (etapa), (2) Instrument (instrumento), y (3) Purpose (propósito). En la Figura 2 se presenta la ontología y las taxonomías de cada dimensión. En ella se observa que la combinatoria de cada elemento de las dimensiones genera un total de 4.116 áreas del cuerpo de la disciplina IR. A modo de ejemplo, un sector de investigación es el *Modelamiento de procesos por medio de una metodología de análisis para la identificación de requerimientos (Model process by analysis method for identify requirements),* o bien, *Simular el ambiente/contexto por medio de herramienta de monitoreo para limitar requerimientos (Simulate environmente/context by a monitoring tool for limit requirements).*

Para efectos de un mejor entendimiento en la Tablas 1 y Tabla 2 construimos un glosario de la ontología. En estas tablas se explicita una descripción detallada de cada sub-ontología, dimensión y taxonomía. Estas descripciones se obtuvieron por medio de un análisis exhaustivo del cuerpo del conocimiento y por medio de la deconstrucción y construcción de este.

Una vez definida esta ontología y las definiciones de cada elemento que la compone, realizamos un mapeo a nivel de taxonomías de cada uno de los artículos revisados por Nuseibeh & Easterbrook (2000), Cheng & Atlee (2007) y Matyokurehwa et al (2017), contabilizando un total de 187 publicaciones relacionadas con la IR. Los resultados que muestran los *brigh* y *blind/blink side* se muestran en la siguiente sección. Cabe destacar que previo al proceso de codificación, la ontología fue probada con una muestra artículos, de esta manera se llegó a un marco robusto antes de realizar el mapeo de todas las publicaciones.

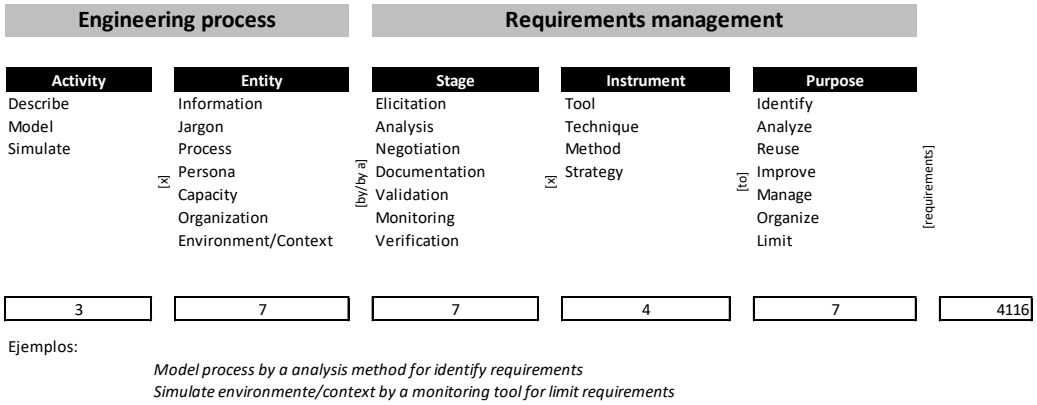


Figura 2 – Ontología de la Gestión de la Ingeniería de los Requerimientos (*Requirements Engineering Management*).

Engineering process	
Se entiende por proceso de ingeniería como el arte y técnica de aplicar los conocimientos científicos a la invención, diseño, perfeccionamiento y manejo de nuevos procedimientos en la industria y otros campos de aplicación científicos.	
Activity	**Actividades del proceso de ingeniería**
Describe	Caracterizar y definir el estado del objeto en estudio.
Model	Simplificar comportamientos de algún elemento en estudio. También es utilizado para explicar tales comportamientos.
Simulate	Mostrar, de manera simplificada y bajo varios supuestos, el comportamiento esperado del objeto de estudio.
Entity	**Objetos de estudios que podrían formar parte de cualquier proceso de ingeniería**
Information	Información/datos
Jargon	Jerga o lenguaje utilizado para la comunicación y transferencia de información/datos
Process	Procesos de negocios que agrupan actividades relacionadas lógicamente. En las actividades pueden participar personas, tecnologías y sistemas.
Persona	Usuario o persona.
Capacity	Tecnologías y sistemas de información.
Organization	Unidades de negocios. Estas indican el orden jerárquico que existe entre las personas.
Environment/Context	Entorno, contexto y/o ambiente.

Tabla 1 – Descripción de la sub-ontología del Proceso de Ingeniería, sus dimensiones y taxonomía.

Requirements management	
La gestión de requerimientos se entiende como el conjunto de actividades que permite definir, documentar y mantener los requerimientos de los usuarios/clientes en un proceso de diseño del sistema de información.	
Stage	**Etapas que componen el ciclo de gestión de los requerimientos**
Elicitation	Identificación y captura de los requerimientos.
Analysis	Análisis de los requerimientos elicitados.
Negotiation	Negociación para el filtro y priorización de los requerimientos identificados.
Documentation	Especificación y escritura final de los requerimientos priorizados.
Validation	Validación con la contraparte de los requerimientos que están especificados en la documentación.
Monitoring	Monitoreo y trazabilidad de los requerimientos en el proceso de diseño e implementación de los sistemas de información.
Verification	Verificación que el SI entregado, que fue levantado con los requerimientos anteriormente especificados, se ajuste y cumpla con los requerimientos.
Instrument	**Instrumentos utilizados en la gestión de los requerimientos.**
Tool	Elemento elaborado con el objetivo de hacer más sencilla una determinada actividad o labor mecánica.
Technique	Conjunto de saberes prácticos o procedimientos relacionados para obtener los requerimientos deseados.
Method	Medio o seres de pasos claramente definidos utilizados para llegar a un fin.
Strategy	Proceso seleccionado a través del cual se prevé alcanzar el estado deseado. A diferencia de la metodología, éste no es una receta de pasos a seguir.
Purpose	**Propósito de la gestión del requerimiento.**
Identify	Identificar requerimientos para su definición
Analyze	Analizar requerimientos previamente definidos
Reuse	Reutilizar requerimientos previamente definidos
Improve	Mejorar requerimientos previamente definidos
Manage	Gestionar los cambios y monitorear el uso de los requerimientos definidos.
Organize	Organizar de manera estructurada los requerimientos.
Limit	Limitar el alcance de los requerimientos.

Tabla 2 – Descripción de la sub-ontología de la Gestión de Requerimientos, sus dimensiones y taxonomía.

4. Resultados

El mapeo de la muestra (Figura 3) de artículos revela que, a nivel taxonómico, los artículos mapeados revelan un enfoque en la descripción por sobre el modelamiento y simulación en lo que respecta a la dimensión Actividad del Proceso de Ingeniería. Por

otro lado, analizando la dimensión Entidades de la misma subontología, las taxonomía principalmente estudiada son Personas (entendiéndose como los *stakeholders)* y Procesos de Negocio. Esto se explicaría por el objetivo principal de la IR en un contexto de gestión de los Sistemas de Información, que es la captura de las necesidades de los *stakeholders* del sistema que se quiere desarrollar en un contexto empresarial. Por otro lado, en la dimensión de las etapas, las más abarcadas son la de Captura y Análisis, mientras que la taxonomía principalmente estudiada en el dominio del propósito de la gestión de IR es el Análisis de los requerimientos. De esta manera al combinar estas taxonomías podemos establecer que las áreas mayormente investigadas en la muestra estudiada son:

1. *Descripción de personas por medio de técnicas de análisis para el análisis de los requerimientos* (*Describe persona by analysis technique to analyze requirements*).
2. *Descripción de procesos por medio de técnicas de análisis para el análisis de los requerimientos* (*Describe process by analysis technique to analyze requirements*).
3. *Descripción de procesos por medio de técnicas de elicitación para el análisis de los requerimientos* (*Describe process by elicitation technique to analyze requirements*).
4. *Descripción de personas por medio de técnicas de elicitación para el análisis de los requerimientos* (*Describe persona by elicitation technique to analyze requirements*).

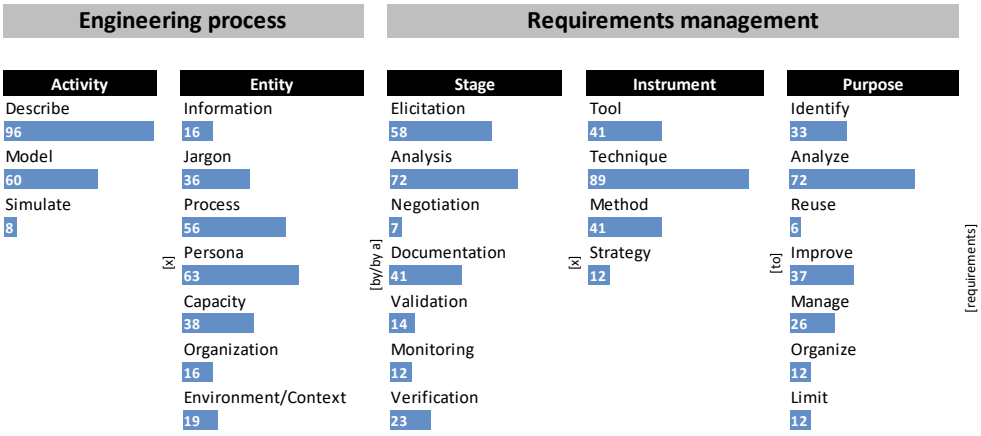


Figura 3 – Frecuencia de mapeo de artículos en las taxonomías.

Por otro lado, las áreas menos cubiertas se pueden identificar por medio de las taxonomías con menos artículos mapeados. Respecto a las actividades del proceso de ingeniería, la taxonomía simulación es la menos investigada en la muestra. En el caso de las entidades, la información y organización son las menos estudiadas. De la misma manera, en las dimensiones de la gestión de los requerimientos la etapa menos considerada es la de negociación, mientras que el instrumento más dejado de lado es la

estrategia. Finalmente, el propósito de reutilización de requerimientos es la que menos ha sido abarcada en la muestra considerada.

Asimismo, podemos establecer que las siguientes áreas serían claves para investigar, dada su casi nula cobertura en la literatura:

1. *Simulación de información por medio de estrategias de negociación para la reutilización de los requerimientos (Simulate information by a negotiation strategy for reuse requirements).*
2. *Simulación de organización por medio de estrategias de negociación para la reutilización de los requerimientos (Simulate organization by a negotiation strategy for reuse requirements).*

En un análisis a un nivel de taxonomías consideradas en conjunto al momento de investigar (es decir, pares de elementos del cuerpo del conocimiento de la IR que son estudiadas en un mismo artículo) podemos ver que son bastantes las duplas que han sido menos cubiertas o prácticamente no han sido estudiadas, revelando una concentración en áreas específicas de investigación en la muestra. A modo de ejemplo, algunas de estas duplas son:

1. *Negociación (etapa) – Estrategia (instrumento)*
2. *Validación (etapa) – Estrategia (instrumento)*
3. *Simulación (actividad) – Monitoreo (etapa)*
4. *Información (objeto)/Jerga (objeto) – Monitoreo (etapa)*
5. *Simulación (actividad)/Información (objeto)/Jerga (objeto) – Negociación (etapa)*

Por otro lado, al analizar las duplas de taxonomías que han sido más abarcadas en la literatura, pudimos identificar que estas son:

1. *Describir (actividad) - Técnica (instrumento)*
2. *Análisis (etapa) – Analizar requerimientos (propósito)*
3. *Describir (actividad) – Analizar requerimientos (propósito)*
4. *Describir (actividad) – Análisis (etapa)*

5. Conclusión y Discusión

La ontología definida en el artículo permite visualizar el cuerpo del conocimiento de la Ingeniería de Requerimientos. Considerando a esta disciplina como compuesta por dos grandes áreas y sus respectivas taxonomías, Proceso de Ingeniería y Gestión de Requerimientos, podemos identificar un total de 4.116 posibles áreas para ser abordadas.

Los resultados mostraron una concentración en las siguientes áreas en el proceso de ingeniería: descripción, personas y procesos de negocio; mientras que a la elicitación y análisis de los requerimientos en el caso de la gestión de los requerimientos. Una explicación de esto el objetivo puede darse por la misión de la IR en el desarrollo de tecnológico, la cual puede ser definida como la identificación de las necesidades de los *stakeholders* para el desarrollo de los Sistemas de Información dentro de un contexto empresarial. Sin embargo, existen otros elementos que son también importantes para

			ENGINEERING PROCESS										REQUIREMENTS MANAGEMENT																	
			Activity			Object							Stage						Instrument				Purpose							
			Describe	Model	Simulate	Information	Jargon	Process	Persona	Capacity	Organization	Environment/Context	Elicitation	Analysis	Negotiation	Documentation	Validation	Monitoring	Verification	Tool	Technique	Method	Strategy	Identify	Analyze	Reuse	Improve	Manage	Organize	Limit
ENGINEERING PROCESS	Activity	Describe																											
		Model	8																										
		Simulate	1	2																									
	Object	Information	10	5	0																								
		Jargon	24	8	1	1																							
		Process	40	14	3	4	6																						
		Persona	39	23	3	6	7	18																					
		Capacity	14	15	4	4	5	9	10																				
		Organization	6	9	1	1	2	4	5	3																			
		Environment/Context	10	9	1	2	1	4	12	2	3																		
REQUIREMENTS MANAGEMENT	Stage	Elicitation	30	18	2	6	15	13	20	18	5	8																	
		Analysis	49	15	2	8	16	27	28	7	0	7	19																
		Negotiation	4	1	0	0	0	1	3	0	0	1	4	2															
		Documentation	17	13	2	2	6	7	8	13	2	1	13	11	3														
		Validation	8	4	1	1	6	3	2	2	0	0	6	8	0	2													
		Monitoring	2	4	0	0	0	3	2	2	3	1	5	5	2	5	0												
		Verification	16	3	2	3	6	9	8	2	1	2	3	9	0	3	6	0											
	Instrument	Tool	22	13	3	2	12	12	15	11	0	2	13	15	4	13	6	1	10										
		Technique	60	21	2	8	18	30	36	15	4	8	28	42	2	16	4	7	7	8									
		Method	17	17	1	2	8	12	12	8	6	7	14	14	1	7	4	1	6	7	8								
		Strategy	3	4	1	2	0	4	2	1	4	1	4	2	0	2	0	3	2	1	0	3							
	Purpose	Identify	22	10	0	2	9	5	13	8	2	3	21	12	1	7	3	0	4	9	16	10	0						
		Analyze	50	13	3	5	17	26	27	8	2	7	22	54	1	11	8	3	9	18	42	15	4	17					
		Reuse	0	4	0	0	0	0	2	1	1	0	1	1	0	3	0	0	0	3	0	2	0	1	1				
		Improve	12	14	2	2	8	11	11	9	1	2	8	9	1	11	0	1	2	6	17	9	4	3	9	0			
		Manage	11	9	1	4	2	5	5	5	5	4	9	6	2	6	0	3	1	6	7	11	5	4	9	3	4		
		Organize	3	3	1	1	1	3	2	3	2	0	6	4	0	3	0	5	0	3	6	2	2	1	3	1	2	6	
		Limit	4	2	1	0	2	4	2	2	0	2	8	4	1	3	1	1	0	2	5	2	2	1	6	0	3	1	2

Tabla 3 – Frecuencia de mapeo de artículos a pares de taxonomías.

su estudio, que pueden complementar o hacer más robusto los nuevos instrumentos generados para la solución del problema que plantea IR, tales como la simulación y negociación de requerimientos.

El mapeo de los artículos de la muestra en la ontología permite generar una guía para futuras investigaciones. En otras palabras, la identificación de los elementos del cuerpo del conocimiento menos abarcados en la literatura permite hallar temas que pueden ser claves para futuras investigaciones, o bien, relacionar elementos que han sido ampliamente investigados con aquellos que no lo han sido o existen muy pocos artículos que lo traten; el impacto de estas investigaciones podría ser mucho mayor para el desarrollo la disciplina respecto de aquellas investigaciones que abarcan temas ampliamente estudiados. A modo de ejemplo, promover herramientas o técnicas de simulación permitiría mejorar el análisis y negociación de los requerimientos, pudiendo llegar a una mejor documentación de los requerimientos, o bien, nuevos instrumentos de monitoreo permitirían una mejor gestión de los requerimientos, sobre todo, en ambientes dinámicos.

Es importante destacar algunas limitantes de este estudio. Sin desmerecer el trabajo disciplinado de los autores respecto de la construcción del marco ontológico, es importante declarar la existencia de una posible sobre especificación o falta de completitud del marco. No obstante, la ontología construida fue discutida con expertos, obtenida por medio del estudio del cuerpo del conocimiento de la IR y probada en el proceso de mapeo de los artículos; por tanto, creemos que estos problemas han sido reducidos a un nivel razonable. En este último aspecto, y con el objetivo de facilitar las futuras investigaciones que busquen robustecer la ontología, se ha realizado un

esfuerzo en construir un glosario, el cual puede ser utilizado como recurso de entrada en el proceso de mejoramiento.

Finalmente, los artículos considerados en este estudio están basados en las RLs de tres artículos, los cuales cuentan con un gran número de citas (3.175 citas según Google Scholar); sin embargo, esto no resguarda que no existan investigaciones importantes a considerar, por tanto, hacemos un llamado a robustecer esta muestra, incluyendo más artículos.

Agradecimientos

Quisiéramos agradecer a Isidora España, estudiante del programa de Ingeniería en Información y Control de Gestión de la Universidad de Valparaíso, por su desempeño en el apoyo de las actividades de colección y procesamiento de artículos científicos y datos que fueron utilizados en esta investigación.

Referencias

Alexander, M. (2016). How to pick the best requirements management tool. *CIO Magazine*. Recuperado de: https://www.cio.com/article/3053082/project-management/how-to-pick-the-best-requirements-management-tool.html.

Ambreen, T., Ikram, N., Usman, M., & Niazi, M. (2018). Empirical research in requirements engineering: trends and opportunities. *Requirements Engineering*, 23(1), 63-95.

Cancino, C. A., La Paz, A. I., Ramaprasad, A., & Syn, T. (2018). Technological innovation for sustainable growth: An ontological perspective. *Journal of Cleaner Production*, *179*, 31-41.

Cheng, B. H., & Atlee, J. M. (2007). Research directions in requirements engineering. *Future of Software Engineering,* IEEE Computer Society, 285-303.

Coronado, F., La Paz, A., Ramaprasad, A., & Syn, T (2015). Navigating the complexity and uncertainty of higher education systems: Ontology mapping of Chile's universities. *HERDSA, Learning for Life and World in a Complex World*, *Australia*.

Daneva, M., Damian, D., Marchetto, A., & Pastor, O. (2014). Empirical research methodologies and studies in Requirements Engineering: How far did we come?. *Journal of systems and software*, *95*, 1-9.

Guarino, N., Oberle, D., & Staab, S. (2009). What is an ontology?. In *Handbook on ontologies* (pp. 1-17). Springer, Berlin, Heidelberg.

Jesson, J., Matheson, L., & Lacey, F. M. (2011). *Doing your literature review: Traditional and systematic techniques*. Sage.

La Paz, A., Ramaprasad, A., Syn, T., & Vasquez, J. (2015). Editorial: An Ontology of e-Commerce – Mapping a Relevant Corpus of Knowledge. *Journal of Theoterical and Applied Electronic Commerce Research*, 10(2).

Manzano, C. A., Ramaprasad, A., & Syn, T. (2018). Information Systems to Manage Local Climate Change Effects: A Unified Framework. *Proceedings of PACIS 2018 (forthcoming)*.

Matyokurehwa, K., Mavetera, N., & Jokonya, O (2017). Requirements Engineering Techniques: A Systematic Literature Review. *International Journal of Soft Computing and Engineering, 7 (1), 14-20*.

Núñez, A., Neriz, L., Mateo, R., Ramis, F., & Ramaprasad, A. (2018). Emergency departments key performance indicators: A unified framework and its practice. *The International journal of health planning and management*.

Nuseibeh, B., & Easterbrook, S. (2000). Requirements engineering: a roadmap. *Proc. of the IEEE Int. Conf. on Soft. Eng, 35-46*.

Parnas, D. L. (1999). Software engineering programs are not computer science programs. *IEEE software*, *16*(6), 19-30.

Ramaprasad, A., & Syn, T. (2013). Ontological meta-analysis and synthesis. *Proceedings of the Nineteenth Americas Conference on Information Systems*, Chicago, Illinois, USA., 1-8.

Ramaprasad, Arkalgud and Syn, Thant (2015). Ontological Meta-Analysis and Synthesis. *Communications of the Association forInformation Systems*, 37, Article 7.

Ramaprasad, A., & Papagari, S. S. (2009). Ontological design. In Proceedings of DESRIST 2009

Ramaprasad, A., Win, K. T., Syn, T., Beydoun, G., & Dawson, L. (2016). Australia's National Health Programs: An Ontological Mapping. *Australasian Journal of Information Systems*, 20.

Sastry, N. K. B., Madhumitha, M., Ramaprasad, A., & Syn, T. (2017). National healthcare programs and policies in India: an ontological analysis. *International Journal Of Community Medicine And Public Health*, *4*(2), 307-313.

Yaco, S., & Ramaprasad, A. (2018). Cultural Heritage Semiotics. In *Digitisation of Culture: Namibian and International Perspectives* (pp. 49-64). Springer, Singapore.

Zamudio, L., Aguilar, J. A., Tripp, C., & Misra, S. (2017). A Requirements Engineering Techniques Review in Agile Software Development Methods. In *International Conference on Computational Science and Its Applications,* Springer, Cham, 683-698.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 14/09/2018
Aceitação/Acceptance: 26/11/2018

Uso de Arquitecturas Empresariales en e-Government Municipal: Un Mapeo Sistemático de la Literatura

Daniela Gallegos-Baeza[1], Ignacio Velásquez[1], Alfonso Rodríguez[1], Angélica Caro[1]

dagalleg@egresados.ubiobio.cl, ignaciovl@outlook.com, mcaro@ubiobio.cl, alfonso@ubiobio.cl

[1] Departamento de Ciencias de la Computación y Tecnologías de Información, Universidad del Bío-Bío, Andrés Bello 720, Chillán, Chile.

Pages: 816–829

Resumen: La Arquitectura Empresarial es un instrumento que permite a las empresas visualizar, para una mejor gestión, el alineamiento de su estructura organizacional, procesos de negocio, sistemas de información e infraestructura TI. En este artículo se presenta un Mapeo Sistemático de la Literatura tendiente a establecer el uso de Arquitecturas Empresariales en el contexto del desarrollo del gobierno electrónico municipal. El resultado de este estudio ha permitido determinar que sólo en un treinta y uno por ciento de los trabajos analizados se utilizan arquitecturas empresariales. Adicionalmente, sorprende la falta de uso de notaciones para el modelado de las arquitecturas empresariales, donde sólo 3 de los nueve estudios que especifican el uso de alguna notación usan una dedicada a representar arquitecturas empresariales.

Palabras-clave: Arquitectura Empresarial, e-government municipal, Servicios de e-government, Mapeo Sistemático de la Literatura.

Use of Enterprise Architectures in Municipal e-Government: A Systematic Mapping of Literature

Abstract: The Enterprise Architecture is an instrument that allows companies to visualize, for a better management, the alignment of their organizational structure, business processes, information systems and IT infrastructure. This article presents a Systematic Mapping of Literature tending to establish the use of Business Architectures in the context of the development of municipal electronic government. The result of this study has made it possible to determine that only thirty-one percent of the analyzed works use business architectures. Additionally, the lack of use of notations for the modeling of business architectures is surprising, where only 3 of the nine studies that specify the use of any notation use one dedicated to representing business architectures.

Keywords: Enterprise Architecture, Municipal e-government, E-government Services, Systematic Mapping of Literature.

1. Introducción

Los ciudadanos usan cada vez más las Tecnologías de la Información y Comunicación (TICs) para el desarrollo de su vida cotidiana, incluyendo actividades sociales, laborales y también en su relación con el gobierno central, que involucra a todo el país, o el gobierno local referido a comunidades más acotadas, las que normalmente corresponden a municipios (Sá, Rocha, & Cota, 2016). Estos gobiernos locales se han convertido durante los últimos años en proveedores de servicios muy variados, en los que el uso de sistemas online posibilita, a diferencia de los mecanismos tradicionales de comunicación presencial, un servicio continuo a lo largo de todo el día (Navarro, Cañavate, & Bleda, 2008), de manera que el uso de las TICs resulta esencial para cumplir este propósito.

En el ámbito local el gobierno electrónico (e-government en inglés) adquiere su máxima dimensión y relevancia dado que los municipios son las organizaciones públicas más cercanas a los ciudadanos y les proporcionan directamente más servicios que ninguna otra administración (Sandoval-Almazán & Colín, 2011; Rocha & Sá 2014). Así mismo, el uso de las TICs en e-government debe ser permanentemente medido y evaluado, para determinar los beneficios reales que se prestan a la ciudadanía (Batlle Montserrat, Abadal, & Blat, 2011). En este sentido resulta fundamental tener una visión global, evidenciada explícitamente, de cómo se vinculan los aspectos motivacionales de un municipio, con sus procesos, servicios y tecnología que los soporta.

Por su parte, las Arquitecturas Empresariales (AEs), definidas como un instrumento de dirección e integración de las TICs para toda la empresa (Lankhorst, Proper, & Jonkers, 2010) han sido sugeridas como un marco para las iniciativas de administración electrónica (Ask & Hedström, 2011; Hornnes, Jansen, & Langeland, 2010; Janssen & Wagenaar, 2003) y como un camino a seguir para lograr estandarización, convergencia e interoperabilidad en los gobiernos locales, puesto que es posible lograr una visión holística del gobierno local que incluye información, personas y tecnología.

Nuestro interés de investigación está centrado en conocer el uso de AEs asociado al desarrollo de servicios municipales en el contexto de e-government local ya que se espera conocer la forma en que estos modelos están siendo utilizados. Consecuentemente, se identificará y catalogará el material bibliográfico utilizando un Mapeo Sistemático de la Literatura (MSL) (Kitchenham et al., 2009) como método de investigación.

El resto del artículo se encuentra organizado de la siguiente forma. En la Sección 2 se presentan conceptos relacionados con el trabajo. La planificación, desarrollo y resultados del MSL es presentado en la Sección 3 y en la Sección 4 respectivamente. Finalmente, en la Sección 5 se presentan las conclusiones y trabajos futuros.

2. Conceptos Relacionados

A continuación, se definen dos conceptos relevantes para esta investigación: e-government y arquitectura empresarial.

2.1. e-government

El e-government es uno los conceptos más interesantes introducidos en el campo de la administración pública a fines de la década del 90 (Moon, 2002). En general se entiende

como el uso de TICs por parte del gobierno, para poner a disposición de los ciudadanos, empresas y otras entidades gubernamentales, servicios a través de la Web.

Aunque existen diversas definiciones de e-government hemos rescatado tres de ellas que nos han parecido las más completas: (i) utilización de las TICs en las tareas y procesos de la administración pública con el objetivo de cambiar la relación entre el gobierno y los ciudadanos, con la prestación de servicios públicos, para mejorar la eficiencia interna e incrementar la participación de los ciudadanos (Batlle Montserrat et al., 2011); (ii) el e-government aborda la mejora de los servicios públicos, la gobernanza y la democracia mediante el desarrollo enfoques centrados en el cliente y la mejora de las interacciones internas y externas a través del uso de las TICs (Dias, 2011) y (iii) e-government corresponde al uso de las TICs para permitir y mejorar la eficiencia de los servicios gubernamentales que se ofrecen a los ciudadanos, empleados y empresas (Carter & Bélanger, 2005).

El e-government es un medio para mejorar la gestión que impulsa una modificación sustancial de la forma de trabajo utilizando la tecnología para transformar y no para automatizar las formas tradicionales de gestión. Además, entrega beneficios directos a la comunidad en general, tales como la eliminación de las barreras de tiempo y espacio, la facilidad en las comunicaciones, transparencia, el acceso igualitario a la información, disponibilidad de información del gobierno, reducción de costos, el incremento en la producción de bienes y servicios; en suma, una mejor calidad de vida para la ciudadanía (Sandoval-Almazán & Colín, 2011; Zuiderhoek, Otter, Bos, & Brinkkemper, 2006).

La relación más estrecha entre el gobierno y la ciudadanía se establece a través de los municipios (Sandoval-Almazán & Colín, 2011; Rocha & Sá 2014). Los municipios se han convertido durante los últimos años en proveedores de servicios muy variados, en los que el uso de sistemas online y TICs posibilita un servicio continuado a lo largo de todo el día (Navarro et al., 2008).

Finalmente, es importante tener presente que la provisión de servicios a través de e-government, no puede ser concebida sin considerar los procesos que coordinan las acciones estatales que los producen, los sistemas de información que les dan vida y la infraestructura tecnológica que asegura su viabilidad y sustentabilidad (Poggi, 2013).

2.2. Arquitectura Empresarial (AE)

Una Arquitectura Empresarial es un instrumento de dirección e integración para toda la empresa que ayuda a orientar y optimizar las inversiones de TI de una organización y permite traducir las estrategias comerciales en soluciones tecnológicas implementables (Lankhorst et al., 2010). La AE, está apoyada en principios, métodos y modelos que se utilizan en el diseño y la realización de la estructura organizacional de la empresa, procesos de negocio, sistemas de información e infraestructura (The Open Group, 2017).

Las aplicaciones importantes de la AE son básicamente el (i) análisis de problemas en el estado actual de una empresa (As Is, en inglés), (ii) determinar el estado futuro o deseado (To Be, en inglés) y (iii) asegurar que los proyectos de desarrollo dentro de los programas de transformación estén de acuerdo con los estados futuros que se quieren lograr. Esto implica que en los modelos de AE, la coherencia y la visión general son más importantes que la especificidad y el detalle (Lankhorst et al., 2010).

3. Mapeo Sistemático de la Literatura (MSL)

Un MSL proporciona una visión global sobre un tema de interés, e identifica la cantidad, tipo de investigación y resultados disponibles sobre el mismo. Asimismo permite identificar temas en los que la evidencia empírica sea escasa y sea necesario realizar más estudios (Genero, Cruz-Lemus, & Piattini, 2014).

La metodología de MSL propuesta por (Kitchenham et al., 2009), permite realizar un estudio objetivo y riguroso de la literatura y consiste basicamente de tres etapas: planificación de la revisión, realización de la revisión e informe de los resultados. La aplicación apropiada de esta propuesta, aplicada de manera rigurosa, permite reducir el riesgo de prejuicios en la conducción de un estudio y asegurar la completitud de los resultados de la revisión.

A continuación se describe la planificación del presente MSL y su desarrollo.

3.1. Identificación de la necesidad de revisión

El objetivo del MSL a desarrollar es identificar material bibliográfico que permita conocer el uso de AEs asociado al desarrollo de servicios municipales en el contexto de e-government local. Basándose en este objetivo, se definieron las siguientes preguntas de investigación (**Q**):

Q1. ¿Se utilizan AEs para apoyar el desarrollo de e-government municipal?
Q2. ¿Cuáles son los servicios de e-government municipal identificados en la literatura y sus clasificaciones?

3.2. Desarrollo del protocolo de búsqueda

Los términos de búsqueda y combinaciones se generaron en tres pasos: (i) derivación de términos importantes de las preguntas formuladas, (ii) inclusión de sinónimos y conceptos relacionados a los términos derivados y (iii) combinación de términos utilizando operadores lógicos AND y OR. En la Tabla 1 y Tabla 2, se muestran los términos y combinaciones utilizados.

ID	Término	ID	Término
T1	E-Government	T2	Electronic Government
T3	Local Government	T4	Municipality
T5	City Council	T6	Services
T7	Electronic Services	T8	Service Model
T9	Enterprise Architecture	T10	ArchiMate

Table 1 – Términos de búsqueda en inglés

ID Combinación	Términos y Combinaciones
CI-1	(T3 OR T4 OR T5) AND (T1 OR T2)
CI-2	(T3 OR T4 OR T5) AND (T1 OR T2) AND (T6 OR T7 OR T8)
CI-3	T9 AND (T3 OR T4 OR T5) AND (T1 OR T2)
CI-4	T9 AND (T3 OR T4 OR T5) AND (T1 OR T2) AND T10

Table 2 – Combinaciones de búsqueda

Para poder realizar el MSL, se utilizaron 4 fuentes bibliográficas: Springer Link, Scopus Science, Web of Science y Google Scholar. Por cada motor de búsqueda y combinación de términos, tanto en inglés como español, se revisaron los primeros 200 resultados de la búsqueda. Los resultados se registraron en una planilla, donde se detalla la cantidad de artículos encontrados, cuántos de ellos fueron seleccionados y cuántos rechazados. Además, se utilizó un gestor de referencias para almacenar la información bibliográfica. Para la selección de trabajos se aplicaron los criterios de inclusión y exclusión presentados en la Tabla 3.

Criterios de Inclusión	Trabajos que respondan a las preguntas establecidas. Trabajos en inglés y español.
Criterios de Exclusión	Estudios que no respondieran la pregunta planteada y los repetidos del mismo estudio. Estudios que, a pesar de contener los términos de búsqueda o combinación de ellos, no contienen información relevante sobre el tema.

Table 3 – Criterios de inclusión y exclusión

3.3. Resultados de la revisión

El proceso de selección de estudios primarios fue realizado entre los meses de marzo y mayo del año 2018. Fueron realizadas un total de 32 búsquedas, de las cuales 10 dieron como resultado más de 200 artículos, en las 22 restantes los resultados fueron menores a 200. Las revisiones parciales que se realizaron fueron aproximadamente de 2.404 resultados, sin embargo, como se puede observar en la Tabla 4 y en la Tabla 5 un gran número de resultados fueron descartados por no cumplir con los criterios de inclusión establecidos.

Q	Motor de búsqueda	Resultados	Revisados	Descartados	Pre-seleccionados	Seleccionados
Q1	**Springer Link**	118	118	104	14	10
	Scopus	13	13	7	6	4
	Web of Science	0	0	0	0	0
	Google Scholar	1.484	304	265	39	34
Total		1.615	435	376	59	48

Table 4 – Resultados para interrogante Q1

Q	Motor de búsqueda	Resultados	Revisados	Descartados	Pre-seleccionados	Seleccionados
Q2	**Springer Link**	2.908	400	399	1	1
	Scopus	1.611	412	405	7	5
	Web of Science	456	357	328	29	29
	Google Scholar	43.060	800	754	46	40
Total		48.035	1.969	1.886	83	75

Table 5 – Resultados para interrogante Q2

Finalmente, se obtuvo un total de 123 artículos seleccionados de los cuales 48 corresponden a la interrogante Q1 relacionada a AEs y 75 a la interrogante Q2 relacionada a servicios. Un documento con el detalle de cada artículo seleccionado puede ser encontrado en http://colvin.chillan.ubiobio.cl/mcaro/gallegos/.

4. Presentación y Análisis de Resultados

Los estudios seleccionados van desde el año 2000 hasta el presente año. En general, cabe destacar el interés en la comunidad en abordar el tema de e-government, evidenciado por el creciente número de estudios que tratan el tema según el análisis de la tendencia de las publicaciones realizado. En la Figura 1 se muestra la cantidad de publicaciones por año, junto con la tendencia, para cada interrogante de investigación y la suma de publicaciones entre ambas. En ambos casos se puede observar cierta alza en la cantidad de publicaciones durante el año 2011, para luego verse reducida durante los años siguientes y volviendo a tener un aumento significativo en los años 2015 (para la interrogante respecto a servicios de e-government) y 2016 (para la interrogante sobre arquitecturas empresariales en e-government).

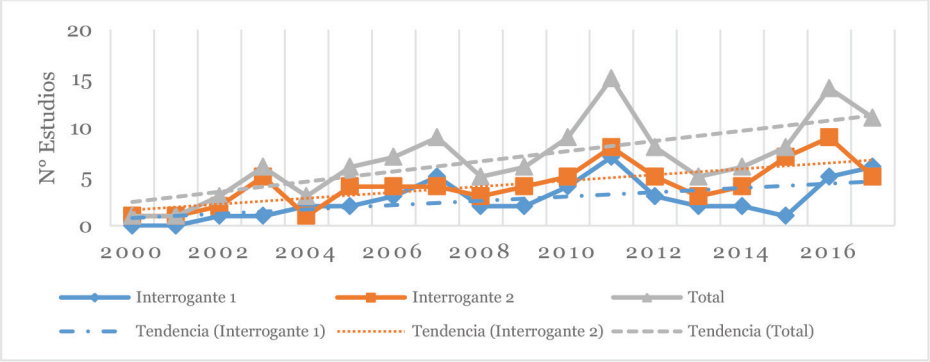


Figura 1 – Cantidad de publicaciones de e-government por año

A continuación, se muestran los resultados por cada interrogante establecida y se presentan los principales hallazgos de la investigación.

4.1. Uso de AEs en el contexto de e-government

En las siguientes subsecciones se detallan los resultados encontrados en el uso de AEs en el contexto de e-government.

4.1.1. Organizaciones cubiertas por las investigaciones

La Figura 2 muestra los estudios agrupados por tipo de organización involucrada, la mayoría de ellos se enfocan en el sector público en general, 26 estudios (54%), seguido de municipios, 15 estudios (31%).

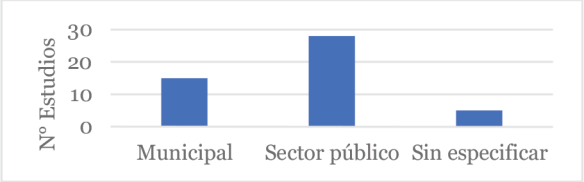


Figura 2 – Enfoque de los estudios de AE

4.1.2. Tipos de documentos sobre AEs

Respecto de los tipos de documentos encontrados, la mayoría de las publicaciones respecto a AEs en e-government corresponden a artículos de conferencia (27), seguido de artículos de revista (10), tesis (7) y capítulos de libro (4).

4.1.3. Uso de notaciones de modelado

Respecto del uso de notaciones de modelado de AEs. En la Tabla 6 se puede apreciar que en sólo 10 artículos (21%) se utiliza alguna notación para describir aspectos de la AE, entre ellas, BPMN es la notación de modelado más utilizada.

Referencia	BPMN	Casos de uso	ArchiMate	Diagramas varios
(Syynimaa, 2017)	x		x	
(Bach. Baez Medina, 2017)	x			
(Jiménez, 2016)	x			
(Aguilar Lugo, Díaz Cruz, & LLanos Sánchez, 2015)	x			
(Rodriguez Ortiz, 2012)	x	x		x
(Bouwman, Houtum, Janssen, & Versteeg, 2011)			x	
(Gallardo, 2011)	x			
(Koning, Bos, & Brinkkemper, 2008)				x
(Lankhorst & Derks, 2007)			x	
(Zuiderhoek et al., 2006)				x

Table 6 – Uso de notaciones de modelado en estudios de AE

4.1.4. Uso de framework en AEs

La Figura 3 muestra los estudios analizados en relación a los framework que son utilizados para el desarrollo de arquitecturas empresariales, se destaca que 25 artículos, 52% de los estudios, no especifica el framework que utilizan para el desarrollo de propuestas. Además, se observa que el 17% de los estudios usa el marco de arquitectura empresarial de gobierno (GEA).

4.1.5. Clasificación de las publicaciones sobre AE por país

Se analizó el país al que correspondía cada publicación, ver Tabla 7. Para facilitar su análisis, los países fueron posteriormente agrupados en base a la región geográfica a la

que pertenecen, lo cual puede ser observado en la Figura 4. Se destaca que la mayoría de los estudios sobre AE en e-government han sido realizados en la región de Europa, seguido de estudios realizados a nivel global, mientras que solo existe un estudio en la región de América del Norte.

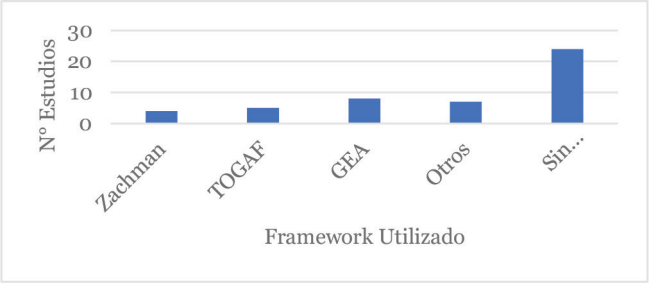


Figura 3 – Uso de framework en los estudios de AE

4.1.6. Clasificación de las publicaciones sobre AE por país

Se analizó el país al que correspondía cada publicación, ver Tabla 7. Para facilitar su análisis, los países fueron posteriormente agrupados en base a la región geográfica a la que pertenecen, lo cual puede ser observado en la Figura 4. Se destaca que la mayoría de los estudios sobre AE en e-government han sido realizados en la región de Europa, seguido de estudios realizados a nivel global, mientras que solo existe un estudio en la región de América del Norte.

País	Art.	País	Art.	País	Art.	País	Art.
Holanda	8	Indonesia	2	Hungría	1	Palestina	1
Finlandia	7	Suecia	2	Irán	1	Reino Unido	1
Perú	3	Corea del Sur	1	Italia	1	Multi-País	6
Australia	2	Dinamarca	1	México	1	Global	6
Colombia	2	Grecia	1	Noruega	1		

Table 7 – Cantidad de publicaciones sobre AE en e-government por país

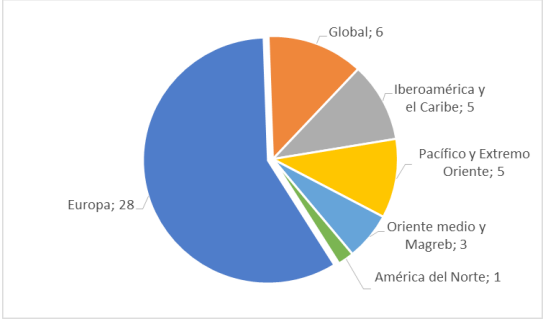


Figura 4 – Cantidad de publicaciones sobre AE en e-government por región geográfica

4.2. Servicios de e-government en la literatura

A continuación se presentan los resultados de Q2 (servicios de e-government).

4.2.1. Foco de los servicios de e-government

Cabe resaltar que la gran mayoría (82,7%) de los artículos identificados se enfocan únicamente en e-government local, mientras que solo un 4% se enfoca únicamente en e-government central. Por otro lado, un 13,3% de los estudios consideran tanto e-government local como central.

4.2.2. Tipos de documentos sobre servicios de e-government

Un total de 46 de los estudios relacionados a servicios de e-government son artículos de revista, seguido de 18 pertenecientes a artículos de conferencias, 5 capítulos de libro y 5 tesis, además de un único documento correspondiente a literatura gris (sin referato).

4.2.3. Clasificación de los estudios seleccionados sobre servicios de e-government

La Figura 5 muestra una clasificación de los estudios de acuerdo con el enfoque en que se abordaron los servicios de e-government. De las investigaciones seleccionadas la mayor fuente de servicios electrónicos en la literatura, que corresponde a 43 estudios (57%), proponen modelos de evaluación. Su objetivo, es delimitar el desarrollo de los sistemas de información en fases para determinar el grado de implementación de e-government. Por otro lado, los estudios que abarcan la implementación (19%), calidad (4%) y factores de éxito (9%), aportan servicios en menor cantidad de acuerdo a sus enfoques respectivos. Finalmente, la clasificación Otros (11%) incluye estudios variados que tratan temas como problemas y beneficios del uso de TICs, clasificación de servicios, entre otros.

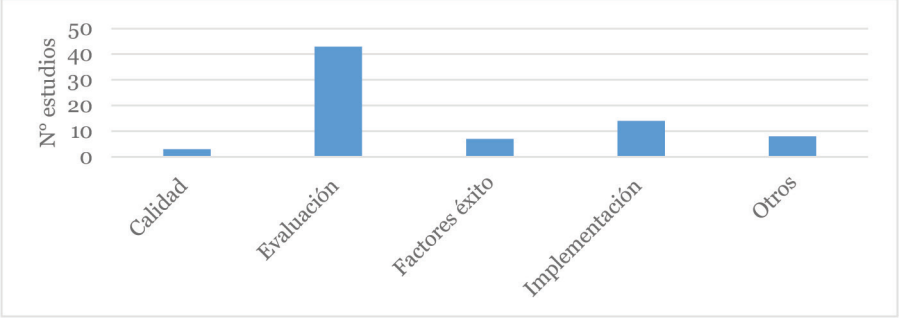


Figura 5 – Clasificación de los estudios sobre servicios en la literatura

4.2.4. Clasificación de los servicios de e-government

En base a la lectura de los artículos, se logró identificar una serie de clasificaciones utilizadas por los distintos autores para categorizar los distintos servicios de e-government que pueden ser ofrecidos. La mayoría de estas categorías tenían características similares entre sí por lo que las homologamos en las siguientes:

- Presencia: presencia en la Web y servicios donde se provee de información al ciudadano y la incorporación de esquemas de búsqueda básicas (Batlle Montserrat et al., 2011; Gelanzé Sevilla, 2014).
- Información: Servicios de relación unilateral donde se provee información perteneciente al gobierno local (Osio & Pineda, 2015; Gómez, 2007).
- Interacción: Se ofrecen servicios con los que se puede interactuar de forma simple con los gobiernos (Batlle Montserrat et al., 2011).
- Transacción: Interacción más sofisticada y de tipo transaccional que involucra autenticación y el pago de servicios en línea (Batlle Montserrat et al., 2011; Gelanzé Sevilla, 2014; Osio & Pineda, 2015).
- Participación: Servicios de participación democrática donde existe un único punto de acceso a todos los servicios (Gelanzé Sevilla, 2014; Osio & Pineda, 2015; Gómez, 2007).
- Automatización: El servicio de gobierno es provisto proactivamente a los ciudadanos, basándose en un registro previo de algún evento o del perfil del ciudadano (Fogli & Provenza, 2012).

En la Tabla 8 se muestran las clasificaciones identificadas y la cantidad de artículos que las usan. Cada clasificación se presenta asociada a las categorías previamente definidas. Cabe señalar que la única clasificación presentada con los nombres de sus categorías originales es la de (Esteves, 2005) (ID 5), utilizada 8 veces.

ID	Clasificación	Cantidad de Artículos
1	Información, Interacción, Participación	3
2	Información, Interacción, Transacción	10
3	Información, Interacción, Transacción, Automatización	3
4	Información, Interacción, Transacción, Participación	10
5	Presencia, Información, Interacción, Transacción, Participación	8
6	Otros	9
7	Sin Especificar	32

Table 8 – Clasificaciones de servicios de e-government

4.2.5. Clasificación de las publicaciones sobre servicios de e-government por país

Se analizó el país al que correspondía cada publicación, lo cual es presentado en la Tabla 9. Para facilitar su análisis, los países fueron posteriormente agrupados en base

País	Art.	País	Art.	País	Art.	País	Art.
Chile	9	Argentina	3	China	1	Reino Unido	1
Venezuela	6	Italia	3	Colombia	1	Tailandia	1
Estados Unidos	5	Grecia	2	Ecuador	1	Multi-País	6
Holanda	5	Perú	2	Eslovaquia	1	Global	7
Portugal	5	Suecia	2	Gales	1		
España	4	Turquía	2	Irlanda	1		
México	4	Bangladesh	1	Montenegro	1		

Table 9 – Cantidad de publicaciones sobre servicios de e-government por país

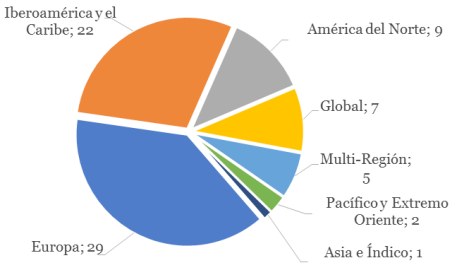


Figura 6 – Cantidad de publicaciones sobre servicios de e-government por región geográfica

a la región geográfica a la que pertenecen, lo cual puede ser observado en la Figura 6. Se destaca que la mayoría de los estudios sobre servicios de e-government han sido realizados en las regiones de Europa e Iberoamérica y el Caribe, mientras que muy pocos lo han sido las regiones del Pacífico y Extremo Oriente, y Asia e Índico.

4.2.6. Servicios de e-government

Como se mencionó anteriormente, la mayor fuente de servicios electrónicos en la literatura (57%), son artículos que buscan determinar el grado de implementación de e-government a través de modelos de evaluación. A partir de las clasificaciones del apartado anterior se extrajo un total 29 servicios que se dividieron en 4 categorías que corresponden a 14 servicios informativos, 5 de interacción, 7 de transacción y 3 de participación los cuales se muestran en detalle en la Tabla 10.

Categorías	Trabajos encontrados por servicio en cada categoría
Informativos	Información general (21); Información sobre salud pública (5); Presencia redes sociales (2); Información sobre educación y cultura (9); Información sobre deporte (4); Información sobre servicios sociales (3); Búsqueda de empleo (7); Transparencia (9); Información sobre transportes (10); Información urbana (11); Información estadística (6); Información sobre seguridad (4); Asamblea municipal (9) Documentos (16)
Interacción	Foros de discusión (10); Contacto (17); Noticias (6); Debates (2) y Retransmisión de los plenos (5)
Transacción	Reserva o compra de tickets (3); Población (9); Territorio, urbanismo y obras (6); Certificados (16); Seguimiento (7); Circulación, vehículos y transportes (11) y Comercio, industria y consumo (18)
Participación	Registro (4); Votación electrónica (6) y Participación ciudadana (11)

Table 10 – Tipos de servicios de e-government

4.3. Principales hallazgos

Se encontró que sólo 15 estudios utilizan AEs para apoyar el desarrollo de e-government municipal por lo que se amplió la búsqueda al sector público. Un aspecto relevante es la

falta de uso de notaciones de modelado en el desarrollo de AEs, donde solo 10 estudios utilizan alguna como BPMN, casos de uso, diagramas o ArchiMate, siendo esta última una de las principales notaciones para modelar AEs.

Los estudios que identifican servicios electrónicos en la literatura y que responden de mejor manera a la interrogante Q2 son los que se detallan en la sección 4.2.4 y 4.2.6. Estos estudios proporcionan la mayor fuente de servicios electrónicos en la literatura y corresponde al 57% de los estudios seleccionados. En general buscan determinar el grado de implementación de e-government a través de modelos de evaluación. Estos modelos proporcionan por cada etapa los tipos de servicios de e-government con los que deberían contar los gobiernos.

5. Conclusiones

En este artículo se desarrolló un Mapeo Sistemático de la Literatura, para identificar material bibliográfico que permita conocer el uso de Arquitecturas Empresariales asociadas al desarrollo de servicios municipales en el contexto de e-government local. Como resultado se obtuvieron y analizaron 123 artículos, de los cuales 48 corresponden a AEs y 75 abordan servicios de e-government.

A partir de los resultados hemos podido llegar a varias conclusiones. En primer lugar, respecto a las organizaciones cubiertas por las investigaciones, se destaca la escasez de trabajos (15) sobre el uso AEs para apoyar el desarrollo de e-government municipal. En segundo lugar, se evidencia la falta de uso de notaciones para el modelado de AEs, donde sólo 3 de los nueve estudios que especifican el uso de alguna notación usan una dedicada a representar arquitecturas empresariales. En tercer lugar, en relación a los servicios de e-government identificados en la literatura, se extrajo un total 29 servicios que se dividieron en las categorías de información, interacción, transacción y participación.

Identificada la forma en que se usan las AEs, como trabajo futuro, se espera estandarizar su uso en relación a un conjunto de servicios determinado que permitan a cualquier municipio acceder a una arquitectura de referencia.

Referencias

Aguilar Lugo, K. Y., Díaz Cruz, A. M., & LLanos Sánchez, A. M. (2015). Propuesta de modelo de arquitectura de negocio usando un marco de arquitectura empresarial para una entidad pública, 269 pags.

Ask, A., & Hedström, K. (2011). *Taking initial steps towards enterprise architecture in local government*. International Conference on Electronic Government and the Information Systems Perspective, 26–40.

Bach. Baez Medina, Erick Ricardo. 2017. “Diseño de Arquitectura Empresarial En La Municipalidad Distrital de La Victoria–Chiclayo.”

Batlle Montserrat, J., Abadal, E., & Blat, J. (2011). Benchmarking del e-gobierno local: limitaciones de los sistemas de evaluación comparativa. *El Profesional de la Informacion, 2011, 20(3), 251–259*.

Bouwman, H., Houtum, H. V., Janssen, M., & Versteeg, G. (2011). Business Architectures in the Public Sector: Experiences from Practice. *cais 29*, 23.

Carter, L., & Bélanger, F. (2005). The utilization of e-government services: citizen trust, innovation and acceptance factors. *Information systems journal, 15*(1), 5–25.

Dias, G. P. (2011). *Local e-government information and service delivery*. Paper presented at the Information Systems and Technologies (CISTI), 2011 6th Iberian Conference on, 1-6, IEEE.

Esteves, J. (2005). *Análisis del desarrollo del gobierno electrónico municipal en España*. Instituto de Empresa, 5–32.

Fogli, D., & Provenza, L. P. (2012). A meta-design approach to the development of. *Journal of Visual Language and Computing, 23*(2), 47–62.

Gallardo, H. (2011). Mejora De Servicios De Una Entidad Del Estado Utilizando Six Sigma Y Arquitectura Empresarial. Caso servicio de pensionamiento Onp. 201.

Gelanzé Sevilla, F. (2014). Gobierno electrónico municipal en el Estado Carabobo. República Bolivariana de Venezuela. *Memoria Política, 3*, 253–268.

Genero, M., Cruz-Lemus, J., & Piattini, M. (2014). Métodos de investigación en ingeniería del software. *Editorial RA-MA: Madrid, Spain*, 171–199.

Gómez, P. (2007). El gobierno electrónico en el municipio venezolano. Enlace, 4(2) 67–80.

Hornnes, E., Jansen, A., & Langeland, Ø. (2010). How to Develop an Open and Flexible Information Infrastructure for the Public Sector?. International Conference on Electronic Government , 301–314.

Janssen, M., & Wagenaar, R. (2003). Towards a Flexible ICT-Architecture for Multi-Channel E-Government Service Provisioning. *System Sciences, 2003. Proceedings of 36th Annual Hawaii International Conference*, 1–10.

Jiménez, J. (2016). Diseño de un Modelo para la Creación de Secretaría TIC en Entes Territoriales Colombianos de Categoría 1 y 2, Basado en Arquitectura Empresarial. 134.

Kitchenham, B., Brereton, O., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering–a systematic literature review. *Information and Software Technology, 51*(1), 7–15.

Koning, H., Bos, R., & Brinkkemper, S. (2008). A Lightweight Method for the Modelling of Enterprise Architectures. *International Conference on Service-Oriented Computing,* 375-387, *Springer, Berlin, Heidelberg*.

Kotusev, S. (2017). Enterprise Architecture: What Did We Study?. *International Journal of Cooperative Information Systems, 26*(04), 1730002.

Lankhorst, M. M., & Derks, W. L. A. (2007). Towards a Service-Oriented Architecture for Demand-Driven e-Government. *11th IEEE International Enterprise Distributed Object Computing Conference*, 214–225.

Lankhorst, M. M., Proper, H. A., & Jonkers, H. (2010). The anatomy of the ArchiMate language. *International Journal of Information System Modeling and Design (IJISMD), 1*(1), 1–32.

Moon, M. J. (2002). The evolution of e-government among municipalities: rhetoric or reality? *Public administration review, 62*(4), 424–433.

Navarro, C. C., Cañavate, A. M., & Bleda, A. M. (2008). La gestión de información en las sedes web de los ayuntamientos españoles. *Revista española de documentación científica, 31*(4), 612–638.

Osio, L., & Pineda, P. L. (2015). Evolución del e-gobierno en la Alcaldia del Municipio Naguanagua, estado Carabobo: periodo 2004-2014. *Enl@ce: Revista Venezolana de Informacion, Tecnologia y Conocimiento, 12*(2), 96–114.

Poggi, E. (2013). Gobierno electrónico, gobierno local y gestión tecnológica. *Pando, D. y Fernández Arroyo, N.(comp.), El gobierno electrónico a nivel local. Experiencias, tendencias y reflexiones*, 71–95.

Rocha, A., & Sá, F. (2014). Planning the information architecture in a local public administration organization. *Information Development*, 30(3), 223–234.

Rodriguez Ortiz, I. (2012). Infraestructura Tecnológica y Sistemas de Información para la Oferta de Servicios de Gobierno Electrónico a Nivel Municipal. 65.

Sandoval-Almazán, R., & Colín, J. M. (2011). *Gobierno electrónico en México: una exploración Municipal 2010*. Paper presented at the AMCIS.

Syynimaa, N. (2017). Method and Practical Guidelines for Overcoming Enterprise Architecture Adoption Challenges. *ICEIS*, 488–514.

Sá, F., Rocha, Á., & Cota, M. P. (2016). Potential dimensions for a local e-Government services quality model. *Telematics and Informatics, 33*(2), 270-276.

The Open Group. (2017). ArchiMate® 3.0.1 Specification.

Zuiderhoek, B., Otter, A., Bos, R., & Brinkkemper, S. (2006). Framework for Dutch municipalities to ensure business IT alignment using enterprise architecture. *Proceedings of the 6th European Conference on e-Government, 457–466*.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 05/09/2018
Aceitação/Acceptance: 27/11/2018

Tecnologias educacionais para cegos e tecnologias emergentes: um estudo bibliográfico

Graziela de Souza Sombrio[1], Vania Ribas Ulbricht[2]

graziela.sombrio@gmail.com, vrulbricht@gmail.com

[1] Instituto Federal de Santa Catarina, Xanxerê, 89820-000, SC, Brasil

[2] Universidade Federal de Santa Catarina, Florianópolis, 88037-130, SC, Brasil

Pages: 830–840

Resumo: Este artigo discute a utilização de tecnologias no processo de ensino e aprendizagem de cegos. Para isto, foi realizada uma revisão sistemática de literatura com o objetivo de identificar as tecnologias que vem sendo utilizadas. Com uma análise dos relatórios do Horizon Report para a Educação Básica, foi possível verificar quais das tecnologias encontradas estão dentre as tecnologias emergentes. Por fim, são sugeridas tecnologias emergentes presentes nos relatórios que podem ser utilizadas no ensino e aprendizagem de cegos.

Palavras-chave: Cegos; tecnologias; inclusão; Horizon Report.

Educational technologies for blind people and emerging technologies: a bibliographic study

Abstract: This article discusses the use of technologies in the teaching and learning process of the blind. For this, a systematic literature review was carried out with the objective to identify the technologies that have been used. With an analysis of the reports of the Horizon Report for Basic Education, it was possible to verify which of the technologies found are among emerging technologies. Finally, emerging technologies are suggested in the reports that can be used in teaching and learning the blind.

Keywords*: Blind; Technologies; inclusion; Horizon Report.*

1. Introdução

O censo do Instituto Brasileiro de Geografia e Estatística – IBGE (2012) publicado em novembro de 2011 indica que no Brasil existem 45,6 milhões de pessoas com algum tipo de deficiência, o que representa 23,9% da população. Destes, mais de 35,8 milhões tem deficiência visual, dentre os quais 6,6 milhões tem deficiência visual severa e quase 730 mil são cegos.

No Brasil, a matrícula de pessoas com deficiência em escolas regulares é garantida pela legislação (Constituição Federal (Brasil, 1988), pela Lei de Diretrizes e Bases da

Educação Nacional – LDB (Brasil, 1996) e pela Lei no 13.146, de 6 de julho de 2015, que trata da inclusão da pessoa com deficiência (Brasil, 2015)). Como consequência, o número de alunos com deficiência matriculados em escolas regulares tem aumentado muito nos últimos anos.

Atualmente, dentre as escolas brasileiras, 57,8% têm alunos com deficiência, transtornos globais do desenvolvimento ou altas habilidades, incluídos em classes comuns. Esse percentual, em 2008, era de apenas 31% (INEP/MEC, 2016). Embora esses dados tragam informações quantitativas quanto ao número de matrículas, eles não garantem a qualidade no processo de ensino e aprendizagem.

Para garantir a efetiva inclusão das pessoas com deficiência na escola regular, é necessário que a realidade desse aluno seja conhecida e as especificidades de aprendizagem para cada tipo de deficiência seja considerada. As práticas pedagógicas, bem como os materiais utilizados, precisam ser adaptadas de acordo com essas especificidades.

2. O cego e a tecnologia

Um dos problemas encontrados pelos cegos, é que eles recebem a significação das coisas pelos videntes, os quais utilizam a visão como principal fonte de conhecimento (Nunes, 2004). Sendo assim, o cego está em constante conflito, precisando aliar as suas percepções com as informações que lhes são passadas. As tecnologias podem ser utilizadas de forma a proporcionar mais autonomia para esse público diminuindo assim, a dependência de auxílio dos videntes, até mesmo para tarefas cotidianas.

Dentre as tecnologias existentes, alguns aplicativos para celular foram desenvolvidos com o objetivo de contribuir com a melhoria da qualidade de vida dos cegos. O *BlindTool*, por exemplo, é um reconhecedor de objetos. Para utilizá-lo basta apontar o celular para seu entorno. O mesmo vibrará quando encontrar um objeto reconhecível e o aplicativo verbalizará qual objeto encontrou. O *Be my eyes* é um aplicativo que possibilita que pessoas videntes auxiliem os cegos a identificar coisas pontuais. Pessoas videntes se cadastram como voluntários no sistema e, ao ser chamados pelo telefone através do aplicativo, podem descrever o que as pessoas cegas mostram através da câmera do celular. O aplicativo *ColorID* reconhece os variados tons de cores e verbaliza-os para o usuário. O objetivo é auxiliar em tarefas como reconhecer as cores das roupas e identificar, por exemplo, se uma fruta está ou não madura. Desenvolvido por alunos do curso de Engenharia da Computação da Universidade Federal do Pará, o *BusID* é um aplicativo que identifica a linha de um ônibus na cidade de Belém – PA, e verbaliza para o usuário. Com isso, o cego consegue mais independência para se locomover utilizando o transporte público da cidade.

Enquanto várias ações são feitas para facilitar a vida diária do cego, o mesmo ainda se sente excluído durante a vida escolar. Segundo Takimoto (2014), há uma falta de comunicação entre aluno/professor e uma falta de preparo por parte dos professores para lidar com a situação.

De forma geral, há pouca informação sobre o cego, tanto na escola, quanto na família e na sociedade. A escola não está preparada para receber o aluno com deficiência. É preciso que quem ensina, compreenda o processo de aprendizagem e as necessidades

educacionais dos cegos, considerando sempre que esse público não tem referências visuais.

É importante considerar que a aprendizagem não é limitada pela deficiência. A limitação, de acordo com Obregon, Vanzin e Ulbricht (2011), ocorre pela não percepção da singularidade e das necessidades desse público. As dificuldades encontradas e os baixos resultados que possam ocorrer no processo de ensino e aprendizagem, de acordo Smole (2000), devem servir para guiar no aprimoramento das capacidades existentes. Afinal, o potencial para aprendizagem e desenvolvimento dessas crianças são normais e é necessário que deficiências secundárias sejam evitadas (Batista & Enumo, 2000). As barreiras são criadas, muitas vezes, pelo fato de que o mundo é estruturado no padrão de normalidade, tornando o processo de aprendizagem altamente excludente (Obregon, Vanzin, & Ulbricht, 2011). Para evitar que isso ocorra, é preciso que, em todas as etapas da vida, profissionais auxiliem nesse desenvolvimento e considerem que, de acordo com Cunha (1997), é a partir da fala e da linguagem que a criança com deficiência visual controla o que está fora de seu alcance visual, sendo esses fatores essenciais para o seu desenvolvimento.

Quanto às tecnologias, a utilização de sistemas e-learning por pessoas com deficiência é desafiadora para pesquisadores e educadores. Um dos obstáculos encontrados é a natureza gráfica desses sistemas. De acordo com Obregon, Vanzin e Ulbricht (2011), os ambientes virtuais estão descontextualizados e não atendem as especificidades dos usuários com deficiência visual. Isto ocorre pelo fato de que a sociedade elege, devido ao padrão cultural imposto "a visão e a audição como capacidades essenciais para atuar em contextos virtuais" (Obregon, Vanzin, & Ulbricht, 2011, p. 62).

3. Metodologia

Este artigo faz uma análise das tecnologias utilizadas no processo de ensino e aprendizagem de cegos. Para identificar tais tecnologias, foi realizada uma revisão sistemática de literatura.

A Revisão Sistemática de Literatura, segundo Galvão e Pereira (2014),tem como objetivo identificar, selecionar, avaliar e sintetizar as evidências relevantes disponíveis sobre um determinado tema. Para este artigo, a busca foi realizada no Portal de Periódicos Capes. O Portal de Periódicos da Capes é uma biblioteca virtual com um acervo de mais de 38 mil títulos com texto completo, 134 bases referenciais, 11 bases dedicadas exclusivamente a patentes, livros, enciclopédias, normas técnicas, estatísticas e conteúdo audiovisual[1].

A busca foi realizada utilizando os termos "*technology AND learning AND blind*". O quadro abaixo mostra como foram realizadas as constrições para seleção dos artigos que possuem relação com o conteúdo desse artigo.

Após as etapas citadas na Tabela 1, 11 artigos foram lidos na íntegra. Apenas 3 deles, por contribuírem com o tema pesquisado, foram incluídos no texto.

[1] https://www.periodicos.capes.gov.br/index.php?option=com_pcontent&view=pcontent&alias=missao-objetivos&Itemid=126

Constrição utilizada	Quantidade de artigos encontrados
Quantidade encontrada no Periódico Capes	110.414
Periódicos revisados por pares	49.987
Apenas os artigos	48.416
Publicados entre os anos de 2007 e 2018	32.654
Tópicos incluídos: learning, blinds, blindness, assistive technology	2.773
Tópicos incluídos: education, students, educational technology, visual impairments	746
Tópicos incluídos: teaching methods, teachers, teaching, disabled students, special education	353
Pré-selecionados os 5% mais relevantes	18
Retirados os repetidos	17
Analisados os resumos	11

Fonte: produzido pelos autores

Tabela 1 – Resultado da revisão sistemática de literatura

4. Resultados

A revisão sistemática de literatura resultou na seleção de 3 artigos que contribuíram com a presente pesquisa. Esses artigos apresentam tecnologias que estão sendo utilizadas por cegos no contexto escolar.

A tabela 2 apresenta uma síntese com as tecnologias utilizadas, bem como os pontos positivos e negativos de cada uma delas.

Título	Autores	Tecnologia apresentada	Pontos positivos	Pontos negativos
Revealing the whiteboard to blind students: An inclusive approach to provide mediation in synchronous e-learning activities	André P. Freire; Flávia Linhalis; Sandro L. Bianchini; Renata P.M. Fortes; Maria da Graça C. Pimentel	Software IWB: ferramenta para usuários remotos compartilharem anotações durante e depois da aula ou em sessões de brainstorming	- Captura as anotações do instrutor durante uma aula ou palestra. - Pode ser utilizado em tablets, PCs e computadores pessoais. - Permite ao usuário utilizar fontes diferentes, formas geométricas pré-definidas. - Disponíveis as opções de copiar, recortar, colar, mover e apagar objetos. - É possível refazer e desfazer, criar novos slides, duplicar slides existentes ou navegar pelos slides.	- A utilização do "Tab" foi um dos principais problemas identificados. - A usabilidade poderia ser melhorada (JAVA). - O professor passa a se preocupar menos com a uma atenção direta e especial aos alunos cegos.

Título	Autores	Tecnologia apresentada	Pontos positivos	Pontos negativos
Enabling People Who Are Blind to Experience Science Inquiry Learning through Sound-Based Mediation	S.T. Levy; O. Lahav	Listening to Complexity: é um ambiente de aprendizagem baseado em mediação sonora	- Pode apoiar a aprendizagem de química por pessoas que são cegas. - São utilizadas as sonorizações auditivas alertas e ecrãs de áudio espaciais. - A sonificação melhora a aprendizagem em ambientes interativos. - Facilita a interação entre cegos e videntes.	- Falta de limitação da quantidade de informações que podem ser integradas em uma única audição. - Falta a opção de retomar os conteúdos nos quais foram identificadas dificuldades. - Não utiliza feedback apropriado instruções explícitas para conduzir a uma retomada dos conceitos. - O ambiente não pode ser considerado exploratório pois não é controlado pelo usuário.
i-Math: Automatic math reader for Thai blind and visually impaired students	Wararat Wongkiaa; Kanlaya Naruedomkulb Nick Cercone	i-Math: ferramenta educacional para estudantes cegos e deficientes visuais (VI), para facilitar o acesso a materiais de matemática	- É um leitor de tela, que produz saída de voz em um computador, capaz de ler documentos matemáticos em voz alta. - Proporciona a possibilidade de estudar e praticar matemática de maneira independente e confortável. - Não é necessário conhecer o Braile.	- Dificuldade em lembrar tudo o que foi verbalizado pelo i-Math. - Proporcionou desconforto em alguns usuários. - Não permite pausar a verbalização em andamento. - As pronúncias ainda não são ideais.

Tabela 2 – Síntese das tecnologias encontradas na RSL

Freire, Linhalis, Bianchini, Fortes, & Pimentel (2010) apontam os leitores de tela como a principal tecnologia assistiva utilizada para a aprendizagem. Entretanto, os autores afirmam que imagens não são lidas pelos leitores. Dessa forma, recursos gráficos que muitas vezes são utilizados como forma de facilitar o ensino e aprendizagem, se tornam dificultadores desse processo quando o aluno for deficiente visual.

Uma tecnologia que tem sido utilizada como forma de promover a interação assíncrona entre alunos, professores e outras pessoas envolvidas no processo de ensino e aprendizagem, são as lousas interativas. O software *IWB* permite registrar anotações

de um professor e revisá-las posteriormente. Fornece opções de utilização de diferentes cores e espessuras de tinta, formas geométricas predefinidas, possibilidade de copiar, recortar, cola, mover e apagar objetos, refazer e desfazer, criar novos slides, duplicar splides e navegar pelos já existentes. Com essas possibilidades, o software *IWB* é uma ferramenta muito útil para usuários remotos compartilharem anotações durante e depois da aula ou em sessões de brainstorming.

O objetivo de Freire, Linhalis, Bianchini, Fortes, & Pimentel (2010) é adaptar essa tecnologia para que seja utilizada por cegos. Para isso, propuseram que os alunos com deficiência visual participassem de sessões de aprendizagem síncronas, à distância, com um mediador. A figura 1 mostra a lousa interativa sendo utilizada.

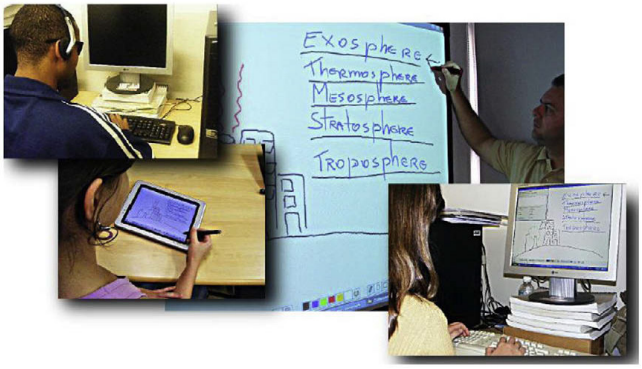


Figura 1: Cenário de uso do IWB, incluindo alunos com diferentes habilidades e tecnologias. À esquerda, no sentido horário: aluno cego ouvindo o texto adicionado pelo mediador, instrutor usando o software IWB, mediador adicionando texto e aluno observando as anotações do quadro usando um tablet PC.
Fonte: Freire, Linhalis, Bianchini, Fortes, & Pimentel (2010)

Na avaliação feita pelos autores, os mesmos puderam observar que, tanto professor quanto mediador, consideram crucial que ambos discutam anteriormente o conteúdo que será abordado na aula. Os pontos positivos dessa tecnologia estão apresentados na tabela 2.

Um fator preocupante, que consta como ponto negativo na tabela 2, é o professor considerar que ele não precisa mais se preocupar tanto com o aluno cego durante a aula. Embora a tecnologia facilite o processo de ensino e aprendizagem, o professor não pode desconsiderar a presença de um aluno cego em sala de aula.

Uma outra tecnologia utilizada é o *Listening to Complexity'1* (L2C). Como forma de compensar a falta de visão, a audição é utilizada como apoio para a aprendizagem de sistemas complexos na área de ciências. O estudo feito por Levy & Lahav (2012) investigou maneiras pelas quais os ambientes de aprendizagem baseados em mediação sonora podem apoiar os cegos na aprendizagem de química. Foi utilizado o conceito de sonificação, que é a apresentação de informações utilizando som, porém sem fala (Kramer , 1994). Sonorizações auditivas, de acordo com Nees e Walker (2009), são classificadas em:

1. Alertas
2. Objeto e indicadores de status, menus auditivos
3. Representação de dados
4. Ecrãs de áudio espaciais
5. Paisagens sonoras e som de fundo
6. Artes e entretenimento

No ambiente L2C são utilizadas as classes de (1*)* a (4*)*. Como resultado, o estudo mostra melhorias no ensino de ciências. Os autores recomendam que a quantidade de informações que podem ser integradas em um único áudio, seja limitada.

Quanto aos pontos negativos, o fato de o ambiente não ser exploratório, não permitindo que o usuário retome conteúdos apresentados anteriormente, é limitante para o processo de ensino e aprendizagem. Feedbacks podem ser utilizados para que o usuário explore melhor o ambiente e, consequentemente, melhore a aprendizagem.

O terceiro artigo analisado traz a tecnologia *i-Math*, que é uma ferramenta educacional que faz leituras de expressões matemáticas, na língua tailandesa. O objetivo é melhorar a acessibilidade de materiais de matemática para deficientes visuais (Ramos & De Andrade, 2016). A avaliação feita pelos autores mostra que o *i-Math* pode ser utilizado por aqueles que não conhecem o Braille, facilitando o processo de ensino e aprendizagem. Também propicia uma maior autonomia para alunos cegos no cumprimento das tarefas escolhidas.

Embora a ferramenta verbalize expressões matemáticas, há dificuldade por parte do usuário em lembrar da expressão completa. Como fator dificultador, tem-se ainda o fato de que não é possível pausar a verbalização.

5. Discussão

A *New Media Consortium - NMC*, formada por especialistas em tecnologia educacional, desde 2002 tem trabalhado no *Projeto Horizon*. Anualmente, três relatórios são publicados: Educação superior, Educação Básica (K-12) e Museus. Neles estão identificadas as principais tendências e as tecnologias emergentes que trarão impacto na educação ao longo dos 5 anos seguintes.

O relatório para a Educação Básica surgiu em 2009, nos mesmos moldes dos relatórios voltados para a Educação Superior, já publicados. Analisando cada um dos relatórios publicados entre 2009 e 2017, foi possível identificar as tecnologias constantes dos relatório e que estão mostradas na Figura 2.

As tecnologias emergentes que são apresentadas nos relatórios *HR* são tendências feitas para a educação em um tempo máximo de 5 anos. Entretanto, estudos posteriores, indicando se essas tendências se concretizaram ou não, não são realizadas. Estudos mostram, de forma independente, a utilização de cada tecnologia, não sendo possível verificar qual a mais utilizada, por exemplo.

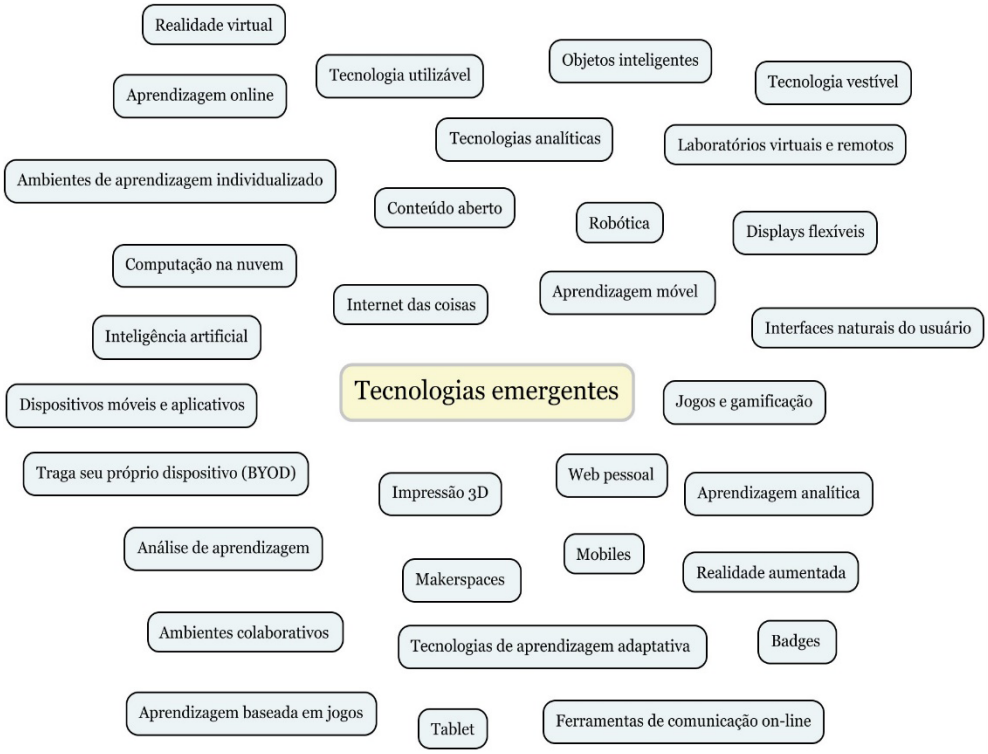


Figura 2: Tecnologias emergentes – Horizon Report – K-12
Fonte: produzida pelos autores

Dentre as tecnologias encontradas nos artigos selecionados na revisão sistemática de literatura, a lousa interativa, da forma como foi utilizada nos estudos feitos por Freire, Linhalis, Bianchini, Fortes, & Pimentel (2010), é uma das tecnologias emergentes citadas no Horizon Report. É possível afirmar que se trata de um ambiente colaborativo, pois cada uma das pessoas envolvidas no processo de ensino e aprendizagem pode contribuir com a produção do material utilizado. Um ambiente colaborativo educacional é um local virtual de trabalho, no qual alunos e professores podem se comunicar, compartilhar informações e, até mesmo, trabalhar juntos. Favorece meios para que alunos e professores trabalhem com colegas, que podem ou não estarem no mesmo ambiente, que trabalhem de forma criativa e se relacionem com seus pares. O trabalho colaborativo na educação é considerado moderno e tem levado muitos professores a buscarem ferramentas que facilitem o trabalho em equipe e a interação entre seus membros. O uso do tablet também está presente, como pode ser visto na Figura 1. O tablet, dispositivo onde o software IWB pode ser utilizado, é uma das tecnologias emergentes apresentadas. O *Horizon Report* aponta o tablet como uma oportunidade de melhorar as experiências de aprendizagem. Tem como vantagem a utilização de aplicativos disponíveis para dispositivos menores,

mas em formato maior, e a não interrupção das atividades por fatores como telefone tocando ou mensagens chegando.

As outras duas tecnologia encontradas na revisão sistemática e selecionadas para esse artigo, não estão entre as tecnologias emergentes citadas no *Horizon Report*. Entretanto, elas poderiam ser melhoradas, através da integração com tecnologias emergentes.

O *Listening to Complexity* poderia ser melhorado com a utilização das tecnologias realidade virtual, *badges* e aprendizagem baseada em jogos. Com isso o ambiente seria mais exploratório e poderia ser controlado pelo usuário, o que o manteria motivado. Feedbacks, cuja falta foi apontada como ponto negativo do ambiente, são comuns em aprendizagem baseada em jogos e podem auxiliar na forma como o ambiente é explorado

O *i-Math*, por ser um leitor capaz de ler expressões matemáticas, deixaria de ser caso adotasse uma dentre a maioria das tecnologias emergentes. Entretanto, sua utilização em tablets ou outros dispositivos móveis, que são tecnologias que fazem parte do *Horizon Report*, deve ser incentivada.

Muitas das tecnologias emergentes presentes nos relatórios do *Horizon Report* poderiam ser utilizadas no processo de ensino e aprendizagem de pessoas com deficiência, incluindo os cegos. Como exemplo podemos citar a impressora 3D e as Interfaces naturais de usuários.

A impressora 3D pode ser utilizada para que, a partir de objetos impressos, o cego possa fazer uso do tato para conhecer o que está sendo apresentado. Ela tem sido utilizada para que cegos conheçam objetos fora do contexto escolar, como por exemplo, obras de arte [2] e fotos[3]. No contexto escolar, ela pode ser utilizada para a impressão de materiais para as aulas de geometria, de história, de geografia, entre outros.

Já as interfaces naturais permitem que os computadores respondam a gestos, movimentos do corpo, expressões faciais, voz, som e outros movimentos corporais, e estão substituindo o teclado e o mouse como padrão para interação entre computador e humanos. As várias tecnologias que permitem interfaces de usuário naturais estão tornando as interações com dispositivos computacionais muito mais intuitivas e, muitas vezes, tão simples que não são necessárias instruções para usá-las. O dispositivo como deve ser a interação homem-máquina. Devido às suas características, pode ser utilizado por pessoas com deficiência, incluindo os cegos.

6. Considerações finais

As tecnologias estão cada vez mais presentes no contexto escolar. Os relatórios do Horizon Report apontam as tecnologias emergentes para a educação para um prazo de 5 anos, mas não tem como foco a inclusão das pessoas com deficiência. Entretanto, várias das tecnologias presentes nos relatórios podem ser utilizadas com esse propósito.

[2] https://www.hypeness.com.br/2015/11/impressora-3d-permite-que-cegos-apreciem-obras-de-arte-classica-pela-primeira-vez/

[3] http://blog.fazedores.com/deficientes-visuais-podem-ver-suas-fotos-impressas-em-3d/

O estudo feito nesse artigo traz uma busca sistemática por tecnologias utilizadas no processo de ensino e aprendizagem de cegos, com a qual, foi possível observar que a utilização das tecnologias emergentes ainda é muito tímida. É preciso um olhar mais atencioso dos pequisadores na área de inclusão para essas tecnologias, pois elas abrem um grande leque de possibilidades para o contexto escolar.

Referências

Batista, C. G., & Enumo, S. R. (2000). Desenvolvimento humano e impedimentos de origem orgânica: o caso da deficiência visual. Em H. A. Novo, & M. C. Meneandro, *Olhares diversos: estudando o desenvolvimento humano* (pp. 157-174). Vitória: Capes.

Brasil. (1988). *Constituição da República Federativa do Brasil de 1988*. Fonte: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm

Brasil. (1996). *LEI Nº 9394, de 20 de dezembro de 1996*. Fonte: Estabelece as diretrizes e bases da educação nacional.: http://www.planalto.gov.br/ccivil_03/leis/L9394.htm

Brasil. (2015). *Lei nº 13.146, de 6 de julho de 2015*. Fonte: Insitutui a lei Brasileira de Inclusão da Pessoa com Deficiência (Estatuto da Pessoa com Deficiência): http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/Lei/L13146.htm

Cunha, A. B. (Agosto de 1997). Promovendo aquisição de linguagem funcional em criança deficiente visual: o efeito de um treinamento de mãe em procedimentos de ensino naturalístico. *Temas em psicologia, 5*(2), 33-56.

Freire, A. P., Linhalis, F., Bianchini, S. L., Fortes, R. P., & Pimentel, M. C. (2010). Revealing the whiteboard to blind students: An inclusive approach to provide mediation in synchronous e-learning activities. *Computer & Education, 54*, pp. 866-876.

Galvão, T. F., & Pereira , M. G. (Jan/Mar de 2014). Revisões sistemáticas de literatura: passos para sua elaboração. *Epidemiologia e Serviços de Saúde, 23*(1), 183-184.

IBGE. (2012). *Censo demográfico 2010: Características Gerais da População, Religião e Pessoas com Deficiência*. Rio de Janeiro: Ministério do Planejamento, Religião e Pessoas com Deficiência.

INEP/MEC. (2016). *Censo Escolar 2016*. INEP, Brasília. Acesso em 24 de Junho de 2017, disponível em INEP: http://portal.inep.gov.br/web/guest/resultados-e-resumos

Kramer , G. (1994). An introduction to auditory display. Em *Auditory Display: Sonification, Audification, and Auditory* (Vol. 18, pp. 1-77). MA: Addison-Wesley.

Levy, S., & Lahav, O. (2012). Enabling People Who Are Blind to Experience Science Inquiry Learning through Sound-Based Mediation. *Journal of Computer Assisted Learning, 28*(6), pp. 499-413.

Nees, M., & Walker, B. (2009). Auditory interfaces and sonification. Em C. Stephanidis, *The Universal Access Handbook* (pp. 507-521). New York: Lawrence Erlbaum Associates.

Nunes, S. d. (2004). *Desenvolvimento de conceitos em cegos congênitos: caminhos de aquisição do conhecimento*. Dissertação (Mestrado), Dissertação (Mestrado em Psicologia) - Instituto de Psicologia, Universidade de São Paulo, São Paulo.

Obregon, R. A., Vanzin, T., & Ulbricht, V. R. (Maio/Agosto de 2011). Ambientes virtuais de aprendizagem e a inclusão de usuários com deficiência visual. *Educaonline, 5*(2), 51-66.

Piekas, M. (2017). *Elementos da linguagem visual e ensino de desenho para crianças cegas*. Tese (Doutorado em Artes Visuais) - Centro de Artes, Universidade Federal de Santa Catarina, Florianópolis.

Ramos, S. M., & De Andrade, A. V. (2016). ICT in Portuguese Reference Schools for the Education of Blind and Partially Sighted Students. *Education and Information Technologies, 21*(3), pp. 625-641.

Takimoto, T. (2014). *A percepção do espaço tridimensional e sua representação bidimensional: a geometria ao alcance das pessoas com deficiência visual em comunidades virtuais de aprendizagem*. Dissertação (Mestrado), Dissertação (Mestrado em Engenharia e Gestão do Conhecimento) - Universidade Federal de Santa Catarina, Florianópolis.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 26/10/2018
Aceitação/Acceptance: 05/12/2018

Arquitectura IoT para la identificación de personas en entornos educativos

Darly Mildred Delgado Agudelo[1], Diego Fernando Girón Timaná[2], Gabriel Elías Chanchí Golondrino[3], Katerine Márceles Villalba[4]

darlydelgado@unimayor.edu.co, dfgiron@unimayor.edu.co, gchanchi@unimayor.edu.co, kmarceles@unimayor.edu.co

[1, 2, 3, 4] Facultad de Ingeniería, Institución Universitaria Colegio Mayor del Cauca, Carrera 7 N° 2-34 Edificio Bicentenario, 190003, Popayán- Cauca, Colombia.

Pages: 841–853

Resumen: En la actualidad se ha evidenciado que en las instituciones educativas, se ven expuestas de manera frecuente a problemas relacionados con el nivel de seguridad física, dada la carencia de barreras que permitan minimizar el riesgo y por ende el impacto al ser vulnerado. A partir de lo anterior, en este artículo se presenta como principal aporte la propuesta de una arquitectura IoT (Internet de las cosas) enfocada en el área de la seguridad, la cual se ha organizado en las capas provistas por la arquitectura Lambda. No obstante esta propuesta busca recopilar las tecnologías y las herramientas más adecuadas para su implementación, las cuales pretenden servir de referencia para el diseño y construcción de futuros servicios IoT relacionados. Finalmente, para evaluar ésta, la arquitectura fue desplegada en diferentes dispositivos, con el objetivo de contrastar y analizar los diferentes tiempos de procesamiento y de respuesta obtenidos.

Palabras-clave: Arquitectura, Identificación de personas, Internet de las cosas, Sistemas biométricos.

IoT architecture for the identification of people in educational environments

Abstract: At present, it has been shown that in educational institutions, they are frequently exposed to problems related to the level of physical security, given the lack of barriers to minimize the risk and therefore the impact of being violated. Based on the foregoing, in this article the proposal for an IoT architecture focused on the area of security, which has been organized in the layers provided by the Lambda architecture, is presented as the main contribution. This proposal seeks to gather the most appropriate technologies and tools for its implementation, which are intended as a reference for the design and construction of future related IoT services. Finally, to evaluate this, the architecture was deployed in different devices, with the aim of contrasting and analyzing the different processing and response times obtained.

Keywords*:* Architecture, Identification of people, Internet of things, Biometric systems.

1. Introducción

En los últimos años la inseguridad se ha vuelto un problema de orden cotidiano, debido al incremento de los índices de hurtos según reportes de la policía y de las autoridades gubernamentales. Esto al parecer es ocasionado por múltiples factores sociales y económicos que las políticas de los gobernantes pasados y actuales no han podido contrarrestar causando desasosiego entre la población. La inseguridad ya no solo pasa por los barrios periféricos y de alta delincuencia de las ciudades, sino en cualquier sector inmiscuyendo a todo tipo de población. Los edificios han tenido que ser protegidos mediante la instalación de cámaras de videos y sistema de vigilancia privada que no son suficientes, ya que el ritmo de la delincuencia no conoce límites y es muy ágil para superar los controles que se implementan. Un sector que, aunque parece respaldado y protegido, ha estado expuesto a este tipo de problemáticas es el de los centros educativos, el cual se ha convertido en objetivo primordial para el robo de elementos propios de las instituciones y de las personas que hacen uso de las instalaciones. Lo anterior se evidencia en diferentes acontecimientos reportados por los medios de comunicación, en los cuales asaltantes vulneran en horarios no laborales la seguridad física de los establecimientos, extrayendo elementos de los planteles educativos (Caracol, 2017) (Carvajalino, 2017).

Una de las alternativas tecnológicas a los anteriores problemas ha sido el uso de sistemas para la detección y localización de objetos, sin embargo, la falta de integración con otros sistemas de adquisición y validación a partir de variables biométricas los hace ineficientes. A partir de lo anterior y considerando la gran cantidad de dispositivos de Internet de las cosas (IoT) que están surgiendo en la actualidad (drones, sensores, cámaras, tarjetas de captura de datos), estos pueden considerarse como herramientas útiles para contrarrestar las diferentes amenazas físicas que pueden llegarse a presentar, para disminuir el riesgo y por ende generar un impacto menor. Una de las principales ventajas en el uso de estos dispositivos, es que pueden articularse en red y trabajar con variables de tipo biométrico para la generación de sistemas de alertas tempranas (Neira Orjuela, 2015).

Dentro del campo de investigación en mención también se encuentran trabajos que buscan y aportan alternativas de seguridad utilizando sistemas biométricos. Algunos de ellos se enfocan en temas afines, como es el caso del trabajo propuesto en (Ayala Rojas, 2017), en el cual se presenta una aplicación de seguridad biométrica para la identificación de usuarios en la web de la Universidad Abierta y a Distancia – UNAD seccional Tunja. En dicho trabajo se concluye que el sistema actual no brinda una adecuada seguridad en cuanto al acceso a la plataforma, por lo cual se sugiere el uso de un sistema basado en métodos de reconocimiento facial, para ofrecer más y mejores garantías a nivel de seguridad, debido a los bajos costos de adquisición, el mantenimiento, la efectividad y la garantía en seguridad de acceso. Del mismo modo, en el trabajo presentado en (JARAMILLO, 2013), se presentan un conjunto de tecnologías y técnicas adecuadas para la implementación de seguridad física en el control de acceso a una residencia.

En virtud a lo anterior, se puede considerar que la seguridad física es de gran importancia puesto que los establecimientos educativos regularmente permiten un acceso manual y no automatizado a las instalaciones. Así mismo, el proceso de validación de identidad no se hace por lo general con una base de datos, ni con sistemas de seguimiento biométrico, por lo que la comprobación de identificación se hace en la mayoría de los casos in situ y

bajo el criterio del personal de seguridad que se encuentre en los accesos, lo que puede provocar que no se tenga un control sobre el personal autorizado que ingresa a los establecimientos.

En consecuencia a lo antes mencionado, en este artículo se presenta una propuesta arquitectónica para la identificación de personas mediante la utilización de un sistema biométrico soportado en el uso de dispositivos tipo SBC (computadora de placa simple). La propuesta arquitectónica está basada en la arquitectura para IoT Lambda, la cual es usada para el tratamiento de grandes volúmenes de datos y cuya estructura está formada por cuatro capas: capa de recolección, capa de almacenamiento, capa de procesamiento y capa de consulta o visualización. De este modo, se puede concluir que se presentan importantes ventajas con la implementación de esta arquitectura, dentro de las cuales se encuentran: integración de tecnologías IoT en el área de la seguridad, seguimiento y control en tiempo real sobre el acceso a las aulas por parte del personal académico, haciendo uso de variables de tipo biométrico y construido bajo un marco metodológico de trabajo ágil alineado a la arquitectura Lambda. Así, la propuesta arquitectónica presentada en este artículo, pretende servir de referencia para la implementación de servicios de seguridad en establecimientos educativos.

El resto de este artículo está organizado de la siguiente forma: en la sección dos se presentan los trabajos relacionados con la temática de estudio, así mismo en la sección tres se muestra la arquitectura soportada sobre IoT, en la sección cuatro la evaluación de la arquitectura propuesta y finalmente en sección cinco las conclusiones y trabajos futuros.

2. Trabajos Relacionados

En este espacio se presentan los trabajos relacionados con la investigación, los cuales se encuentran enmarcados en la temática de reconocimiento facial y su relación con dispositivos IoT en sistemas biométricos desarrollados para diferentes campos de aplicación.

En el trabajo de investigación presentado por (Guillen, 2015), se plantean las ventajas de algunas de las tecnologías para la captura y análisis del rostro. Además, se detalla cómo los sistemas de reconocimiento pueden hacer la comprobación de la identidad mediante la comparación de los histogramas. En este se diseña un sistema, en el cual se logra la identificación de varios usuarios, lo que permite a la investigación tomar como parámetros los resultados obtenidos por los autores para mejorar la propuesta de arquitectura que se propone.

En (Chuquisengo, 2006), se esboza la implementación de sistemas biométricos como elementos de seguridad, especificando su planteamiento en lectores de huellas dactilares, abordando en detalle algunos de los sistemas biométricos que existen en la actualidad, los diseños, evaluación de costos y beneficios; así como la implementación de estos sistemas de seguridad tanto en lugares públicos como privados. Con este análisis se provee un acercamiento al empleo de los sistemas biométricos como medio de seguridad en los centros educativos.

Así mismo en el trabajo presentado en (AGUSTIN, Septiembre 2015), se establece la identificación de los rasgos físicos de las personas en el campo de la criminalística, haciendo uso de retratos hablados y la digitalización de esta información en bases de

datos utilizando softwares y sistemas como IDENTIKIT, ASPLEY 3D E-FITH, CDOZIC entre otros, incluyendo un estudio morfológico de la forma de cabeza, cara, ojos, boca y nariz. Mediante este referente se observa la practicidad de un sistema de identificación con la utilización de técnicas de retratos automatizados.

De esta forma, la información recolectada en esta literatura hace valida la necesidad de proponer una arquitectura IoT que conjugando las tecnologías de identificación facial y utilizando bases de datos y sistemas embebidos como SBC, permita el seguimiento y control sobre el acceso a un sitio público o privado, por ende brinde seguridad física a las personas beneficiadas con la utilización de este tipo de arquitecturas.

3. Arquitectura Soportada en IoT

La propuesta arquitectónica presentada en esta sección ha sido estructurada en tres perspectivas, que son perspectiva funcional, perspectiva de implementación y perspectiva de negocio (ver fig. 1). Cada una de estas perspectivas ha adoptado una de las cuatro capas de la arquitectura Lambda (capa de recolección, capa de almacenamiento, capa de procesamiento y capa de consulta o visualización) y han sido implementadas bajo una metodología de trabajo ágil.

No obstante, en este mismo orden de ideas se describe que en la perspectiva funcional se establecen los diferentes componentes funcionales de la propuesta arquitectónica, en la perspectiva de implementación se tienen los diferentes componentes de hardware y software seleccionados para la construcción de la arquitectura y finalmente en la perspectiva de negocio se encuentra el modelo de negocio de la arquitectura en el ámbito del aseguramiento físico en las aulas de clase de una institución educativa. A continuación, se presenta la descripción de cada una de las perspectivas de la propuesta arquitectónica.

Figura 1 – Arquitectura IoT

3.1. Perspectiva funcional

Aquí se concibe la propuesta arquitectónica basada en las cuatro capas de Lambda, teniendo en cuenta el comportamiento funcional de la arquitectura. La capa de recolección de datos se caracteriza por recopilar la información en tiempo real desde la captura de imágenes tanto de los gestos como los rasgos faciales de personas a través de la cámara digital. La capa de almacenamiento tiene por función recibir los datos, almacenarlos y almacenarlos en la base de datos. La capa de procesamiento se encarga de analizar estos datos por medio de las variables establecidas y validadas desde la librería OpenCV. Dado a lo anterior es posible realizar el reconocimiento e identificación de las personas

registradas y tener así acceso autorizado en las aulas de clase, en la asignatura, hora y día en que se ha matriculado; y la capa de consulta es la destinada para los usuarios finales que les permite mediante la implementación de la interfaz gráfica interactuar y poder visualizar si el acceso a través del sistema biométrico está siendo validado y reconocido como persona registrada y de esta manera lograr el ingreso y además obtener un reporte de las personas que han ingresado por día.

En la capa de recolección, las entradas de los datos se toman con el sistema de Captura de Datos (Raspberry Pi - Cámara digital) y se envía en dos direcciones, a la capa de almacenamiento (gestor de base de datos) y a la capa de procesamiento (herramientas Machine Learning - empleando librería OpenCV para el reconocimiento facial). Este proceso garantiza que, si llegase a ocurrir una pérdida de datos por causas de la conectividad, pueda volver a reproducirse sin problemas. Esta información puede ser accedida por el administrador a través de una conexión a internet por medio de dispositivos como: servicios web y/o aplicaciones móviles, a través de la capa de consulta.

La capa de almacenamiento genera un archivo histórico, guardando los datos que se hayan recolectado. El almacenamiento de éste es de 4 meses periodo que dura un semestre académico, el cual podría ajustarse de acuerdo a las necesidades requeridas por la institución educativa.

La capa de procesamiento trabaja como una instancia que conjuga listas de datos, y streaming, proceso que se realiza en tiempo real. Es importante destacar que la capa de almacenamiento basa su reconocimiento y validaciones en datos almacenados previamente, los cuales son usados por los algoritmos de clasificación provistos por la librería OpenCV. Así mismo, es de destacar que una buena parte de la responsabilidad de esta capa recae sobre el administrador de la arquitectura, quien se encargará de velar que la información recolectada sea la pertinente para tener una autenticación por parte de las personas autorizadas, en la capa se consulta.

Las funciones de cada una de las capas descritas se pueden ver desglosadas en la figura 2, el cual representa la secuencia funcional de la arquitectura basada en IoT para la identificación de personas, de acuerdo a las capas propuestas por la arquitectura de referencia como Lambda.

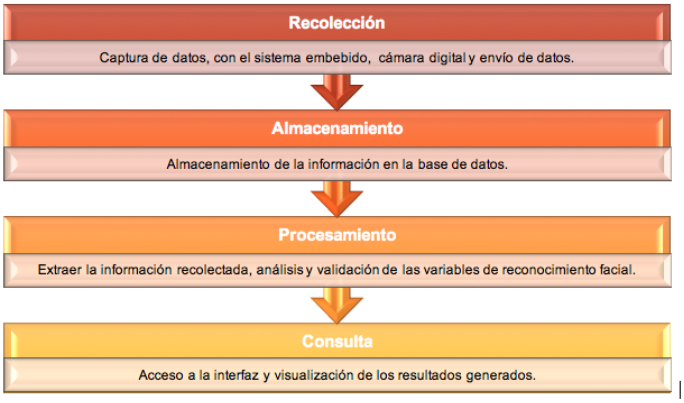


Figura 2 – Secuencia funcional de la arquitectura IoT

3.2. Perspectiva de implementación

En esta perspectiva se presenta los componentes de hardware y software seleccionados para la implementación de la propuesta arquitectónica soportada en IoT, los cuales están organizados de acuerdo a las capas de la arquitectura Lambda.

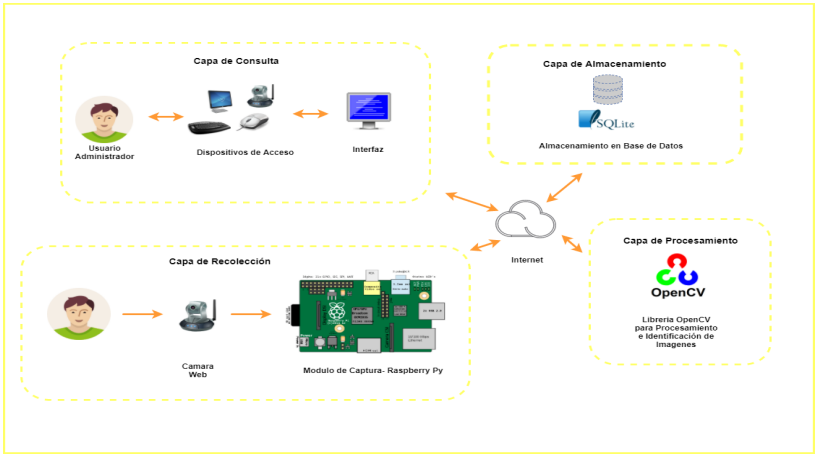


Figura 3 – Perspectiva de implementación de la arquitectura IoT

En la figura 3, específicamente en la capa de recolección se encuentra la cámara digital que permitirá tomar la imagen fotográfica del estudiante, que se gestionará por medio de la tarjeta de IoT Raspberry pi 3, el cual se encuentra configurada con el anterior dispositivo para conformar la capa de recolección de datos. Gracias al SBC escogido se envía en tiempo real las capturas de la imágenes asociadas a un registro de identificación para ser almacenado en una base de datos SQLite sincronizada; así mismo se desarrolló para la capa de procesamiento una aplicación en Python que permitió realizar en línea el procesamiento de los datos recolectados en las capas posteriores, mediante la utilización de la librería OpenCV, permitiendo la validación de variables para efectuar el reconocimiento al momento de hacer un registro e ingreso al aula de clases.

Por consiguiente, en la capa de consulta, es accedida por el administrador para realizar el debido registro a los estudiantes al sistema a través de un aplicativo desarrollado en Python con las librerías Tkinter y herramientas como PAGE, quienes permiten establecer comunicación con la base de datos, y así poder visualizar los datos capturados y dar acceso al aula de clases a los estudiantes registrados.

Dada la importancia de cada uno de los componentes hardware y software asociados a las diferentes capas expuestas se procede a realizar una breve descripción de ellos:

Capa de recolección. En esta propuesta arquitectónica se empleó una tarjeta SBC de bajo costo como es el Raspberry Pi 3, por ser un sistema embebido adecuado para la capa de recolección de datos. Este se complementa con el sistema operativo Raspbian sobre

la plataforma del Raspberry y una configuración vía Wifi, que permite que se realice el envío de los datos de manera inalámbrica a las demás capas de la arquitectura planteada (ver fig. 4).

La cámara empleada que se encuentra conectada al SBC es de marca genérica, en este caso Logitech; ésta se encarga de realizar los registros fotográficos que se validará en la capa de procesamiento más adelante.

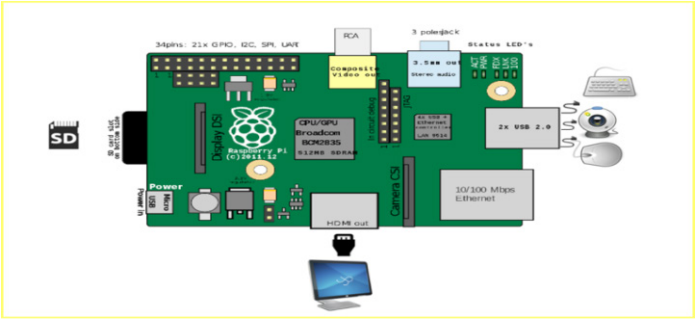


Figura 4 – Esquema de conexión entre dispositivos.

Capa de almacenamiento. En esta capa se cuenta con un gestor de almacenamiento, para lo cual se seleccionó la base de datos SQLite, dado que ofrece una mejor portabilidad, transferencia, importación, exportación, sincronización de datos e informes, de manera fácil, ágil y segura.

Capa de Procesamiento. En esta se procesan los datos recolectados que a partir de los parámetros definidos y validados con la librería OpenCV, realizan el proceso de identificación de la persona siempre y cuando ésta haya sido registrada previamente. Esta librería se caracteriza por ser libre de visión artificial y contar con algoritmos de machine learning. Dada su infraestructura permite realizar un reconocimiento facial, gracias a los diferentes algoritmos de visión artificial o de clasificación que contiene ésta. Además, en esta capa juega un papel importante el gestor de la base de datos SQLite, quien permite el almacenamiento para posteriormente ser procesada y gestionada a través de la función de identificación.

Capa de consulta. En esta capa, se desarrolló una interfaz gráfica de acceso a la aplicación de procesamiento provisto por la librería Tkinter y la herramienta para la generación automática de interfaces Page, las cuales fueron utilizadas con el lenguaje Python para el diseño e implementación de la misma.

No obstante, es aquí donde se lleva a cabo tanto la consulta de visualización si el usuario se encuentra registrado o no, como también el proceso de actualización y registro de un usuario nuevo en el sistema. En la figura 5, se evidencia como se visualiza el proceso de identificación y validación de un estudiante registrado, desde el dispositivo SBC.

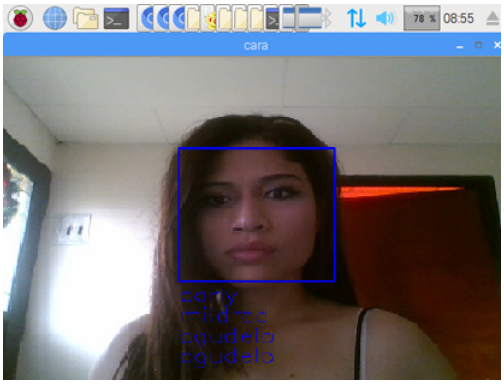


Figura 5 – Visualización de identificación

3.3. Perspectiva de Negocio

En la perspectiva de negocio se muestra la propuesta arquitectónica desplegada para entornos educativos, lográndose observar las diferentes capas de la arquitectura Lambda.

Atendiendo a la propuesta arquitectónica, en la capa de recolección se encuentran las instituciones de educación las cuales son la población objetivo, el dispositivo SBC (Raspberry Pi), Cámara digital, pantalla y la conectividad a internet, que permiten recepcionar y visualizar la información que se ha capturado a través de éstos. La capa de almacenamiento esta comprende el gestor de bases de datos SQLite que permiten guardar la información recepcionada desde la capa de recolección. La capa de procesamiento es aquí donde se tratan los datos almacenados para realizar las validaciones que permita a las personas autenticarse. Finalmente, en la capa de consulta, permite tanto al estudiante como al administrador visualizar el tiempo de respuesta al autenticarse una persona registrada, como el acceso al reporte de las personas que han ingresado por día al recinto que tiene asociado el dispositivo.

4. Evaluación de la propuesta arquitectónica

Finalmente se realizó una serie de evaluaciones a la propuesta arquitectónica en el módulo de procesamiento específicamente en lo que se refiere al algoritmo de identificación de imágenes. Este módulo fue dispuesto a evaluación para conocer los límites de procesamiento y de utilización del recurso RAM dentro de un tiempo arbitrario de 60 segundos con el fin de obtener y analizar el consumo de memoria y la cantidad de tiempo medido en milisegundos que usa en cada una de las invocaciones para ejecutar el método de identificación. Lo anterior con el fin de optimizar a futuro los recursos y diseñar estrategias de optimización del procesamiento del aplicativo.

Para lo anterior, se plantearon tres escenarios en los que la arquitectura fue puesta a prueba con tres tipos diferentes de hardware y/o sistema operativo: para el primer escenario o escenario principal se usó una placa SBC con sistema operativo Raspbian, el segundo escenario se ejecutó sobre un ordenador personal con sistema operativo Windows, y por último el tercer escenario se desplegó sobre una máquina virtual con Sistema Operativo Raspbian.

4.1. Escenario 1

En este escenario de evaluación se empleó una placa SBC de tipo Raspberry Pi 3 con sistema operativo Raspbian.

Evaluación. En la Fig. 6. Se describe el tiempo utilizado en cada una de las invocaciones realizadas al método de identificación de imágenes. En el eje X se encuentra el número de invocaciones al algoritmo, mientras que en el eje Y se define el tiempo en que se realiza cada uno de los procesamientos. Con lo anterior se logró evidenciar que el número de respuestas en el procesamiento de identificación muestra un rango promedio entre 800 y 1000 milisegundos con un salto abrupto hasta de 1727 milisegundos. La línea de distribución de igual manera muestra que los datos permanecen estables durante casi todo el periodo usando la placa SBC, el cual es el hardware para el que fue considerada la arquitectura.

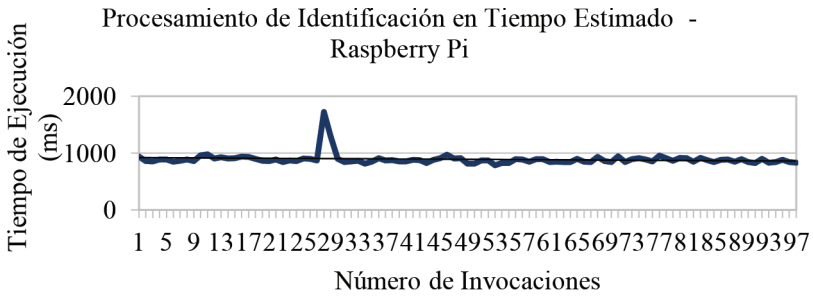


Figura 6 – Tiempo del procesamiento durante la ejecución del módulo de identificación-sobre el Raspberry Pi.

En la Fig. 7 se presenta el porcentaje de memoria RAM empleado por el método de procesamiento e identificación de imágenes. Cabe anotar que el eje X denota el número de invocaciones al método que hace la identificación, mientras que en el eje Y el porcentaje de memoria consumida para la función de procesamiento. De esta manera se comprueba que este componente de la arquitectura utiliza por parte del procesador una memoria entre el 15 y 23 % de su uso, y la línea de distribución también determina que estos datos se mantienen estables, por lo cual se concluye que la propuesta arquitectónica utiliza un bajo uso de recursos de memoria, dado que no sobrepasa el 50 % de ésta, según los valores obtenidos.

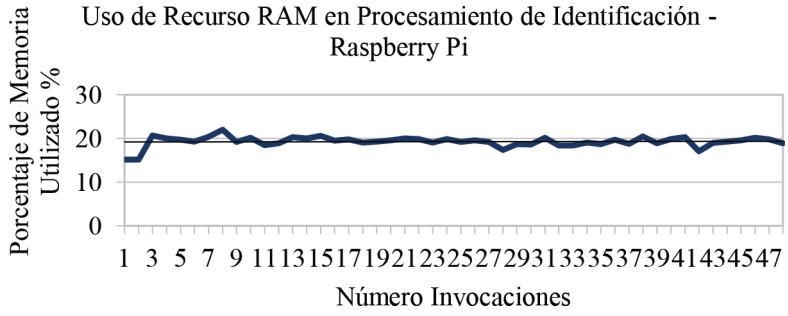


Figura 7 – Porcentaje de uso de RAM en módulo de Identificación.

4.2. Escenario 2

Para el escenario 2, se empleó un ordenador personal de marca LENOVO, con 8 GB de RAM, un procesador de 5 núcleos y la versión de Sistema Operativo Windows 10 Home.

Evaluación. En las Fig. 8. y 9 se utilizaron los mismos parámetros de la evaluación del escenario 1 tanto para la medición del tiempo de procesamiento del módulo de identificación como para la evaluación del consumo y uso del recurso RAM. En la Fig. 8 se encontró que el tiempo utilizado inicialmente para la primera invocación es de 919 milisegundos, luego de lo cual cae a 127 milisegundos en las siguientes invocaciones y permanece posteriormente en un rango entre 90 y 191 milisegundos para los siguientes procesamientos.

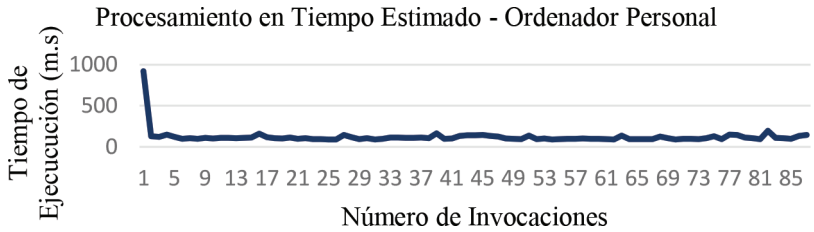


Figura 8 – Tiempo del procesamiento durante la ejecución del módulo de identificación en un ordenador personal.

En la Fig. 9 el porcentaje de memoria RAM empleado para la función evaluada determinó que existen unos picos altos de hasta 93.5% y otros muy bajos de 5%. Lo anterior permite evidenciar que el consumo de memoria RAM no permanece estable durante el procesamiento de la identificación de imágenes bajo el tiempo que duró la prueba. De esta manera se comprueba la inestabilidad en el uso de memoria desplegando la arquitectura propuesta en un ordenador personal con sistema operativo Windows.

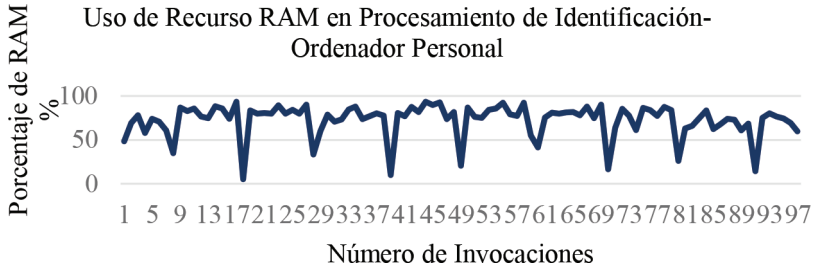


Figura 9 – Porcentaje de uso de RAM en módulo de identificación con ordenador personal.

4.3. Escenario 3

En este último escenario se utilizó una máquina virtual Raspbian sobre la plataforma VirtualBox con una RAM virtualizada de 3 GB y la versión de sistema operativo raspbian jessie 06-2017.

Evaluación. Los parámetros utilizados en esta última prueba son los mismos usados en las evaluaciones anteriores. En la fig. 10 los tiempos de ejecución varían entre 1771 y 967 milisegundos. Aunque en este escenario se obtienen en el mejor de los casos valores inferiores a los 1000 milisegundos, se produce unos picos que no determina estabilidad en los valores obtenidos, con respecto al escenario 1.

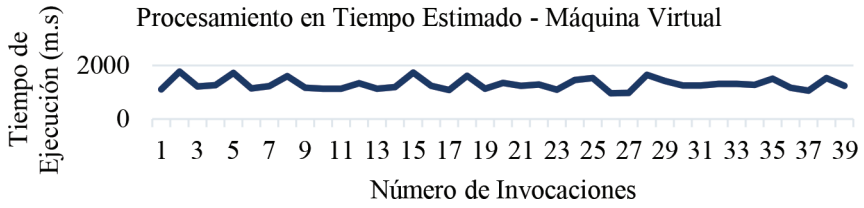


Figura 10 – Tiempo del procesamiento durante la ejecución del módulo de identificación en máquina virtual.

Del mismo modo utilizando las mismas características de las evaluaciones de memoria en los escenarios anteriores, la prueba concluyó como se muestra en la fig. 11. que en este escenario se producen picos altos de 89.2% y se alcanzan valores mínimos de 0%.

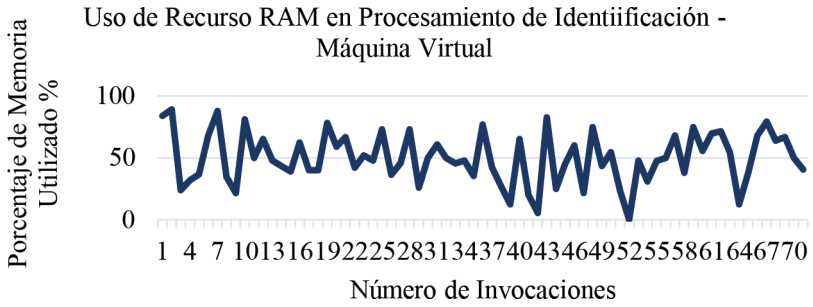


Figura 11 – Porcentaje de uso de RAM en módulo de identificación con máquina virtual.

Las evaluaciones realizadas en los escenarios anteriormente descritos permiten determinar que el tiempo de procesamiento del algoritmo de prueba se relaciona o adapta mejor con el hardware del escenario dos, esto debido a que el ordenador tiene especificaciones mejores para los tiempos de respuesta para el procesamiento de las imágenes. Así mismo, se evidencia que en cuanto al porcentaje del uso de recurso RAM, la opción del Raspberry pi brinda unas mejores condiciones al utilizar un rango bajo y ser mucho más estable que las condiciones dadas por los otros escenarios expuestos. Si se colocan en conjunto todas las opciones analizadas durante el desarrollo de este artículo, tales como: la estabilidad que brindan los dispositivos SBC, los costos económicos que se obtienen con el uso de esta tecnología y los datos en perspectiva, se llegan a determinar que la plataforma de hardware óptima para esta arquitectura es la escogida inicialmente.

5. Conclusiones y trabajos futuros

Con la utilización de los dispositivos de hardware y software libre, se logró diseñar una arquitectura soportada bajo el concepto de Internet de las Cosas. Esta propuesta está basada en la arquitectura Lambda y tiene asociadas las perspectivas de funcionalidad, negocio e implementación. Cada perspectiva cuenta con cuatro capas de operación como son: recolección, que se realiza con las imágenes del rostro facial del usuario; almacenamiento, que permite la persistencia de estas imágenes sobre una base de datos SQLite; procesamiento, que ejecuta tareas de registro y comparación mediante la librería OpenCV para el procesamiento de validación; consulta, la cual permite la visualización de las tareas de reconocimiento en tiempo real mediante interfaces implementadas usando la librería Tkinter.

Esta propuesta de arquitectura se propone como un modelo de seguridad física basado en un sistema biométrico de bajo costo, para que sea implementada en los centros educativos tanto públicos como privados.

De acuerdo a los resultados obtenidos en las pruebas de consumo de memoria y tiempo de respuesta del algoritmo de identificación, se puede concluir que las herramientas seleccionadas son adecuadas para la implementación de este tipo de propuestas arquitectónicas soportadas en IoT.

Es de destacar que uno de los objetivos logrados fue integrar los componentes en una sola arquitectura, teniendo la placa SBC como hardware de bajo costo, complementado con software libre como: la base de datos SQLite y el desarrollo en un ambiente Python, mediante las herramientas PAGE, Tkinter, y la librería OpenCV, ayudado de IDEs como Geane.

Dentro de los trabajos futuros se pretende evaluar la propuesta con varios dispositivos conectados a la vez y ligados a un almacenamiento Icloud, para construir una arquitectura más escalable y diferenciar las alternativas con la propuesta inicialmente planteada para esta investigación. Además, con el conocimiento que se tiene con relación a los sistemas biométricos que carecen de seguridad en el ciframiento de la información resguardada, se propone evaluar un algoritmo criptográfico para determinar el más adecuado para implementar en este tipo de arquitectura y así salvaguardar los datos manteniendo la integridad, confidencialidad y disponibilidad de ellos.

Referencias

Agustin, J. C. (Septiembre 2015). *Identificación a Través de Los Rasgos Biométricos del Rostro*. Huehuetenango, Guatemala: Universidad Rafael Landívar.

Ayala Rojas, N. A. (2017). *Monografía de estudio sobre la aplicación de seguridad biométrica para la identificación de usuarios en entornos web*. Tunja - Colombia: Universidad Nacional Abierta y A Distancia.

Caracol. (22 de 06 de 2017). *Noticias Caracol Tv*. Obtenido de https://noticias.caracoltv.com/bogota/rajados-en-seguridad-diez-ladrones-asaltaron-todo-un-salon-de-clases-en-una-nocturna

Carvajalino, L. (4 de 10 de 2017). *El Heraldo*. Obtenido de Periodico el Heraldo: https://www.elheraldo.co/judicial/colaborenme-con-los-celulares-ladron-estudiantes-de-la-cuc-408977

Chuquisengo, L. E. (2006). *Verificación de Identidad de Personas Mediante Sistemas Biométricos Para El Control de Acceso a Una Universidad*. Lima - Perú: Pontificia Universidad católica del Perú.

Guillen, D. E. (2015). *Reconocimiento facial*. Valparaiso-Chile: Pontificia Universidad Católica de Valparaiso.

Jaramillo, J. E. (2013). *Soluciones inteligentes para el control de acceso físico mediante el uso de tecnología biométrica*. Pereira: Universidad Tecnologica de Pereira.

Neira Orjuela, F. (2015). Biometría y control migratorio en América Latina Cuadernos de H Ideas. *Vol. 9*.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 23/09/2018
Aceitação/Acceptance: 29/11/2018

A evolução cultural graças à adoção da tecnologia e ao fenómeno "Mix-Tech" – Um estudo exploratório baseado na observação

Manuel Au-Yong-Oliveira[1], Frederico Branco[2], Carlos Costa[3]

mao@ua.pt, fbranco@utad.pt, ccosta@ua.pt

[1,3] GOVCOPP, Department of Economics, Management, Industrial Engineering and Tourism, University of Aveiro, Campus Universitário de Santiago, 3810-193 Aveiro, Portugal.

[2] INESC TEC and University of Trás-os-Montes e Alto Douro, Quinta de Prados, 5000-801 Vila Real, Portugal.

Pages: 854–869

Resumo: Indivíduos passam atualmente parte significativa do seu dia interagindo com ecrãs de computador e de smartphone. Este estudo procura contrariar isto e é baseado na observação participante durante uma viagem de 16-28 de agosto a Istambul e à Ásia do Sul. Apuramos que a tecnologia poderá ser um meio que mostra que somos similares em muitos aspetos, apesar de estarmos separados por distâncias grandes e pela língua. A observação é essencialmente uma ferramenta que destapa significados, que um visitante e investigador poderá ver. Em particular, nas Maldivas, na ilha de Meedhoo, vimos uma forma superior de viver. A harmonia era o elemento principal que vimos na ilha, incluindo entre géneros, que vimos como sendo iguais, apesar do que se encontra escrito nos *media* acerca de certas culturas e religiões. No entanto, em Istambul, na Turquia, vimos uma cultura que ainda permanece sendo muito diferente do Ocidente.

Palavras-chave: Cultura; Tradição; Mudança; Aprendizagem; Harmonia; Inovação.

How cultures evolve aided by technology adoption and the "Mix-Tech" phenomenon – An exploratory study based on observation

Abstract: Individuals currently spend a considerable part of their day-to-day lives interacting with computer or smartphone screens. This study seeks to contradict this and is based on participant observation during a trip from 16-28 August 2018, to Istanbul and to South Asia. We have found that the role of technology may be a medium which shows that we are similar in many ways, despite being separated by large distances and by language. Observation is essentially a tool to uncover meanings, which a visitor and researcher may see. In particular, in the Maldives, on Meedhoo island, we saw a superior form of living. Harmony was the main element we saw on the island, including between genders, which we saw as being equal, despite what is to be found written in the media about certain cultures and religions. Albeit, Istanbul, in Turkey, revealed a culture that still remains very different from the West.

Keywords*:* Culture; Tradition; Change; Learning; Harmony; Innovation.

1. Introdução

Um dos objetivos deste estudo é analisar dois estados soberanos muito diferentes, a Turquia e as Maldivas – diferentes tanto em tamanho quanto em geografia – aplicando a ferramenta de pesquisa conhecida como observação a esses dois destinos de férias.

Adicionalmente procuramos dissertar acerca de como as culturas inovam ao longo do tempo, misturando e aprendendo com outras culturas – uma mistura que inevitavelmente ocorrerá, nos casos escolhidos, dada a quantidade de turistas que visitam cada localidade anualmente, mas também através de outras interações, tais como, a Internet e outros meios de comunicação.

A Turquia é membro da Organização das Nações Unidas (ONU), da OTAN (Organização do Tratado do Atlântico Norte), do FMI (Fundo Monetário Internacional) e do Banco Mundial. Além disso, a Turquia é membro fundador da OCDE (Organização para Cooperação e Desenvolvimento Econômico). A Turquia é um país de grande dimensão populacional com mais de 80 milhões de habitantes que tenta entrar na União Europeia (UE) há algum tempo.

Istambul era uma mistura do antigo com o novo (Imagem 1). Um elétrico (*tramcar*) moderno pode ser visto com uma mesquita de fundo. Os elétricos e o metro estão disponíveis para viagens através da compra de um único cartão, que pode ser carregado nas máquinas com dinheiro para realizar viagens.

Imagem 1 – Uma mistura do antigo e do novo em Istambul - Um elétrico moderno e uma mesquita ao fundo

As Maldivas, por outro lado, são uma nação soberana arquipelágica, localizada no Oceano Índico, no sul da Ásia. As Maldivas são na verdade o país asiático mais pequeno, em termos de população (cerca de 400.000 habitantes) e território. Além disso, as Maldivas são membros da ONU e membros fundadores da Associação do Sul da Ásia para Cooperação Regional (SAARC). Durante a permanência nas Maldivas passamos algum tempo num resort na ilha Ugoofaaru (Kaafu, Maldivas) e na ilha de Meedhoo (com cerca de 2.000 habitantes) habitada por uma comunidade local.

A Imagem 2 mostra como os habitantes locais das Maldivas aderiram à tecnologia (e.g., smartphones) enquanto mantêm as tradições.

Imagem 2 – Cidadãos das Maldivas gravando e interagindo com seus smartphones no decorrer de um festival local tradicional - ilha Meedhoo

No que diz respeito ao Índice de Desenvolvimento Humano (IDH), "criado para enfatizar que as pessoas e suas capacidades devem ser os critérios principais para avaliar o desenvolvimento de um país, não se cingindo apenas ao crescimento econômico" (Human Development Reports, 2017), a Turquia ficou em septuagésimo primeiro lugar e as Maldivas em centésimo quinto, no que concerne ao seu desempenho de 2016. No topo da lista estavam a Suíça, em segundo lugar, e a Noruega, em primeiro.

A plataforma Hofstede Insights (2018) classifica a Turquia como tendo um elevado *Power Distance* (pontuação de 66 – o que significa que o poder distribuído desigualmente é aceite e que o chefe ideal é paternalista; o poder é centralizado) e ainda um maior *Uncertainty Avoidance* relativo (pontuação de 85 – o que significa que as leis e regras são uma necessidade importante, além disso, os rituais desempenham também um papel importante na sociedade, incluindo os religiosos). Por outro lado, a Turquia tem

baixa pontuação no *Individualism* (pontuação de 37, o que significa que é uma sociedade coletivista). As Maldivas não são classificadas pela plataforma Hofstede Insights (2018).

2. Metodologia

O autor principal fez uma viagem de férias a Istambul, na Turquia, e às Maldivas, com a sua família e alguns amigos, entre 16 e 28 de agosto de 2018. Este artigo é um estudo descritivo – "projetado para produzir uma representação precisa de pessoas, eventos ou situações" (Saunders & Lewis, 2018, p.116) – enquanto, adicionalmente, procura ser explicativo – "estudar uma situação ou um problema para explicar as relações entre variáveis" (Saunders & Lewis, 2018, p. 118). O artigo baseia-se em observações pessoais face-a-face feitas durante as interações culturais com os habitantes locais, principalmente no setor de serviços, mas também em observações feitas durante outros encontros sociais. Um desses encontros ocorreu durante uma visita guiada, com um guia especializado, e a interação com outros residentes nativos. O autor participou da atividade observada. A questão de pesquisa que abordamos é: Como é que os estrangeiros nas suas relações promovem uma interação social aberta e empática? Apesar do envolvimento com os locais não ter sido muito profundo, a interação foi suficiente de forma a retirar benefícios mútuos sem entrar em conflito. Neste manuscrito, descrevemos alguns episódios que ocorreram e que nos marcaram durante a estadia. É nossa intenção abordar o desenvolvimento da teoria de forma indutiva – "uma abordagem de pesquisa que envolve a construção da teoria a partir da análise de dados já recolhidos" (Saunders & Lewis, 2018, p. 113). Todas as fotografias exibidas foram obtidas durante a viagem. Foram gravados vídeos de vários eventos para suporte à análise posterior. As observações foram desestruturadas e registadas num bloco de notas (Saunders & Lewis, 2018). Para quase todas as fotos foi solicitada permissão para registar as observações.

3. Um olhar sobre a literatura relativa à cultura e mudança, inovação aberta e tecnologia

Cultura é definida como "Compreensões compartilhadas manifestadas em ato e artefacto" (Redfield, 1948, as quoted in House, Hanges, Javidan, Dorfman, & Gupta, 2004). Schein (2010, p.10) complementa a definição afirmando que a cultura é "A aprendizagem compartilhada acumulada de um dado grupo, cobrindo elementos comportamentais, emocionais e cognitivos do funcionamento psicológico total dos membros do grupo". Durante a nossa estadia não observamos nada que implicasse qualquer tipo de perigo ou algo realmente fora do comum e incompreensível.

No que diz respeito à inovação, "as raízes da inovação aberta remontam à história [...] Nem a utilização de contributos de pessoas externas para melhorar os processos internos de inovação, nem o procurar de oportunidades de comercialização externas para o que foi desenvolvido internamente é novo" (Huizingh, 2011, p.2), a inovação aberta já foi anteriormente utilizada para o estudo da antropologia cultural (Huizingh, 2011). Consideramos que as culturas estão predispostas para adotar inovações e novos valores e práticas de culturas externas, e de facto vêm fazendo isso há milênios. "Para inovar, as empresas muitas vezes precisam aprender e colaborar com um grande número de atores externos à sua organização." (Laursen & Salter, 2014, p.867). Os países não são

diferentes, e os aspetos culturais fazem parte integrante das afirmações anteriores, já que as culturas que não inovam e se adaptam estão condenadas a desaparecer. Neste estudo, vemos exemplos de como as culturas adotaram a tecnologia e outras amenidades modernas (imagens 1 e 2), mantendo, no entanto, outras práticas culturais centrais para a sua religião e crenças.

"A criação de inovações geralmente requer abertura, mas a comercialização de inovações requer proteção." (Laursen & Salter, 2014, p.867). Assim, vemos aqui um paradoxo, como mencionado por Laursen & Salter (2014), as culturas e religiões devem ser abertas e inovadoras, mantendo-se em simultâneo protetoras dos seus valores e instituições fundamentais, os quais não devem perder, por receio de perderem a identidade. O que deve ser permitido e o que não deve? Perguntas difíceis em tempos de mudança tecnológica rápida e duradoura.

Kratzer, Meissner, e Roud (2017, p.128) afirmam que "o conhecimento e a tecnologia são desenvolvidos e concebidos pelas pessoas, o que evidencia a centralidade e a importância do fator humano no processo de inovação e, consequentemente, na gestão da inovação". Ao falar de inovar em relação às culturas nacionais e regionais, o fator humano não poderia ser mais central. Ao considerar a sobrevivência a longo prazo das culturas e da inovação, "a abertura interna faz a diferença" (Kratzer et al., 2017, p.128).

Alvesson e Sveningsson (2015, p.15) veem a mudança cultural como "um resultado direto de mudanças externas". Forças externas relevantes incluem: político, tecnológico, cultura, demográfico, econômico e relacionadas com o mercado. Para as culturas estudadas e descritas neste manuscrito foram observados símbolos, valores e crenças de outras culturas, ao longo desses vetores, especialmente o vetor da tecnologia que impulsiona a mudança, como veremos a seguir.

A tecnologia auxilia o processo de ensino (Dočekal & Tulinská, 2015), estando esta conectada com a mudança e a aprendizagem. De facto, Dočekal and Tulinská (2015) falam da educação como um fator diferenciador na medida em que permite aos indivíduos compreender a informação, o que representa uma forma de riqueza simbólica. O desenvolvimento de tecnologia "cria a necessidade de formação" (Dočekal & Tulinská, 2015, p.3765) e pode afastar aqueles que não são capazes de adotar e usar a tecnologia. Uma das tecnologias mais transformadoras da sociedade introduzidas recentemente no mercado de massa (pela Apple, em 2007) foi o smartphone - que engloba várias funcionalidades num único dispositivo - como uma câmara, uma calculadora, uma lanterna, uma agenda e diário, um bloco de notas, um álbum de fotos e vídeos, um computador, um dispositivo de Internet, um telefone, a funcionalidade de "Walkman" para ouvir música, entre outros. Serão as culturas de todo o mundo capazes de adotar e usar o smartphone com sucesso?

4. Dados recolhidos por meio de observação participante e comentários de inovação aberta

De seguida, são descritos uma série de eventos para familiarizar o leitor com as culturas das Maldivas e da Turquia – destinos turísticos populares, embora por diferentes razões. As Maldivas são reconhecidas como destino de praia e mergulho. Uma das principais cidades da Turquia, Istambul é um importante destino urbano, com uma localização

única, tanto na Ásia como na Europa, e com monumentos antigos e uma cultura antiga que vale a pena explorar.

4.1. Episódio 1

"Qual é a diferença entre um rapaz e uma rapariga? Somos diferentes?" perguntou uma filha de dez anos, do nosso grupo, aos seus pais, repentinamente autoconsciente e ciente de que talvez houvesse diferenças entre os géneros. Esta pergunta foi feita depois de ver, na Mesquita Azul, no centro de Istambul, as áreas a que as mulheres estavam circunscritas, muito diferentes de onde os homens podiam ir e orar (na câmara principal, de muito maior dimensão).

Observação sobre inovação aberta e lição aprendida: Sim, existem diferenças entre os géneros, seja qual for a perspetiva, as mulheres têm filhos e os homens não. Portanto, no decorrer das nossas vidas e até que que haja alterações em gerações futuras, os géneros nunca serão os mesmos. Além disso, como as mulheres rezam em casa, como nos informaram, elas necessitam de menos espaço numa mesquita. Nas Maldivas também notamos que havia uma clara distinção entre homens e mulheres e que, na sua maioria, havia uma separação entre géneros nas atividades cotidianas (Imagem 3). Além disso, as mulheres tinham as cabeças cobertas, enquanto os homens não tinham (Imagem 3).

Imagem 3 – À esquerda: Reunião de mulheres para assistir a um festival tradicional local "Eid"; À direita: Uma *selfie* com residentes – Meedhoo Island

4.2. Episódio 2

Os funcionários de diferentes serviços, em Istambul, em mais do que uma ocasião, não nos deram o troco referente aos pagamentos das compras realizadas. Enquanto nos EUA, por exemplo, aprendemos que os trabalhadores das atividades associadas ao setor de serviços esperam uma gorjeta e ficamos à vontade para dar mais ou menos dinheiro, ou nenhum; na movimentada Istambul, eles desejam o dinheiro dos turistas, pelo que percebemos, ficando com ele sem pedir autorização para tal. Numa dessas ocasiões, na loja da Nike, no centro de Istambul, o empregado tentou ficar com 15 liras (a lira é a moeda local – e 15 liras rondavam os 3 euros, de acordo com a taxa de câmbio à data em Portugal); e só perante a nossa insistência é que entregou o troco. Noutra ocasião,

um vendedor de bebidas disse que não dispunha de troco (durante a compra de duas bebidas, na rua, a este vendedor ambulante) e, portanto, queria reter as 4 liras do troco. Talvez não seja muito dinheiro para um europeu (cerca de um euro), no entanto, consideramos preocupante a sua insistência em ficar com o dinheiro que nos pertencia. Neste caso, o vendedor mostrou uma só moeda e retorquiu não dispor de troco. "Bem, pode começar por me dar essa lira", afirmei. Foi então que chegou outro cliente, que já tinha ouvido a nossa conversa, o que levou o vendedor a entregar a lira que apresentara. "Pode entregar ao cavalheiro as minhas duas liras", acrescentou ainda o novo cliente, piscando-me o olho. O vendedor anuiu. Eu ainda pretendia obter a lira em falta, por uma questão de princípio, regressando mais tarde para a recolher. Talvez os comerciantes locais tenham conhecimento da tradição de dar gorjetas em algumas culturas, mas ainda não compreendam por completo essa prática.

Observação sobre inovação aberta e lição aprendida: Alguns países não estão acostumados a receber gorjetas, mas pretendem tê-las. Os governos podem ajudar a melhorar esta questão, proporcionando formação relativa à prestação de serviços e dos mecanismos de gratificação (i.e., gorjetas). Adicionalmente, caso se praticassem salários mais elevados, os trabalhadores teriam menor necessidade de se apropriarem do dinheiro (e.g., troco) dos turistas. Precisamos de porventura trabalhar no sentido de haver um salário mínimo global, pelo menos na Europa e, talvez, no resto do mundo. As tremendas desigualdades sentidas ao nível global são uma forte causa de apoquentação social.

4.3. Episódio 3

Naquela noite jantamos num restaurante muito moderno no centro de Istambul, no telhado de um edifício. A conta foi de cerca de vinte euros por pessoa, um valor muito elevado para Istambul. Éramos um grupo de dez. O serviço não foi nada bom. Esqueceram-se de nós por um período considerável de tempo. Não trouxeram bebidas em conformidade com o pedido. Passaram todo o jantar a roçar o limiar da boa educação. Para nossa surpresa, apresentaram-nos, escrito à mão, uma conta com uma taxa adicional de serviço de 10% (uma gorjeta obrigatória de 10%). Decidimos pagar e não discutir. De seguida, pedimos para ficar em posse do recibo, no entanto o empregado levou-o. Corremos atrás do empregado para resgatar o recibo. Este respondeu que já o iria entregar. O que ocorreu, cerca de dez minutos depois, mas entregaram-nos um recibo sem a taxa adicional de serviço escrito à mão, como na versão anterior. O valor do recibo foi, portanto, 10% inferior ao que realmente foi pago. A gorjeta estava, obviamente, "isenta de impostos" e talvez em violação de alguns regulamentos. Certamente, no menu, não houve menção de uma taxa de serviço de 10%. No seguimento do fraco serviço que nos foi prestado, de modo algum daríamos gorjeta. A qualidade do serviço obtido durante toda a nossa estadia, em Istambul, foi consistentemente muito ruim, independentemente de se tratar de bens ou serviços – exceto talvez no Grande Bazar, e no mercado tradicional adjacente – onde um esforço de marketing de relacionamento com o cliente foi superior ao dedicado à venda puramente transacional.

Observação sobre inovação aberta e lição aprendida: Em similitude à situação descrita acima, os salários precisam ser mais altos, globalmente, o que deve levar a funcionários mais satisfeitos. Além disso, a noção de serviço precisa ser incutida nos serviços prestados aos cidadãos. As gorjetas devem ser conquistadas e merecidas e os

fornecedores de serviços precisam frequentar formação para melhorar os padrões de qualidade.

4.4. Episódio 4

Na receção do hotel, a seguir ao check-out, solicitamos um quarto adicional, pois o avião só levantava voo às 2 da manhã e, depois de um longo dia em visita por Istambul, queríamos tomar um duche. "Se deseja manter o quarto atual, terá que pagar dois quartos, pois são quartos interligados", informou-nos o gerente do hotel. Como queríamos um só quarto fomos obrigados a mudar de andar. O gerente do hotel tinha tentado impulsionar as vendas ao oferecer os quartos interligados – apesar de haver uma porta com uma fechadura entre estes, tornando-os verdadeiramente independentes. Assim, não aceitamos a oferta e preferimos aguentar o incómodo de ter de trocar as malas entre quartos em andares diferentes. Portanto, não houve um esforço para ser gentil e deixar-nos permanecer no quarto atual, ainda que a porta de comunicação entre quartos pudesse ser trancada e não interferisse com a chegada dos [eventuais] próximos clientes. "O hotel está cheio?", Perguntei. "Não, não está", respondeu o gerente do hotel. Paguei em dinheiro e nenhuma gentileza particular foi trocada. Nada além da interação necessária para a transação comercial ocorrer. Não havia empatia alguma entre nós.

Observação de inovação aberta e lição aprendida: Os hotéis são empresas e existem para ganhar dinheiro com turistas. A maioria dos turistas procura uma nova experiência todos os anos e, portanto, não serão clientes fiéis, por melhor que seja o serviço, em muitos dos casos. É, portanto, legítimo tentar ganhar mais dinheiro com os turistas. No entanto, fóruns on-line, como o TripAdvisor, são plataformas que reúnem comentários de turistas e essas são uma força a ser considerada. Portanto, os hotéis fariam bem em manter os seus clientes satisfeitos.

4.5. Episódio 5

"Eu dei ao taxista a minha palavra de que lhe ligaria para realizar a viagem de [25 minutos] de volta ao hotel [depois de jantar no restaurante]". Disse à minha esposa. "Então, vou manter a minha palavra e fazer exatamente isso". O taxista assegurou que demoraria cerca de 30 a 40 minutos, no máximo, para nos recolher. No entanto, 35 minutos depois, ainda aguardávamos. "Ligue para o taxista para ver se está tudo bem", afirmou a minha esposa. "Tive um problema, apanhe outro táxi por favor ", declarou o motorista de táxi, simplesmente, após lhe ter telefonado. Motoristas de táxi aparentemente não são confiáveis. Não se deu ao trabalho de telefonar para o meu número de telemóvel – o qual lhe tinha dado anteriormente. Poderíamos ter aguardado por um longo período de tempo, até percebermos que não estava vindo. Em abono da verdade, sempre que perguntamos no hotel como poderíamos chegar a diversos destinos, invariavelmente a sugestão era a utilização do elétrico ou metro. Os táxis deviam ser evitados, aparentemente, especialmente durante o dia (quando há também tráfego significativo e confusão nas ruas de Istambul). No entanto, pensamos que o problema está para além da questão do tráfego, como demonstra a nossa próxima história (episódio 6).

Observação sobre inovação aberta e lição aprendida: Os taxistas em Istambul talvez precisem receber formação para compreenderem a noção de lealdade do cliente. Tivemos bons motoristas de táxi – sendo que em alguns casos não podiam, no entanto,

falar inglês – o que constituía um problema à parte – o da comunicação eficaz num ambiente internacional.

4.6. Episódio 6

Esta ocorrência também envolve uma viagem de táxi. Noutra ocasião, eu e a minha filha, queríamos deslocarmo-nos para um destino, de manhã, que distava cerca de 3 km. Pedimos a três motoristas de táxi para nos levar, não sendo aceite o serviço por nenhum (não consideraram o serviço apelativo). Um quarto taxista aceitou, mas fixou o preço em 60 liras (cerca de 12 euros). Recusei-me, indignado, e acabámos por aceitar o conselho do rececionista do hotel e fomos de elétrico – por 5 liras cada um. O nosso destino estava somente a cinco paragens de onde nos encontrávamos. Mais tarde, ao final do dia, fomos de táxi para um restaurante que se encontrava a 25 minutos de distância. Essa viagem custou 47 liras. Sendo esta uma distância muito superior ao da manhã, o preço solicitado, anteriormente, pelo quarto taxista, foi ainda muito superior. Era, em suma, uma proposta que vejo como desonesta.

Observação sobre inovação aberta e lição aprendida: Em várias ocasiões, víamo-nos como meros contribuintes financeiros para os prestadores de serviços e nada mais além disso. Isso precisa de ser alterado, de modo a proporcionar uma experiência mais genuína e satisfatória ao cliente turista.

4.7. Episódio 7

Cerca de meia hora depois de sairmos do restaurante em Istambul, percebemos que tínhamos deixado um smartphone topo de gama na mesa de jantar. Corremos de volta. O telefone ainda estaria lá? Sim, o gerente tinha guardado o equipamento. Ficamos muito aliviados.

Observação sobre inovação aberta e lição aprendida: A honestidade é uma característica que pode ser encontrada em toda a parte. Teria sido fácil dizer que não tinham visto o telefone. O restaurante era muito escuro e era frequentado por muita clientela. Poderia ter sido fácil justificar não encontrar o smartphone. No entanto, eles optaram pela honestidade.

4.8. Episódio 8

A cultura, por exemplo nas Maldivas, era muito voltada para o futuro e de acordo com uma cultura mais global típica do Ocidente (Imagem 4). A importância da tecnologia é enfatizada e consequentemente a necessidade de socializar. Acima de tudo, pensar criticamente e criativamente é recompensado na escola Raa Atoll – Meedhoo island. A educação na ilha Meedhoo é gratuita para todos, tanto para os níveis básicos como para os níveis superiores, o que é muito louvável – se os alunos passarem. Se os alunos reprovarem, terão que pagar para realizarem novos exames. No entanto, a maioria das pessoas passa (obtendo uma avaliação de 50% ou mais). Apesar disso, a educação de nível universitário tem que ser autofinanciada. A escola também é acessível a todos, incluindo pessoas que se deslocam em cadeiras de rodas através da existência de rampas. Além disso, todos na ilha têm acesso a cuidados de saúde gratuitos. Estes são traços de sociedades avançadas e humanas. A religião (Islã) é um aspeto unificador na escola e na

ilha. Algo tido como positivo, pois ensina às crianças que existe um ser superior que está acima de nós e que procura a justiça e recompensa dos indivíduos que praticam o bem. Diversidade religiosa, no entanto, não é vista como uma opção.

Observação sobre inovação aberta e lição aprendida: Nós, no Ocidente, podemos pensar que encontramos uma solução para a vida em sociedade. Podemos pensar e julgar outras culturas a partir de uma posição de superioridade. No entanto, o que vi na ilha Meedhoo transcendeu o que vi no Ocidente. Eu vi uma sociedade harmoniosa e uma cultura maravilhosa, que não esqueceu as tradições.

Imagem 4 – As principais competências que as crianças devem aprender na Escola Raa Atoll – ilha Meedhoo – não são diferentes, em muitos aspetos, quando comparados com o que os países ocidentais tentam doutrinar

4.9. Episódio 9

Apesar de meros visitantes da ilha Meedhoo, durante o festival Eid, lá sentimo-nos bem-vindos e ninguém parecia ressentir-nos. Éramos o único grupo de turistas naquele dia, acompanhados por um guia turístico local. Os moradores permitiram fotografá-los, sozinhos e com membros do nosso grupo (Imagem 5). Foram acolhedores de uma maneira genuína, numa região que ainda não está deteriorada pela atividade comercial excessiva.

Observação sobre inovação aberta e lição aprendida: Existem outras culturas acolhedoras de turistas, como Portugal, atualmente um dos principais destinos turísticos da Europa. Talvez os países que estão menos preocupados em ganhar dinheiro acima de tudo, sejam bons destinos turísticos. Não há substituto para estar satisfeito com o que temos. Isto é o que eu vi na ilha Meedhoo.

Imagem 5 – À esquerda: mulheres em trajes tradicionais; À direita: *selfies* são realmente universais - ilha Meedhoo

4.10. Episódio 10

Durante o festival de Eid na ilha de Meedhoo, foi transportado um presente e, eventualmente, oferecido pelo homem responsável às mulheres. As mulheres devolveriam o gesto no dia seguinte, oferecendo, por sua vez, um presente aos homens (Imagem 6).

Observação sobre inovação aberta e lição aprendida: Isto é indicativo da igualdade de género e é, de facto, em si mesmo muito refrescante – especialmente depois da publicidade negativa que ouvimos sobre as culturas islâmicas – que, segundo o que vi, está incorreta. Não vi frustração nem infelicidade, ligada a diferenças de género.

Imagem 6 – Eid festival – À esquerda: Bailarinos que serão substituídos no dia seguinte por bailarinas; À direita: Um presente a ser oferecido pelos homens às mulheres – que devolveram o gesto no dia seguinte

5. Além da transação e a criação de relacionamentos com clientes [de longo prazo?]

Realmente a nossa estadia em Istambul foi bastante agradável, apesar do mau serviço que recebemos consistentemente. Sentimo-nos seguros em todos os momentos, embora tenhamos tomado algumas precauções, como evitar lugares escuros à noite. Como mencionado acima, no Grande Bazar e no mercado de especiarias adjacente, os comerciantes eram bons, esforçando-se para motivar os clientes na compra dos seus artigos através de uma relação cordial. Geralmente pareciam ser genuinamente amigáveis. Obviamente alcançam os seus objetivos, fruto do trato prestado aos turistas e da manutenção rotineira desse padrão. Ao longo da estadia, regressamos aos mesmos locais para repetir compras. Esta foi uma prova de como os relacionamentos foram desenvolvidos, apoiados por produtos de qualidade superior. Um dos membros do grupo, que se deslocara a Istambul, admitiu amar o local, apesar do serviço reconhecidamente ruim de alguns prestadores de serviços.

Assistimos a uma demonstração de força durante a estadia em Istambul, pelas autoridades e, consequentemente, pela tecnologia avançada. A Turquia teve alguns

problemas com o terrorismo, assim como alguns problemas políticos internos e agitação social (aparentemente resolvidos no momento em que este artigo é redigido) (Imagem 7). Durante uma estadia em Barcelona, em 2017, pouco antes do ataque a Las Ramblas, vimos muitos mais soldados e policias patrulhando as estradas e locais públicos do que vimos em Istambul em 2018.

Imagem 7 – À esquerda: Carro blindado patrulhando uma importante rua comercial pedestre em Istambul; À direita: Garantindo a segurança dos turistas – Um caminhão do exército localizado no centro de Istambul

As Maldivas cumpriram com o prometido – belas praias, sol, peixes nadando nos corais ao nosso lado. Paraíso, em uma palavra. No entanto, o que realmente chamou a atenção foi o quão agradável eram as pessoas – apesar de nós, turistas, virmos apropriar-nos das suas terras, das suas praias e da maior parte do seu país.

O nível de educação e a ordem geral das coisas (a clínica médica, a esquadra da polícia, o tribunal, entre outros) que vimos na ilha Meedhoo foram uma surpresa muito agradável. Jurei que voltaria algum dia. Fui fisgado por uma gestão de relacionamento superior com o cliente – ou, se preferir, pela bondade humana genuína.

6. Discussão e conclusões

Sobre o grande escritor do século 20 George Orwell, Ricks (2017, p.128) escreveu, referindo-se a um período durante a Batalha da Grã-Bretanha, por volta de 1940-1941, que "Orwell sempre adorou observar, e agora havia tanto novo e diferente para ver e contemplar". Esta tem sido a nossa metodologia de investigação – observar, enquanto participamos.

Além disso, pesquisar ou escrever não é fácil, como Orwell comentou: "Para ver o que está à frente do nosso nariz obriga a uma luta constante" (Ricks, 2017, p. 40).

Procuramos com este estudo discutir como as culturas inovam ao longo do tempo, misturando e aprendendo com outras culturas. Que as culturas adotam práticas de outras culturas é evidente nas fotografias tiradas. Podemos confortavelmente afirmar que a modernidade, por exemplo, incorporada no smartphone, foi adotada mesmo nas

comunidades mais remotas (incluindo a ilha Meedhoo). Lembremos que o primeiro smartphone comercializado com sucesso em todo o mundo foi o iPhone 1, da Apple, em 2007, produto desenvolvido e supervisionado pelo falecido Steve Jobs que, como mostram as nossas evidências, conseguiu realmente mudar o mundo. Em Istambul, vimos turistas visitarem e gravarem imagens de diversos cenários com smartphones, como faria qualquer outro turista. Na ilha de Meedhoo vimos os residentes utilizarem smartphones muito sofisticados para gravar eventos.

Se a cultura é "O que resta quando o que foi aprendido é inteiramente esquecido" (Selma Lagerlof (Petit, 1960, p.100) as quoted in Usunier, 1998, p.16) então sim, as culturas permanecem inalteradas, mas também vimos evidências de culturas sustendo-se ao longo dos tempos. E isso só pode representar algo bom.

A tecnologia é vista como um fator de união entre as culturas. Juntamente com a bondade humana inata (visível nos olhos das pessoas nas várias imagens). Todos nós nos aproximamos desses fatores e procuramos satisfazer as mesmas necessidades. Interagir, gravar, comunicar, desfrutar – são todas interações que garantem que as culturas nacionais e regionais inovem ao longo do tempo, estimuladas pela tecnologia.

Segundo a pesquisa de Salehan, Kim e Lee (2018), existe uma relação entre o aspeto sociocultural da globalização e a tecnologia. Isto é, a tecnologia afeta valores culturais nacionais. Este estudo em particular também revelou que a tecnologia é um condutor (ou "driver", em inglês) importante da convergência cultural. Devido à tecnologia culturas nacionais ficarão, em concreto, mais individualistas e com menor distância de poder. O nosso estudo confirma estes resultados, pois mostramos indivíduos a guardar imagens e vídeos e a interagir com o ambiente através do seu próprio telemóvel (smartphone) ao invés de em grupo (Imagem 2); enquanto que os indivíduos se assumem mais próximos uns dos outros, incluindo de turistas do Ocidente – o suficiente para tirarem *selfies* juntos, com toda a naturalidade e apesar de terem religiões, línguas, e vivências aparentemente muito diferentes (Imagens 3 e 5).

Encorajamos jovens investigadores, em particular, a observarem enquanto interagem em situações novas e a gravar e comentar sobre as observações. Tal exercício irá tornar-nos mais conscientes do que está a ocorrer ao nosso redor – envolvendo uma maior consciência da mudança à medida que esta ocorre.

Embora haja mais interações do que possa aparentar, as imagens são capazes de contar histórias que valem mil palavras, cada uma. Num cenário ideal – sem a limitação de tempo ou restrições financeiras – a observação pode significar imergir na cultura local por períodos de um ano ou mais. Enquanto o falecido George Orwell fez isso em variadas ocasiões (Ricks, 2017), esse tipo de observação, compreensivelmente, não é para todos. Até que ponto estamos preparados para ir em nome da ciência e para contribuir para o avanço do conhecimento das sociedades e influenciar o modo como vivemos no seu seio? Uma pergunta difícil, de facto, à qual só o leitor, individualmente, poderá responder.

A Figura 1 resume o que observamos em relação às culturas e como elas inovam e evoluem para se manterem vivas. Quando a Cultura A encontra a Cultura B, elas efetivamente misturam-se e aprendem através dessa interação. Resultados da inovação – novos hábitos, símbolos e valores. A Cultura A torna-se Cultura A1 e a Cultura B torna-se Cultura B1. O papel da tecnologia é importante, pois funciona como catalisador para

aproximar pessoas desconhecidas e culturalmente distantes. Por exemplo, o fenómeno "*selfie*" é global – garantindo a captura de momentos do cotidiano não excluindo ninguém da fotografia. É o verdadeiro "momento de partilha" nas interações de grupo. Escolhemos designar ao processo de evolução das culturas, auxiliadas pela adoção da tecnologia, de fenómeno "Mix-Tech".

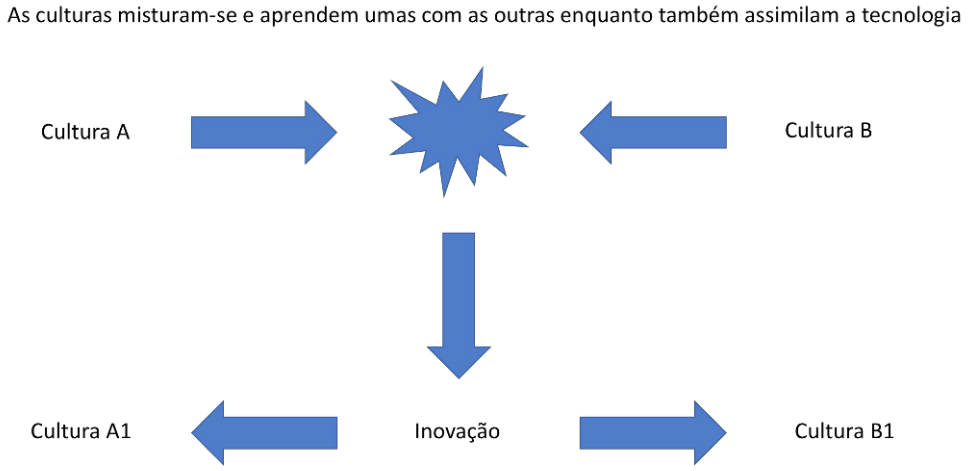


Figura 1 – O fenómeno "Mix-Tech" – As culturas evoluem misturando e aprendendo umas com as outras e convergem auxiliadas pela adoção de tecnologia

7. Limitações e sugestões para investigações futuras

As limitações do estudo incluem o período de tempo reduzido para a sua realização – cerca de duas semanas de observações. Sugerimos que estudos futuros sejam conduzidos por um período mais alargado de tempo. O caso atual foi limitado no tempo por fatores financeiros e outros. Finalmente, seria interessante entrevistar mulheres e homens da Turquia e das Maldivas para indagar sobre os seus pontos de vista relativos à inovação cultural e o papel da tecnologia na sociedade.

Agradecimentos

Agradecemos ao nosso guia local na ilha Meedhoo – Mohamed Yamaanee – por ter partilhado connosco o seu conhecimento sobre o seu povo e a sua cultura. Gostaríamos também de agradecer ao João Barreiros pela sua revisão e crítica muito positiva de uma versão anterior do artigo.

Referências

Alvesson, M., & Sveningsson, S. (2015). *Changing organizational culture: Cultural change work in progress*: Routledge.

Dočekal, V., & Tulinská, H. (2015). The impact of technology on education theory. *Procedia-Social and Behavioral Sciences, 174*, 3765-3771.

Hofstede Insights. (2018). What about Turkey? Retrieved from https://www.hofstede-insights.com/country/turkey/

House, R. J., Hanges, P. J., Javidan, M., Dorfman, P. W., & Gupta, V. (2004). *Culture, leadership, and organizations: The GLOBE study of 62 societies*: Sage publications.

Huizingh, E. K. (2011). Open innovation: State of the art and future perspectives. *Technovation, 31*(1), 2-9.

Human Development Reports. (2017). Human Development Index (HDI). *United Nations Development Programme.* Retrieved from http://hdr.undp.org/en/content/human-development-index-hdi

Kratzer, J., Meissner, D., & Roud, V. (2017). Open innovation and company culture: Internal openness makes the difference. *Technological Forecasting and Social Change, 119*, 128-138.

Laursen, K., & Salter, A. J. (2014). The paradox of openness: Appropriability, external search and collaboration. *Research Policy, 43*(5), 867-878.

Ricks, T. E. (2017). *Churchill & Orwell – The fight for freedom*. New York: Penguin Books.

Salehan, M., Kim, D.J., Lee, J-N. (2018). Are there any relationships between technology and cultural values? A country-level trend study of the association between information communication technology and cultural values. *Information & Management, 55*, 725-745.

Saunders, M. N., & Lewis, P. (2018). *Doing research in business & management: An essential guide to planning your project* (2nd edition ed.). Harlow: Pearson.

Schein, E. H. (2010). *Organizational culture and leadership* (Vol. 2): John Wiley & Sons.

Usunier, J.-C. (1998). *International and cross-cultural management research*: Sage.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 27/09/2018
Aceitação/Acceptance: 16/11/2018

Sistema de Gestión de las emociones en un salón de clases inteligente basado en Modelos de confianza y reputación

Jorge Cordero[1], Jose Aguilar[2], Kristell Aguilar[2], Miguel Martinez[2]

jmcordero@utpl.edu.ec, aguilar@ula.ve, kristell153@gmail.com, martinezpmg9@gmail.com

[1] DCCE, Universidad Técnica Particular de Loja, San Cayetano Alto, 11-01-608, Loja, Ecuador.

[2] CEMISID, Universidad de Los Andes, La Hechicera, 5101, Mérida, Venezuela.

Pages: 870–885

Resumen: En este artículo, se propone un modelo de confianza y otro de reputación, para la gestión de las emociones en un Salón de Clases Inteligente (SaCI). En general, un SaCI puede modelar como agentes a sus diferentes componentes de hardware y de software. Por otro lado, en los ambientes inteligentes modelados con agentes, es necesario definir los mecanismos de interacción usados en las conversaciones entre los agentes. Además, las conversaciones deben ser enriquecidas con modelos de confianza y reputación de los agentes involucrados en las mismas, para que cada agente pueda decidir qué información, procedente de los distintos agentes, considerar al momento de ejecutar sus procesos de toma de decisión. Particularmente, se desarrollan modelos de confianza y reputación para los agentes de SaCI. Esos modelos se usan en una conversación que realiza la gestión de las emociones en SaCI, la cual permite definir la emoción que prevalece en un momento dado en él, a la que llamamos emoción social. La conversación es probada en varios casos de estudio, con la finalidad de validarla su capacidad para establecer la emoción social de SaCI. Los resultados experimentales con el sistema de gestión de las emociones basado en modelos de confianza y reputación, demuestran que es capaz de determinar la emoción social en SaCI sin importar las características de su contexto y aprendiendo del comportamiento de los agentes, lo que le da una gran robustez.

Palabras-clave: modelo de confianza, modelo de reputación, gestión de emociones, emoción social, aula inteligente, sistemas multiagente.

Approaches to identify student learning styles through emotions in a classroom

Abstract: In this article, we propose a model of trust and reputation, for the management of emotions in a Smart Classroom (SaCI). In general, a SaCI can model its different hardware and software components as agents. On the other hand, in intelligent environments modeled with agents, it is necessary to define the interaction mechanisms used in conversations between agents. In addition, the conversations should be enriched with models of trust and reputation of the agents

involved in them, so that each agent can decide what information coming from the different agents, to consider when executing their decision-making processes. In particular, this paper defines models of trust and reputation for the SaCI agents. These models are used in a conversation that performs the management of emotions in SaCI, which allows to define the emotion that prevails at a given moment in it, which we call social emotion. The conversation is tested in several cases of study, in order to validate its ability to establish the social emotion of SaCI. The experimental results with the emotion management system based on models of trust and reputation, demonstrate that it is able to determine the social emotion in SaCI regardless of the characteristics of its context and learning from the behavior of agents, which gives it great robustness.

***Keywords**:* trust model, reputation model, emotion management, social emotion, smart classroom, multiagent systems.

1. Introducción

La confianza y la reputación se han convertido en criterios cruciales para seleccionar socios en entornos abiertos, distribuidos y dinámicos, como los Ambientes Inteligentes (AmI) y los Sistemas MultiAgente (SMA). En la literatura, se han propuesto muchos mecanismos computacionales de confianza y de reputación, para SMA (Lijian et al/, 2010; Nguyen et al., 2013, 2015a, 2015b) En general, la confianza y reputación se han usado para ayudar a los agentes a tomar decisiones sobre cuanto creer sobre su entorno (otros agentes, información del ambiente, etc.) en diferentes contextos (D'Angelo et al, 2017)

Un AmI tiene como finalidad implementar espacios tecnológicamente enriquecidos, que apoyen a los usuarios en su vida diaria. Existen múltiples aplicaciones de los AmI: hogar inteligente, servicios en el sector de salud, de gobierno, de aula, etc. En particular, en este trabajo estamos interesados en un Salón de Clases Inteligente (SaCI), el cual es un aula de clase que integra dispositivos tecnológicos, software especializado, sensores, conectividad, entre otras cosas, para coadyuvar al proceso de enseñanza-aprendizaje. Los diferentes componentes de SaCI pueden ser modelados como agentes, que interactúan entre sí y con agentes externos, para facilitar y mejorar los procesos de enseñanza y aprendizaje (Valdiviezo et al., 2015; M. Sánchez et al , 2015a). Así, SaCI modela como agentes a sus diferentes componentes de hardware y de software. Todos estos agentes deben colaborar entre sí, con el fin de mejorar el proceso de enseñanza-aprendizaje que se da en SaCI. Por ejemplo, para realizar una adecuación efectiva de las condiciones ambientales por parte de los agentes encargados de dicha tarea, es necesario la comunicación con los dispositivos que adecuan la temperatura y luz ambiental, entre otros.

En los SMA los distintos agentes deben desempeñar distintos roles, y seguir ciertas normas y estándares, para garantizar un desempeño óptimo del sistema como un todo. En el caso de las interacciones con otros agentes, algunas de las preguntas que un agente podría hacerse son ¿Hasta qué punto puedo confiar en un agente? ¿Qué tan confiable es la información que me suministra un agente? Para solventar estos problemas, los agentes pueden usar modelos de confianza y reputación, los cuales se deberían emplear en las distintas conversaciones en las que participan, para guiar sus decisiones. En específico, los modelos de confianza y reputación han sido comúnmente usados para inferir el comportamiento futuro de otras entidades a partir de información basada en experiencias pasadas (Hermoso et al., 2013). En el contexto de SMA, cada agente

es consumidor o proveedor de servicios. Estos agentes tienen sus propios objetivos e intereses y, por lo tanto, no son completamente fiables. Es por esto que es necesario el uso de modelos de confianza y reputación sobre los agentes en un ambiente distribuido.

En cuanto a las emociones en el contexto de AmI, para que las emociones contribuyan en él, es necesario definir un modelo para reconocer y gestionar las emociones. En esta investigación, se presenta un modelo para gestionar las emociones en AmI. El problema de gestión implica determinar y filtrar las fuentes de información emocional en el ambiente, manejar los múltiples estados emocionales de los diferentes usuarios, usar la información sobre las emociones para mejorar los procesos adaptativos del AmI, entre otros aspectos.

En este artículo, se analiza como valorar las emociones de cada agente, como priorizarlas, como determinar su validez, usando para ello modelos de confianza y reputación. Para ello, se usa el concepto de consenso social (punto de referencia social) propuesto en (Aguilar, 2017a), aplicado a SaCI, con el fin de obtener un consenso sobre la emoción imperante, considerando lo aportado por todos los agentes. Los agentes decisores definirán este punto de referencia (consenso social), para establecer un compromiso entre todos, para establecer un modelo de enseñanza-aprendizaje colectivo. Así, la idea es que los agentes decisores puedan establecer ese punto de consenso emocional en SaCI, basado en los modelos de confianza y de reputación de los agentes presentes en las conversaciones.

En los trabajos revisados en la literatura ((Burnett et al. , 2010; Hang et al., 2009; Hermoso et al., 2013; Victor et al., 2009; Nguyen et al., 2015a, 2015b), se pone de manifiesto que los mismos no han sido usados para construir consenso social en un AmI, ni para gestionar las emociones en dicho ambiente. Así, este trabajo se propone el uso de modelos de confianza y reputación para la gestión de las emociones en un SaCI. Cada conversación o interacción usa los modelos de confianza y reputación de los agentes, para que ellos puedan decidir qué información considerar, al momento de procesarlas y determinar la emoción social imperante en SaCI. En particular, en este trabajo se proponen usar los modelos de tablas de interacción local (o LIT por sus siglas en inglés Local Interaction Tables) y de preferencia, para la gestión de las emociones. LIT es usado para representar la confianza en las emociones de un agente, y el modelo de preferencia es usado para determinar la reputación de las emociones de un agente. A partir de allí, en este trabajo se propone un esquema para que un agente pueda definir la emoción que prevalece en un momento dado en SaCI, a lo que llamamos emoción social.

2. Marco conceptual

2.1. Salón de Clases Inteligente (SaCI)

SaCI es un Salón de Clases Inteligente propuesto en (Valdiviezo et al., 2015; M. Sánchez et al., 2015a), donde se define el modelo conceptual de un aula inteligente centrada en el estudiante, sensible a contexto, que apoya el proceso de enseñanza-aprendizaje a través de dispositivos colaborativos y aplicaciones que facilitan el autoaprendizaje. Para hacer esto, SaCI tiene diferentes tipos de componentes de hardware (p. ej., cámaras, pizarra interactiva, robot social, sensores, etc.) y de software (p. ej., entorno virtual de

aprendizaje (VLE), repositorios de objetos de aprendizaje, sistemas recomendadores, sistema de recursos educativos abiertos, entre otros), que se adaptan e integran en el aula de acuerdo con las necesidades de los estudiantes.

Por otro lado, en (M. Sánchez et al., 2015b) los autores presentan un middleware basado en el paradigma de SMA, que permite administrar a SaCI. El middleware propone tres niveles, uno para la gestión de la comunidad multiagente, otro para gestionar el acceso a servicios, aplicaciones, etc., y el último para caracterizar los diferentes componentes (software y hardware) del entorno educativo inteligente.

Por otro lado, los agentes de SaCI generan una gran cantidad de datos sobre los procesos de aprendizaje, que pueden ser explotados para mejorar dicho proceso. En ese sentido, en (J. Aguilar et al., 2017b; Aguilar et al, 2017c; M. Sánchez et al., 2016) se ha extendido SaCI con los paradigmas de la computación en nube y analítica de aprendizaje, donde proporcionan servicios de análisis de aprendizaje en la nube basado en los datos capturados en SaCI. Los servicios de análisis de aprendizaje han permitido, generar conocimiento sobre los perfiles de estudiantes, detectar necesidades particulares en un curso, entre otras cosas.

2.2. Modelo de Confianza en SMA.

Existen muchos modelos de confianza en la literatura, en este artículo, se propone usar el modelo de confianza basado en Tablas de Interacción Local (LIT, por sus siglas en inglés *Local Interaction Tables*) (Aguilar, 2014). El modelo LIT es un modelo de confianza que parte de la idea que un agente posee un rol en la comunidad, realiza interacciones en la misma, con la suposición que este agente conoce la estructura de la organización. En este caso, se representa que un agente A evalúa la confianza que tiene en otro agente B, jugando este último el rol R en la tarea T. La confianza se denota como tA ® (B, R, T) Î[0..1]. Dentro de este rango, valores cercanos a uno representan para un agente una confianza más alta que valores cercanos a cero.

En el caso del modelo LIT, cada agente mantiene una tabla, en la que va guardando los resultados de las interacciones pasadas. Esa tabla tiene un índice (clave) del tipo agente/rol/interacción (X,Y,Z), y dos columnas con valores asociados de confianza (CA) y confiabilidad (rA). En la tabla, (X,Y,Z) define que un "agente X está jugando el rol Y en la tarea Z"; la confianza en que eso sea cierto es definida por CAà (X,Y,Z), y se actualiza por la propia experiencia del agente; y la confiabilidad rAà (X,Y,Z) mide que tan seguro está el agente acerca de su propio valor de confianza para esa situación (X,Y,Z). Por otro lado, el agente debe calcular el valor de confianza usando una función que evalúe las experiencias pasadas en esa situación determinada (X,Y,Z). El macroalgoritmo basado en el modelo LIT, parte del principio de que: "… un agente A confía que un agente X puede tener el rol Y para hacer la actividad Z". Si no existe esa entrada (Y, Z) para un agente X en la tabla LIT del agente A, ese agente puede:

1. Buscar otro agente similar a X, bajo el principio de que los agentes se comportan de manera similar cuando juegan roles similares o actúan en interacciones similares, por lo que eventualmente también se puede usar el valor conocido de otro agente similar del que si se tiene esa entrada de confianza.
2. De lo contrario, se acude al modelo de reputación del Agente X.

En general, para la determinación de la confianza, se usa la similitud de las tareas y de los roles en diferentes situaciones. Entre más similares son dos tareas T y T' o roles R y R', se supone que más similar será el comportamiento de los agentes realizando esas tareas o con esos roles. Dada una situación (X, R, T) y un agente A, la confianza que estima A sobre la situación (B, R, T) utilizando una especie de modelo de *inferencia*, se define como:

$$t_A \rightarrow (X, \mathrm{R}, T) = \frac{\sum_{(X,Y,Z) \in LIT_A} C_A \rightarrow (X,Y,Z) \cdot W_A \rightarrow (X,Y,Z)}{\sum_{(X,Y,Z) \in LIT_A} W_A \rightarrow (X,Y,Z)} \quad (1)$$

donde, $W_A \rightarrow (X,Y,Z)$ es el peso del valor de confianza del agente A en la situación (X, Y, Z). Los pesos se según ecuación 2.

$$W_A \rightarrow (X,Y,Z) = r_A \rightarrow (X,Y,Z)\ \mathrm{sim}((Y,Z),(R,\mathrm{T})) \quad (2)$$

El cálculo de la similitud se evalúa como una suma ponderada de las similitudes de los elementos individuales de cada situación (Y, Z) con las de la situación que se quiere decidir (R, T):

$$sim\ ((Y,Z),(\ R,T) = \beta\ \cdot\ sim_R(Y,R)\ +\ \gamma\ \cdot\ sim_T(Z,T) \quad (3)$$

donde $sim_R(Y,R), sim_T(Z,T) \in [0..1]$ miden la similitud entre los roles y las tareas, β y γ, con $\beta + \gamma = 1$, son parámetros que especifican la sensibilidad de cada similitud de manera individual.

Por otro lado, una vez realizada la actividad solicitada al agente X, hay un proceso de aprendizaje, para actualizar la Tabla 1. El valor de confianza se actualiza como:

$$C_A \rightarrow (X, \mathrm{R}, T) = \epsilon \cdot C_A \rightarrow (X,Y,Z) + (1-\epsilon) \cdot\ g_A(X,Y,Z) \quad (4)$$

Donde ϵ especifica la importancia que el agente le da al valor de confianza calculado anteriormente con respecto a la evaluación actual (tiene un efecto de evaporación del valor de C_A), y gA(X,Y,Z) es la ganancia que obtuvo A por la acción realizada por el agente X en el rol Y en la tarea Z (puede ser positiva o negativa).

Y el de confiabilidad como:

$$\mathrm{rA}\rightarrow (\mathrm{X,Y,Z}) = \alpha \mathrm{rA}\rightarrow (\mathrm{X,Y,Z}) + (1\text{-}a)(\mathrm{I(X,Y,Z)}) \quad (5)$$

donde a es la taza de reforzamiento, e I(X,Y,Z) es el número de instancias (X,Y,Z) que ha tenido el agente A con respecto a su total de instancias. Esos valores siempre estarán entre 0 y 1.

2.3. Modelo de Reputación en SMA.

En general, la *reputación* es el *prestigio o estima* que se tienen de alguien o algo. La reputación de un agente se conforma de acuerdo con su accionar en el tiempo (Ramchurn et al., 2004). Particularmente, los modelos de reputación basados en preferencias (Aguilar, 2014). Este método asume que todas las opciones son conocidas por todos los individuos, y que lo que resta es determinar sus preferencias. Este mecanismo elige la opción, en base a lo que votan todos los agentes. En el caso más simple, cada individuo tiene un voto, que se lo da a la opción que este individuo prefiere. Agregando todos los votos, se determina las preferencias relativas entre las

diversas alternativas. Este protocolo es conocido como el protocolo de la pluralidad (la opción con mayor votos gana).

En una versión más sofisticada del mecanismo de votación, los individuos pueden distribuir su votación sobre diversas alternativas, proporcionalmente a sus preferencias individuales. Por ejemplo, la alternativa A pudiera conseguir un voto de 0.5, B de 0.3, C de 0.2 y D de 0.0. En ese caso, la función de preferencia colectiva Pcol sobre la opción lj de un grupo de opciones L, es el promedio de las n funciones de preferencias individuales Pi (Aguilar, 2014):

$$P_{col}(l_j) = \frac{1}{n} \sum_{i=1}^{n} P_i(l_j) = \frac{1}{n} \sum_{i=1}^{n} P_i^j \quad (6)$$

El modelo de preferencia de la ecuación 6, utiliza una función de utilidad que representa numéricamente las preferencias individuales de cada agente. Luego, dichas preferencias son unificadas, obteniéndose una decisión que beneficie a la colectividad. El mecanismo de funcionamiento de este modelo es el siguiente:

1. Se determina la preferencia de cada individuo para hacer ciertas cosas.
2. Se procesa el grupo de preferencias de los individuos en la comunidad. Ella representa la preferencia colectiva.

Inspirado en ese modelo de preferencias se construye el modelo de reputación en este trabajo (Aguilar, 2014). Como modelo de reputación, el valor de preferencia P_i^j sintetiza lo que cree cada agente i de otro agente lj, y a partir de allí se obtiene la reputación general de ese agente ($P_{col}(l_j)$).

3. Formalización de la gestión de la emoción social en SaCI usando modelos de confianza y reputación

En esta sección se presenta la formalización de los diferentes elementos que intervendrán en el proceso de determinación de la emoción social en SaCI. En específico, el objetivo es que el VLE pueda conocer el estado de emoción social en SaCI, y a partir de allí presentar recursos de aprendizaje y desarrollar actividades educativas, que permitan mejorar el proceso de aprendizaje. En este caso, se usan todos los agentes que representan dispositivos estáticos en SaCI (pizarra interactiva, cámara), para que envíen información que pueda ser usada para determinar la emoción social. Esos agentes envían la información que perciben al agente gestor de emociones de SaCI, el cual después lo envía al agente VLE, para que él pueda proveer objetos de aprendizaje adecuados al estado emocional de los estudiantes. Con base en la definición dada en (Aguilar et al. 2017a), la emoción social se calcula como una combinación lineal de la emoción reportada por cada agente que representa un dispositivo estático, según la siguiente ecuación.

$$SP_s = \sum_{i=1}^{n} \frac{k_i SPA_i}{} \quad (7)$$

Donde, SPA es la emoción determinada por cada agente i, k_i es el peso que se le da a esa opinión del agente i, basado en su modelo de confianza o reputación. El agente gestor de emociones primero busca en su tabla LIT el valor de confianza, y si no lo consigue o es muy baja la confianza en ese agente para ese rol y tarea, usa entonces el modelo de reputación.

En particular, la conversación donde ocurre el proceso de consenso social sobre la emoción imperante en un momento dado en SaCI, en la cual se usan los modelos de confianza y reputación, es definida en la Tabla 1 y Figura 3.

Elemento	Valor
Objetivo	Reconocer la emoción social de los estudiantes presentes en el aula.
Agentes participantes	gestor de emociones, pizarra interactiva, pizarra personal, escritorio interactivo, cámara
Iniciador	gestor de emociones
Datos intercambiados	Emoción de los estudiantes (SPA).
Precondición	Presencia de estudiantes en SaCI.
Condición de terminación	Realizar acciones de acuerdo con la emoción.
Descripción	Los agentes dispositivos proporcionan el estado emocional de los estudiantes que logran percibir, para que el agente gestor de emociones pueda determinar la emoción social, y después se la informe al VLE para que tome acciones sobre qué recursos de aprendizaje son adecuados para SaCI en ese momento.

Tabla 1 – Conversación para obtener la emoción social

En la conversación de la Figura 1, el agente gestor de emociones solicita las emociones reconocida por cada agente dispositivo de SaCI ($SPA_i SPA_i$), pondera cada una según su modelo de confianza o de reputación de cada agente (ver Tablas 2 y 3), y a partir de allí, usando la ecuación 7, determina la emoción social en ese momento en SaCI. La emoción social que consigue, después se la transmite al agente VLE para que la use en sus procesos de toma de decisión.

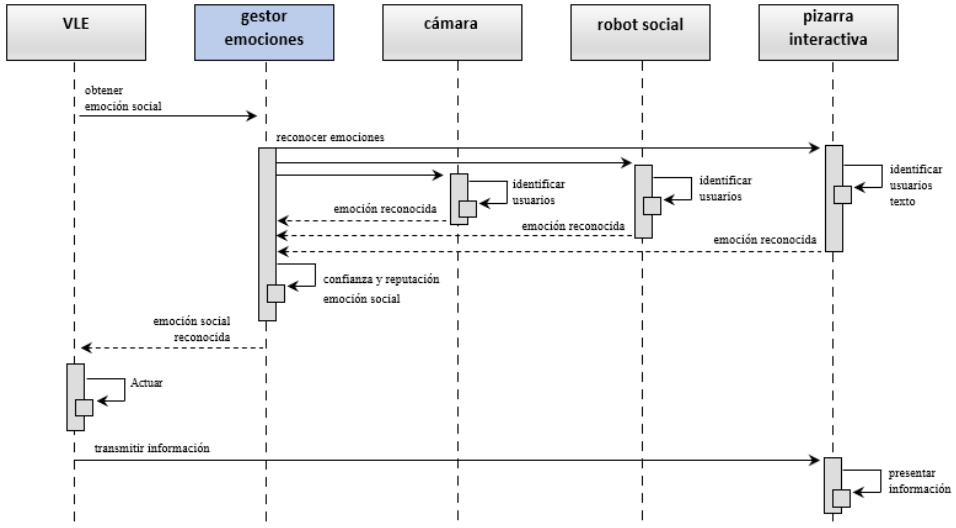


Figura 1 – Conversación para obtener la emoción social

En la Tabla 2, se observa el modelo de confianza del agente gestor de emociones con cada uno de los agentes de SaCI con los que interactúa para determinar la emoción social. En particular, las columnas sim(Y,R) y sim(Z,T) especifican la similitud entre la tarea de reconocimiento de emociones y las otras tareas que puede ejecutar cada agente, y la similitud entre los roles que puede desempeñar cada agente y el que es de interés para reconocer las emociones en SaCI (identificar). Los otros valores son calculados y actualizados, según como se establece en la sección 2.3.

Cod	Agente	R(Rol)	T(Tarea)	sim (Y,R)	sim (Z,T)	CA	rA	WA	tA→(X,R,T)
A1	Profesor	enseñar	proporcionar conocimiento	0.14	0.11	0.47	0.17	0.01	0.459
A1	Profesor	observar	observar comportamiento del estudiante	0.22	0.25	0.45	0.15	0.01	
A2	Estudiante	aprender	adquirir conocimiento	0.03	0.04	0.47	0.17	0.00	0.456
A2	Estudiante	autoformación	resolver cuestionarios	0.10	0.15	0.45	0.15	0.01	
A3	Cámara	identificar	identificar usuarios mediante expresiones faciales	1.00	0.90	0.85	0.75	0.41	0.855
A4	GestorEmociones	reconocer	reconocer emociones básicas	0.00	0.00	0.00	0.00	0.00	0
A4	GestorEmociones	gestionar	obtener emoción social	0.00	0.00	0.00	0.00	0.00	
A5	VLE	gestionar	gestionar recursos y actividades de aprendizaje	0.12	0.35	0.46	0.66	0.07	0.461
A5	VLE	facilitar	facilita comunicación multidireccional	0.19	0.18	0.45	0.65	0.07	
A6	RobotSocial	identificar	identificar usuarios (expresiones faciales)	1.00	0.90	0.20	0.75	0.41	0.415
A6	RobotSocial	identificar	identificar usuarios (voz)	1.00	0.85	0.17	0.75	0.40	
A6	RobotSocial	identificar	identificar usuarios (gestos)	1.00	0.80	0.15	0.65	0.34	
A7	PizarraInteractiva	presentar	presentar información, REAs	0.15	0.14	0.85	0.87	0.07	0.271
A7	PizarraInteractiva	identificar	identificar usuarios (texto)	0.38	0.42	0.75	0.85	0.42	

Tabla 2 – Confianza entre el agente gestor de emociones y el resto de agentes de SaCI

La Tabla 3 muestra el modelo general de reputación de alguno de los agentes de SaCI, en donde se incluyen los agentes con los que interactúa el agente gestor de emociones. El

valor $P_{coi}(l_i)$ en esa tabla se calcula usando la ecuación 6, tal que P_i^j es la reputación que tiene el agente l_j, en el agente i.

Cod	Agente	P_{col}
A1	Profesor	enseñar
A2	Estudiante	vigilar
A3	Cámara	observar
A4	GestorEmociones	aprender
A5	VLE	opinar
A6	RobotSocial	autoformación
A7	PizarraInteractiva	identificar
A8	RecomendarOA	reconocer
	...	

Tabla 3 – Modelo de Reputación de los agentes de SaCI

La importancia de determinar la emoción social en SaCI, es porque le permite al agente VLE definir los recursos y dinámicas de aprendizaje adecuadas para el grupo de estudiantes en SaCI en un momento dado. Esta conversación que determina la emoción social permite obtener esa emoción social, que después otras conversaciones presentes en SaCI las usan en sus interacciones (por ejemplo, la conversación en la que el agente VLE gestiona la dinámica de aprendizaje en SaCI, ver (Aguilar et al., 2017a; Valdiviezo et al., 2015) para detalles de ella).

En general, el objetivo principal de los modelos de confianza y reputación son ofrecer mecanismos para apoyar procesos adaptativos entre los agentes, y disminuir al máximo los efectos negativos de los agentes poco fiables.

4. Experimentación

4.1. Escenario 1: Condiciones ideales

En este caso, se suponen condiciones ideales en SaCI, donde todos sus agentes funcionan correctamente, y pueden comunicarse sin ningún tipo de problema. La Tabla 4 representa el modelo de confianza del agente gestor de emociones para ese escenario en ese momento. Se observa que la cámara proporciona una confianza superior para el proceso de reconocer la emoción. La Tabla 5 describe la reputación general de los agentes con los que interactúa el agente gestor de emociones en ese momento, en esa conversación.

La Tabla 6 describe el proceso de determinación de la emoción social en SaCI. Los valores de las emociones son obtenidos del sistema de reconocimiento multimodal propuesto en (Cordero et al., 2016). En este caso, vemos que la emoción que prevalece en ese momento que se analiza es la sorpresa (46.3%), determinada primordialmente por el agente Cámara. Además, el agente gestor de la emoción social tiene alta confianza en

ese agente (0.85). Con respecto a la felicidad, es reconocida principalmente por el agente RobotSocial, del cual confía medianamente el agente gestor de emociones (0.41), por lo que se usa su modelo de reputación en la comunidad de agentes de SaCI. En cuanto al agente Pizarra Interactiva, tanto el agente gestor de emociones como la comunidad confían poco en él (0.27 y 0.26, respectivamente), por lo que influye poco en el proceso de determinación de la emoción social. De esta manera, el agente gestor de emociones confía esencialmente en el agente cámara, tal que tendrá mucha influencia en la emoción social final reconocida. El agente gestor de emociones le informará al agente VLE que las emociones sociales imperantes en SaCI son la sorpresa (46.3%), principalmente, la felicidad (17.8%) y el miedo (16.2%). Además, con el modelo propuesto se obtiene que en un 16.47% no se sabe cuál es la emoción social, lo cual es relativamente bajo.

Agente	tA → (X,R,T)
Profesor	0.459
Estudiante	0.456
Cámara	0.855
RobotSocial	0.415
PizarraInteractiva	0.271

Tabla 4 – Modelo de confianza del agente gestor de emociones

Agente	Reputación
Profesor	0.459
Estudiante	0.456
Cámara	0.855
RobotSocial	0.415
PizarraInteractiva	0.271

Tabla 5 – Reputación entre los agentes

Cod	Agentes	tA→(X,R,T)	Reputación	Felicidad	Tristeza	Ira	Sorpresa	Miedo
A3	Cámara	0.855	0.491	20.14	2.65	5.41	58.26	18.92
A6	RobotSocial	0.415	0.412	22.65	1.42	2.89	29.03	14.11
A7	PizarraInteractiva	0.271	0.268	2.45	0.61	1.17	0.87	2.15
			SPA	17.842	1.002	2.037	46.384	16.288

Tabla 6 – Emoción Social

4.2. Escenario 2: Condiciones de falla de agentes

En este caso, se suponen fallas en la cámara y una situación diferente en SaCI. La Tabla 7 describe el modelo de confianza del agente gestor de emociones para este escenario para esa nueva situación en SaCI. Se observa que la confianza en el agente cámara disminuye significativamente a un 0.051, por lo tanto, se limita su aporte en el proceso de reconocimiento de la emoción social. La reputación del agente disminuye drásticamente debido a la falla. La Tabla 8 describe la reputación general de los agentes con los que interactúa el agente gestor de emociones en ese momento, para este escenario en esa conversación. La Tabla 9 describe el proceso de obtención de la emoción social en condiciones en las que falla la cámara. Particularmente, al disminuir la confianza, entra en funcionamiento el modelo de reputación de ese agente. Ahora bien, al ser bajo también, implica que lo que ese agente reconozca no será considerado.

Para este caso, el agente gestor de emociones considera principalmente para determinar la emoción social lo que reconoce el Robot Social, ya que su valor de confianza es

relativamente alto (0.707). El agente reconoce principalmente como emociones sociales la ira (41.3%) y la tristeza (18.14), basado en lo que reconoce el robot social. Pero, además, reconoce que hay Miedo (13.4%), porque la Pizarra Interactiva lo detecto, y ese agente tiene una muy buena reputación en la comunidad (0.67).

Agente	tA → (X,R,T)
Profesor	0.459
Estudiante	0.456
Cámara	0.051
RobotSocial	0.607
PizarraInteractiva	0.296

Tabla 7 – Modelo de confianza del agente gestor de emociones en condiciones de falla

Agente	Reputación
Cámara	0.011
RobotSocial	0.495
PizarraInteractiva	0.376

Tabla 8 – Reputación entre los agentes en condición de falla

Cod	Agentes	tA → (X,R,T)	Reputación	Felicidad	Tristeza	Ira	Sorpresa	Miedo
A3	Cámara	0.855	0.491	20.14	2.65	5.41	58.26	18.92
A6	RobotSocial	0.415	0.412	22.65	1.42	2.89	29.03	14.11
A7	PizarraInteractiva	0.271	0.268	2.45	0.61	1.17	0.87	2.15
			SPA	17.842	1.002	2.037	46.384	16.288

Tabla 9 – Emoción Social en condiciones de falla

En este caso, se constata que adecuando al contexto los modelos de confianza y reputación de los agentes en tiempo real, el mecanismo de determinación de la emoción social adecua la emoción social detectada. En la Tabla 9 se observa que la cámara no hace aportes para calcular la emoción social debido a que sus modelos tienen valor bajo, lo que es importante para tener fuentes de información robustas para calcular la emoción social.

4.3. Análisis de casos extremos

En condiciones normales, la emoción social se determina por los eventos capturados por los diferentes agentes dispositivos, a partir de los cuales reconocen un estado emocional. Por otra parte, al dejar de funcionar el agente cámara, como en el segundo escenario, se observa que su confianza y reputación disminuye de manera importante, tal que no es considera en el proceso. Veamos ahora el comportamiento del sistema en la misma situación de SaCI de la prueba 1 (sección 4.1) en tres contextos de confianza y reputación: 1) Todos los agentes se tienen mucha confianza, 2) Todos los agentes se tienen baja confianza y el modelo de reputación es el de la Tabla 7 (regular a baja reputación de los agentes), 3) Todos los agentes se tienen poca confianza y alta reputación. La emoción social que se obtienen en estas pruebas se muestra en la Tabla 10.

Los contextos muestran un aspecto interesante, dado que el modelo es capaz de dar una respuesta consistente todo el tiempo, muy parecida a la emoción social obtenida en la prueba 1. Lo que hace el modelo es determinar cuál fuente de información prevalece en

Contexto	Felicidad	Tristeza	Ira	Sorpresa	Miedo
1	18.41	1.56	3.15	44.72	15.06
2	12.61	1.68	1.38	33.58	11.21
3	19.43	1.15	2.32	43.59	16.81

Tabla 12 – Emoción Social para diferentes contextos de confianza y reputación

cada contexto, para calcula el valor final de la emoción social, según los valores en el modelo de confianza o de reputación de cada agente. Ahora, al ser algunos de los dos modelos sus valores altos (casos 1 y 3), los resultados se parecen bastante. Pero incluso, cuando ambos valores son bajos (caso 2), se estima una emoción social muy parecida, solo que con un mayor porcentaje de incertidumbre del tipo de emoción que no se logra determinar claramente (en el caso 2, se estima que en un 39,79% no se sabe cuál es la emoción social, a diferencia de caso 1 (17.1%) y caso 3 (16.7%), valores muy parecido al primer escenario de prueba). Estos escenarios muestran la robustez del modelo al determinar la emoción social, su versatilidad para manejar diferentes condiciones operacionales en SaCI, que pueden representar problemas diversos en el ambiente en un momento dado.

4.4.Aprendizaje

Una de las características más importante del modelo propuesto es la capacidad de adaptación del modelo de confianza. Esa capacidad reside en su modelo de aprendizaje, basado en los resultados que obtienen los estudiantes en el proceso de aprendizaje (es la base para calcular la variable gA, ver sección 2.3) y en sus experiencias pasadas en situaciones parecidas (determinada por la variable I, ver sección 2.3). El proceso de aprendizaje es por reforzamiento, tal que, si por ejemplo el valor de gA es alto, es debido en parte a que las emociones sociales detectadas fueron las adecuadas. En ese sentido, la confianza en los agentes que actuaron en la determinación de esas emociones sociales debe ser reforzada de manera importante, y para el resto evaporada. Esos mecanismos de aprendizaje se describen en las ecuaciones 4 y 5, los cuales afectan a la ecuación 1.

Para las pruebas en esta fase se supondrán una ganancia de gA de la forma de una función normal alrededor del valor de 0.9, y para I de la forma de una función normal alrededor del valor de 0.5, a través del tiempo, en SaCI, partiendo de la situación inicial del escenario 1, además, para $\epsilon\epsilon$ y a se supondrá 0.8. Solo se mostrará la evolución de la confianza para el rol y la tarea más similar al problema de reconocimiento (identificar). En el caso del agente RobotSocial se consideró identificar por expresiones fáciles (él puede reconocer de diferentes formas). La dinámica del proceso de aprendizaje se puede ver en la Figura 2, para las ecuaciones 1 y 5. En general, en algunos casos, en las primeras iteraciones hay algunos ajustes (caso agente cámara), pero después se estabilizan los valores de esas métricas, indicando que la confianza y confiabilidad en ese valor son relativamente estable en el resto del proceso.

Los valores de CA y rA se adaptan dependiendo de los valores de gA e I, según las ecuaciones 4 y 5. Ese proceso de adaptación por reforzamiento, depende de la calidad del proceso de aprendizaje en SaCI, medido a través de los resultados de los estudiantes

(gA) y experiencias pasadas (I). Si son buenas ambas métricas, los valores de CA y rA se refuerzan, de tal manera de mejorar la confianza en las capacidades de esos agentes para reconocer las emociones.

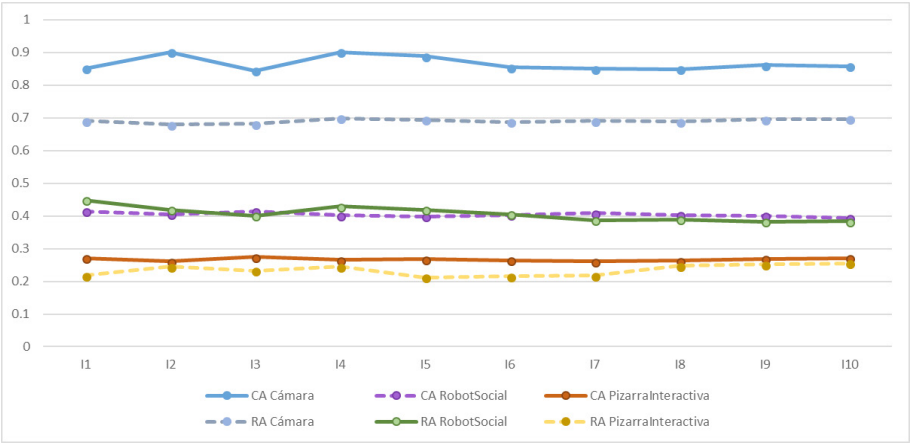


Figura 2 – Proceso de Aprendizaje

En concreto, como es la cámara en la que se tiene más confianza en ese escenario (ver sección 4.1), y gA es alto, sus valores de CA y rA son reforzados para siempre mantenerse altos, mientras que esos valores en los otros agentes se refuerzan poco para que mantengan sus niveles actuales. Eso se va haciendo a través de las iteraciones, de tal manera de mantener ajustado al modelo de confianza. En el caso del modelo de reputación, su ajuste es mucho más simple, ya que es una actualización sobre el valor P_i^j usando una expresión matemática como la de la ecuación 5.

5. Conclusiones

En este artículo se propone un enfoque de gestión de las emociones en un SaCI, basado en el uso de modelos de confianza y de reputación. El enfoque parte del principio de que múltiples agentes dispositivos en SaCI pueden percibir el estado emocional de los estudiantes presentes en SaCI en un momento dado. Partiendo de ese criterio, el agente gestor de emociones recibe los reconocimientos hechos por cada agente dispositivo, y los pondera según la credibilidad que él tenga (confianza), o se tenga en la comunidad (reputación), de cada agente dispositivo.

El modelo parte del concepto de emoción social, entendido por la necesidad de seguir una emoción que sea la más apta colectivamente, para los estudiantes presentes en un momento dado en SaCI. En ese sentido, la emoción social es la emoción que consensualmente es la mejor colectivamente. Eso le permite actuar a SaCI (en este caso, al agente VLE), siguiendo esa emoción que colectivamente es la más adecuada. Esta misma idea puede ser usada para gestionar las condiciones ambientes, entre otras cosas, en SaCI.

Por otro lado, los modelos de confianza y reputación pueden plasmar en ellos información sobre el comportamiento funcional de cada dispositivo basado en sus históricos de fallas, las dificultades presentes en el ambiente para que los agentes se comuniquen en un momento dado, entre otras cosas. De esta manera, esos modelos pueden irse actualizando para seguir la dinámica del sistema. Además, esos modelos son enriquecido por el efecto causado en el proceso de enseñanza-aprendizaje por las emociones reconocidas. En ese sentido, si los resultados del proceso de aprendizaje son exitosos en cuanto a que los estudiantes obtienen buenos resultados, es en parte porque la emoción social reconocida fue la adecuada. Por ello, los modelos de confianza y reputación son reforzados para aquellos agentes que coadyuvaron a identificar la emoción social.

Los dos modelos que se han usado, LIT y preferencia, han permitido tener buenos resultados en el comportamiento del esquema propuesto. El primer modelo, basado en los roles y tipos de tareas, permite gestionar la confianza de cada agente sobre los otros agentes. Además, tiene sus propios mecanismos de actualización en función del impacto de la emoción social detectada en el proceso de aprendizaje. El segundo modelo es basado en la fórmula de preferencia, pero ahora adaptado al contexto. Ambas formulas se combinan perfectamente, según la información disponible en el ambiente, para permitir calcular la emoción social.

En concreto, la innovación propuesta en este artículo es el modelo de emoción social para ambientes inteligentes basado en los modelos de confianza y de reputación de los SMA. Entre las características más resaltantes de ese modelo propuesto están: permite plasmar las características actuales en el contexto inteligente, es consistente en la emoción que detecta, y va aprendiendo los modelos de confianza y de reputación.

Algunos trabajos fututos tienen que ver con evaluar otros modelos de reputación y confianza, a ver si influyen en el proceso de cálculo de la emoción social. Otro trabajo futuro es llevar este modelo de consenso social de la emoción, a otras conversaciones de SaCI (por ejemplo, para establecer sus condiciones ambientales de temperatura, luminosidad, etc.). También, otro trabajo futuro es personalizar los procesos de aprendizaje en los agentes (modificar la ecuación 4), tal que el valor de confianza se actualizaría en función de cuanto participo cada agente en la obtención de la emoción social (valor de la confianza o reputación de ese agente en la ecuación 7). Finalmente, el proceso de adecuación del modelo de reputación debe ser realizado en próximos trabajos.

Referencias

Aguilar, J. (2014). *Introducción a los Sistemas Emergentes*. Talleres Gráficos, Universidad de Los Andes.

Aguilar, J., Camacho, O., & Chávez, D. (2017a). Social set points definition based on trajectory for control systems in smart classrooms. In *EEE 3rd Colombian Conference oAutomatic Control* (pp. 1–6).

Aguilar, J., Cordero, J., & Buendía, O. (2017b). Specification of the autonomic cycles of learning analytic tasks for a smart classroom. *Journal of Educational Computing Research*, *0*(0), 1–26.

Aguilar, J., Sánchez, M., Cordero, J., Valdiviezo-Díaz, P., Barba-Guamán, L., & Chamba-Eras, L. (2017c). Learning analytics tasks as services in smart classrooms. *Universal Access in the Information Society*.

Burnett, C., Norman, T. J., & Sycara, K. (2010). Bootstrapping trust evaluations through stereotypes. In *International Conference on Autonomous Agents and Multiagent Systems: volume 1-Volume 1* (pp. 241–248).

Cordero, J., & Aguilar, J. (2016). Reconocimiento multimodal de emociones en un entorno inteligente basado en crónicas. *COISINT*, 525–541

D'Angelo, G., Rampone, S., & Palmieri, F. (2017). Developing a trust model for pervasive computing based on Apriori association rules learning and Bayesian classification. *Soft Computing*, *21*(21), 6297–6315.

Hang, C.-W., Wang, Y., & Singh, M. P. (2009). Operators for propagating trust and their evaluation in social networks. In *International Conference on Autonomous Agents and Multiagent Systems-Volume 2* (pp. 1025–1032).

Hermoso, R., Billhardt, H., & Ossowski, S. (2013). Trust-based role coordination in task-oriented multiagent systems. *Knowledge-Based Systems*, *52*, 78–90.

Lijian, H., Houkuang, H., & Wei, Z. (2010). A framework to support heterogeneous reputation models interoperation in Multi-Agent System. In *2nd International Conference on Advanced Computer Control* (pp. 183–188).

Nguyen, M. H., & Tran, D. Q. (2013). A combination trust model for multi-agent systems. *International Journal of Innovative Computing, Information and Control*, *9*(6), 2405–2420.

Nguyen, M. H., & Tran, D. Q. (2015a). A trust-based mechanism for avoiding liars in referring of reputation in multiagent system. *Editorial Preface*, *4*(2).

Nguyen, M. H., & Tran, D. Q. (2015b). A trust model for new member in multiagent system. *Vietnam Journal of Computer Science*, *2*(3), 181–190.

Ramchurn, S. D., Huynh, D., & Jennings, N. R. (2004). Trust in multi-agent systems. *The Knowledge Engineering Review*, *19*(1), 1–25.

Sánchez, M., Aguilar, J., Cordero, J., Valdiviezo-Díaz, P., Barba-Guamán, L., & Chamba-Eras, L. (2016). *Cloud computing in smart educational environments: Application in learning analytics as service*. *Advances in Intelligent Systems and Computing* (Vol. 444).

Sánchez, M., Aguilar, J., Cordero, J., & Valdiviezo, P. (2015a). A Smart Learning Environment based on Cloud Learning. *International Journal of Advanced Information Science and Technology*, *39*(39), 39–52.

Sánchez, M., Aguilar, J., Cordero, J., & Valdiviezo, P. (2015b). Basic Features of a Reflective Middleware for Intelligent Learning Environment in the Cloud (IECL). In *Proceedings - 2015 Asia-Pacific Conference on Computer-Aided System Engineering, APCASE 2015*. https://doi.org/10.1109/APCASE.2015.8

Valdiviezo, P., Aguilar, J., Cordero, J., & Sánchez, M. (2015). Conceptual Design of a Smart Classroom Based on Multiagent Systems. In *International Conference on Artificial Intelligence (ICAI)* (p. 471).

Victor, P., Cornelis, C., Cock, M., & Da Silva, P. (2009). Gradual trust and distrust in recommender systems. *Fuzzy Sets and Systems*, *160*(10), 1367–1382.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 22/09/2018
Aceitação/Acceptance: 11/11/2018

Criptoanálisis del algoritmo RSA de al menos 330 bits con dispositivos de bajo costo SBC

Nelsón Darío Pantoja[1], Anderson Felipe Jiménez[2], Siler Amador Donado[3], Katerine Márceles Villalba[4]

Dariopantoja@unimayor.edu.co[1], afjimenez@unimayor.edu.co[2], samador@unicauca.edu.co[3], kmarceles@unimayor.edu.co[4]

[1, 2, 4] Facultad de Ingeniería, Institución Universitaria Colegio Mayor del Cauca, Carrera 7 N° 2-34 Edificio Bicentenario, 190003, Popayán- Cauca, Colombia.

[3] Facultad de Ingeniería Electrónica y Telecomunicaciones, Universidad del Cauca, Carrera 2 #4N-140, 190003, Popayán- Cauca, Colombia

Pages: 886–899

Resumen: Actualmente los dispositivos Single Board Computer (SBC), han sido utilizados en muchas ramas de la ciencia, pero no especificamente en el campo del criptoanálisis, por lo que supone como un reto afrontar, tomando como relevancia componentes como el hardware y el software para realizar distribuciones y llevar a cabo la técnica de criptoanálisis, dado a lo anterior éste artículo busca demostrar ¿Qué tan eficiente en tiempo resulta una técnica de criptoanálisis sobre un SBC que resulte adecuado para quebrantar llaves del algoritmo RSA de al menos 330 bits?, para esto se realizaron pruebas de rendimiento del algoritmo frente a un solo dispositivo SBC confrontando frente a un Clúster comprendido por 3 SBC, mediante la distribución de la ejecución del algoritmo y la medición en tiempo que consume en vulnerar las llaves generadas para el Data set.

Palabras-clave: Bit, Clúster, Criptoanálisis, RSA, SBC.

Cryptanalysis of the RSA algorithm of at least 330 bits with low cost devices SBC

Abstract: Currently Single Board Computer (SBC) devices have been used in many branches of science, but not specifically in the field of cryptanalysis, so it is a challenge to address, taking as relevant components such as hardware and software to perform distributions and carry out the technique of cryptanalysis, given the foregoing this article seeks to demonstrate how efficient in time is a technique of cryptanalysis on a SBC that is adequate to break keys of the RSA algorithm of at least 330 bits ?, for this performance tests of the algorithm were performed against a single SBC device confronting a Cluster comprised of 3 SBC, by means of the distribution of the execution of the algorithm and the measurement in time consuming in violating the keys generated for the Data set.

Keywords*: Bit, Cluster, Cryptanalysis, RSA, SBC.*

1. Introducción

El algoritmo RSA (Rivest, Shamir y Adleman) es uno de los más seguros y esto se debe al empleo de una clave de cifrado numérica bastante grande, estas claves se generan al multiplicar dos números primos cuyo resultado deriva en claves que sobrepasan los 300 bits o 100 dígitos escogidos al azar (Bressoud, 1989). Ya se han establecido medidas de tiempo y eficiencia de criptoanálisis sobre este sistema criptográfico, cuya solidez radica en la dificultad de factorizar números especialmente grandes, estas medidas se han obtenido gracias a pruebas realizadas en sistemas de cómputo tradicionales (Yan, 2008). Dado a lo anterior surge la pregunta ¿ Qué tan eficiente en tiempo resulta una técnica de criptoanálisis sobre un SBC (Single Board Computer) adecuado para vulnerar llaves RSA de al menos 330 bits?.

En la actualidad estos dispositivos (SBC), han sido utilizados en muchas ramas de la ciencia, pero no parece haber muchas referencias sobre los mismos en el camino del criptoanálisis, debido a que los antecedentes apuntan claramente a que la distribución del procesamiento para el criptoanálisis del algoritmo RSA mediante factorización es la más eficiente (Kleinjung et al., 2010), este elemento también supone un reto a afrontar con los dispositivos anteriormente mencionados, pues toman relevancia componentes como el hardware y el software adicionales necesarios para realizar este tipo de distribuciones sin mencionar los elementos que se precisan para llevar a cabo la técnica de criptoanálisis seleccionada, por ello durante el desarrollo del artículo se mostrará un estudio de rendimiento del algoritmo con respecto a la vulnerabilidad de un determinado número de llaves bajo el algoritmo RSA, demostrando que mediante la construcción de un clúster con computadores de tarjeta simple y de bajo costo, se puede realizar un criptoanálisis al sistema criptográfico mencionado.

2. Trabajos relacionados al criptoanálisis

En esta sección se enmarcan el conjunto de trabajos previos, que se tomó como referentes para el proceso de investigación:

En el artículo Raspberry Pi Computer Cluster (Petit, Johnson, Vo, Vo, & Grant, 2010), los autores identifican que en la distribución de procesamiento con SBC's se evidencian una serie de beneficios y limitaciones, por un lado, está la sobrecarga de comunicación, lo cual es entendible y esperable; también se logró observar una mejora notable al comparar el rendimiento de un solo dispositivo y un cluster de los mismos, pero como es evidente nunca supera la relación entre el tiempo y la cantidad de dispositivos.

En el paper Factorization of a 768-Bit RSA Modulus(Kleinjung et al., 2010), se explica el desarrollo del proceso para lograr el objetivo de romper el número de bits de RSA más grande conocido hasta la fecha, además menciona un poco las metas alcanzadas anteriormente, este proyecto fue desarrollado bajo la propuesta del reto RSA propuesto en 1991 para incentivar el ataque al algoritmo con el fin de mejorar su seguridad.

Pointcheival, David plantea en el artículo How to Encrypt Properly with RSA (Pointcheval, 2002) que inicialmente los esfuerzos para realizar un criptoanálisis al algoritmo RSA fueron realizados en 1994 cuando se intentó exitosamente romper una clave pública de 129 dígitos usando 1600 computadores en red, posteriormente en 1999 se quebrantó

una llave de 140 dígitos empleando 185 computadores en red en aproximadamente 8,9 años; en este mismo año con 300 computadores en red se logró romper una llave de 512 bits, y a partir de 2002 se recomienda el uso de llaves de 1024 bits, a partir de lo mencionado en el artículo se puede medir el alcance del proyecto a realizar, reduciendo el número de bits de las llaves a 340 teniendo en cuenta los recursos computacionales de los SBC's.

El Matemático Hugo Scolnik es uno de los principales líderes en la investigación de factorización entera, necesaria para descifrar el sistema criptográfico RSA. Scolnik asegura no necesitar ordenadores cuánticos ni el algoritmo de Peter Shor para romper las claves del mencionado sistema, es más en uno de sus documentos titulado "Fundamentos matemáticos del método RSA" menciona que, si la cantidad de bits es 512, los años para factorizar serían 30.000 suponiendo que se ejecuten 1.000.000 de instrucciones por segundo en una máquina de cómputo (Scolnik, 2004).

En el artículo "The Quadratic Sieve Factoring Algorithm" el reconocido profesor Carl Pomerance explica los fundamentos matemáticos de la técnica de factorización de Criba Cuadrática, la cual se ha demostrado que hasta ahora es la mejor técnica de factorización para números conformados por hasta 110 dígitos (Pomerance, 1985), con lo que dicha técnica se establece como la mejor opción para el presente proyecto, por lo menos hasta los 90 dígitos, pues de ahí en adelante se utiliza NFS como se puede apreciar más adelante.

Pomerance, Carl en 1994 publica un artículo llamado "The Number Field Sieve" (NFS), donde explica el desarrollo y el proceso matemático por el cual hace posible factorizar números conformados por hasta 155 dígitos, correspondientes a 512 bits, siendo este trabajo sucesor del Quadratic Sieve. El NFS ha demostrado ser mucho más eficiente a la hora de factorizar desde las 90 cifras decimales, con lo que se decide añadirlo al proceso de criptoanálisis del presente proyecto.

3. Implementación del Clúster con SBC

Para automatizar la técnica de criptoanálisis escogida NFS al algoritmo RSA con 330 bits, se decidió utilizar dispositivos SBC, debido a los beneficios que otorga su arquitectura modular y adaptable(Pajankar, 2015), además de su reducido costo permite ajustar la potencia óptima a medida de los requerimientos de la técnica previamente mencionada. Aunque los SBC's son los dispositivos seleccionados para implementar la técnica de criptoanálisis, hay algunas desventajas a tener en cuenta con la utilización de los mismos, por ejemplo la fragilidad de sus componentes eléctricos, un detalle bastante a tener en cuenta sobre todo por las condiciones eléctricas de la ciudad en la que se desarrolla el proyecto (Irwin, 1997), también, como se mencionaba anteriormente su costo es muy bajo lo que resulta ventajoso, pero por contraparte se debe lidiar con su limitada potencia requiriendo la utilización simultánea de varios de estos dispositivos.

Teniendo en cuenta los antecedentes en proyectos similares se decidió construir un cluster con los dispositivos mencionados anteriormente, dado a que está comprobado que distribuir el proceso de factorización es mucho más eficiente que hacerlo en un solo

dispositivo independiente a sus características (Dean & Ghemawat, 2008). Para ello se utilizan Raspberry Pi 3 Model B pues el acceso que proporciona a la hora de montar un cluster es una de las razones para escoger este dispositivo (Petit et al., 2010), además de contar con una de las mayores comunidades que le brindan apoyo con gran cantidad de proyectos, también su sistema operativo principal "Raspbian" aporta soluciones eficientes para hacer la distribución de los procesos, pues dada su naturaleza basada en Debian hace posible el uso de SSH, el cual es el soporte para la implementación de operaciones colectivas (Khan, Mohit, & Hina, 2016), ideal para montar la técnica de factorización en el cluster.

4. Experimentos y pruebas

Para el proyecto se necesitó construir un conjunto de llaves RSA que fueron sometidas a un criptoanálisis. Este conjunto de llaves fue construido mediante una librería criptográfica de código abierto denominada OpenSSL.

OpenSSL permite generar llaves RSA académicas para ser analizadas, dado a su naturaleza de código abierto permitió la creación de las llaves necesarias para formar un conjunto de llaves lo suficientemente amplio para realizar un Data set de pruebas.El Data Set se generó haciendo uso también del software para Windows de Miguel Schlereth Martínez de la Universidad Politécnica de Madrid – España llamado RSAManager. Este software al ser ligado con la librería de OpenSSL permitió construir un Data Set con la cantidad de llaves y número de bits deseados.

La evaluación de estas llaves consiste en tomar el módulo n de la llave pública y enviarlo como parámetro de factorización al software Cado-nfs, el cual genera un log de registro donde se encontrarán los números primos p y q para la construcción de la llave privada, esto demuestra que la llave fue rota y no es segura de usar.

Se obtienen los resultados provenientes del software Cado-nfs, este mismo software emplea el algoritmo NFS para optimizar el proceso de factorización y encontrar los factores p y q.

En términos generales, al realizar la configuración adecuada del software junto con las librerías necesarias, es posible ingresar parámetros de números enteros para que Cado-nfs los factorice. Una vez factorizados, se verifica que el número entero ingresado, para este caso en particular el módulo n de las llaves públicas generadas con OpenSSL, tenga solo dos factores, ya que el módulo de las llaves se crea únicamente con la multiplicación de dos factores (p y q), de ser así, se considera que la llave fue rota o quebrantada.

Se ha dispuesto de un Data-set de llaves a factorizar las cuales fueron sometidas a evaluación. Los resultados obtenidos fueron satisfactorios, el *Data-set* que fue puesto a prueba haciendo uso del algoritmo *NFS* con la implementación del software *Cado-nfs* no resultó victorioso con ninguna de sus llaves, todas se lograron factorizar con éxito. A continuación, se muestran las ilustraciones con sus respectivas llaves

y tiempos de factorización de acuerdo a los bits de cada parámetro de entrada empleando un solo SBC:

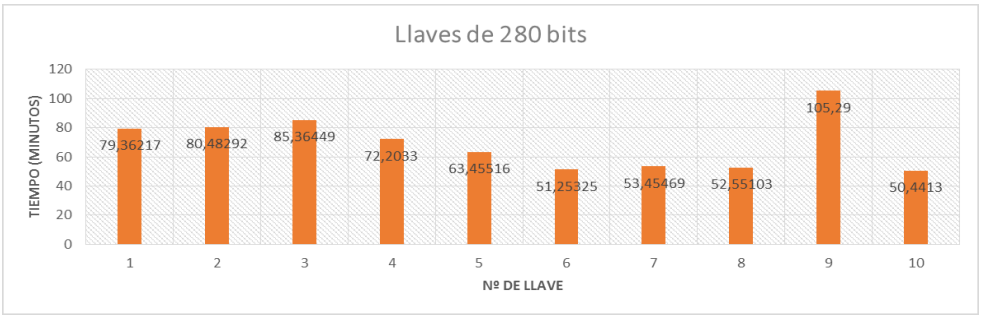


Figura 1 – Llaves de 280 bits quebrantadas con un Raspberry pi 3

En la representación gráfica de las llaves de 280 bits que fueron evaluadas se observa como fueron quebrantadas en su totalidad, disponiendo de un tiempo máximo para la llave número 9 de aproximadamente 105 minutos y un tiempo mínimo para la llave número 10 de aproximadamente 50 minutos.

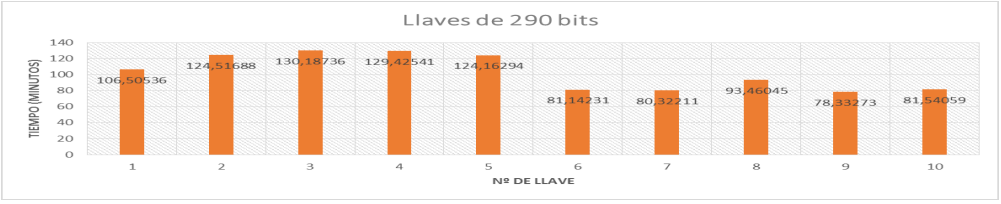


Figura 2 – Llaves de 290 bits quebrantadas con un Raspberry pi 3

Al efectuar el criptoanálisis sobre el segundo grupo de llaves con 290 bits se observa que al igual que en el anterior grupo, estas llaves fueron rotas en su totalidad, dando un tiempo máximo para la llave número 3 de aproximadamente 130 minutos y un tiempo mínimo para la llave número 9 de aproximadamente 78 minutos.

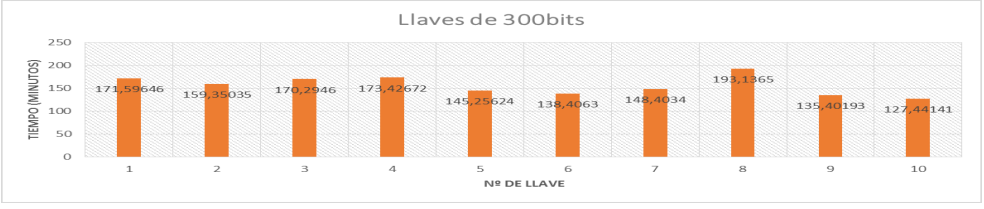


Figura 3 – Llaves de 300 bits quebrantadas con un Raspberry pi 3

En el grupo de llaves de 300 bits se observa que también fueron quebrantadas las llaves en su totalidad, otorgando un tiempo máximo de rompimiento en la llave número

8 con aproximadamente 193 minutos y un tiempo mínimo en la llave número 10 de aproximadamente 127 minutos.

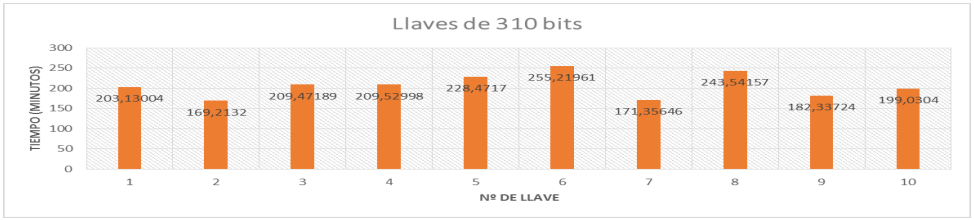


Figura 4 – Llaves de 310 bits quebrantadas con un Raspberry pi 3

En el grupo de llaves de 310 bits se evidencia que fueron rotas en su totalidad las llaves de este grupo evaluado, otorgando tiempo máximo de rompimiento en la llave número 6 de aproximadamente 255 minutos y tiempo mínimo en la llave número 2 de aproximadamente 169 minutos.

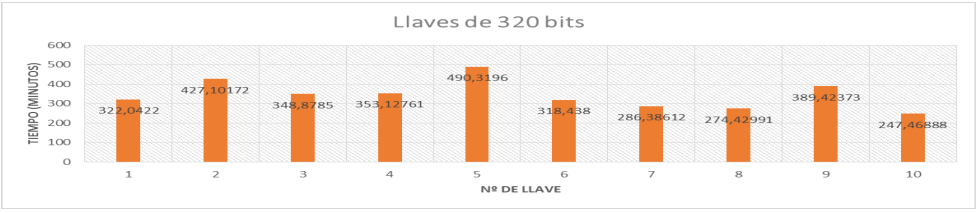


Figura 5 – Llaves de 320 bits quebrantadas con un Raspberry pi 3

De igual forma para las llaves de 320 bits, se evidenció que todas las llaves fueron quebrantadas, otorgando un tiempo máximo de rompimiento para la llave número 5 de aproximadamente 490 minutos y un tiempo mínimo para la llave número 10 de aproximadamente 247 minutos.

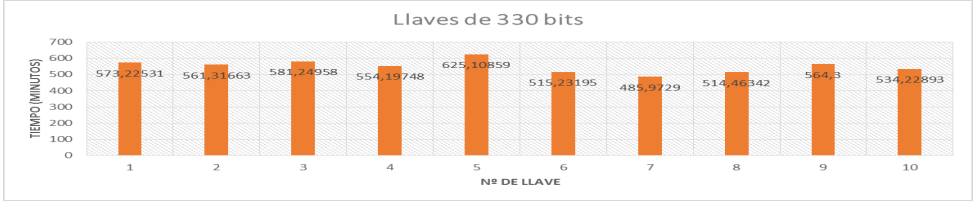


Figura 6 – Llaves de 330 bits quebrantadas con un Raspberry pi 3

Cuando se realiza el criptoanálisis al grupo de llaves más cercano al objetivo propuesto, los tiempos son considerablemente altos, se evidencia un tiempo máximo para la llave número 5 de aproximadamente 625 minutos (Aprox. 11 horas) y un tiempo mínimo para la llave número 7 de aproximadamente 486 minutos (Aprox. 8 horas).

Desde la figura 1 a la figura 6, se puede observar el comportamiento del criptoanálisis con respecto al tiempo de factorización para obtener los factores *p* y *q* del módulo *n* de las correspondientes llaves.

Para esta etapa de la investigación se concluye que el tiempo de la factorización de la llave objetivo es demasiado extenso, como lo muestra la figura 6, lo cual implica mucho trabajo computacional continuo por parte de un solo Raspberry pi 3.

Para solucionar los problemas mencionados se acuerda que es conveniente utilizar procesamiento distribuido, el cual en teoría podría mejorar el tiempo de factorización y además ayudaría a la confirmación de las llaves que ya han sido factorizadas, brindando una segunda verificación de las mismas.

En ese mismo orden de ideas el proceso anterior se procedio a realizarlo sobre el cluster de SBC, como se muetra en la figura 7, con el ánimo de evaluar la eficiencia en tiempo con respecto al quebrantamiento de número de llaves .

Figura 7 – Cluster *Físicamente Construido.*

El primer paso en este proceso como se puede intuir es la generación de llaves, la cual para este proyecto fueron distribuidas por el software Open SSL, cabe resaltar que esta herramienta está disponible en muchos sistemas operativos.

Una vez construido el sistema distribuido por medio de un cluster de Raspberry pi 3 se someten las llaves factorizadas que se encuentran en el Data-set. Esto arrojó los siguientes resultados en términos de tiempo:

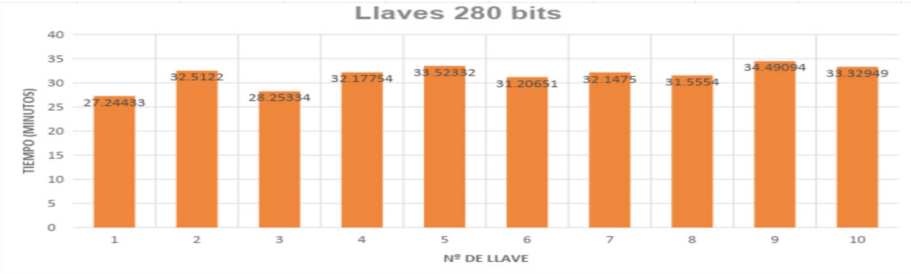


Figura 8 – Llaves de 280 bits quebrantadas con un Clúster de 3 nodos Raspberry pi 3

Al realizar el criptoanálisis con el clúster de 3 nodos a las llaves de 280 bits se obtuvo un tiempo de factorización máximo para la llave número 9 de aproximadamente

34 minutos y un tiempo mínimo para la llave número 1 de aproximadamente 27 minutos.

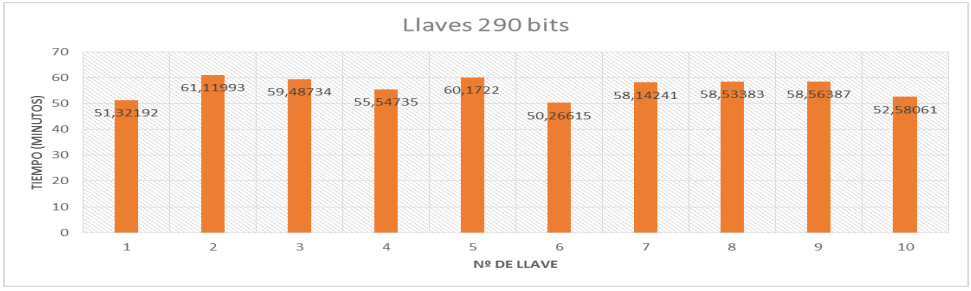


Figura 9 – Llaves de 290 bits quebrantadas con un Clúster de 3 nodos de Raspberry pi 3

En el grupo de llaves de 290 bits se encontró un tiempo máximo de factorización para la llave número 2 de aproximadamente 61 minutos y un tiempo mínimo para llave número 6 de aproximadamente 50 minutos.

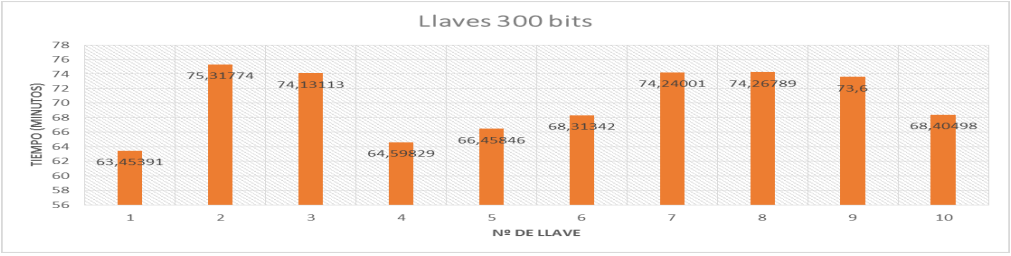


Figura 10 – Llaves de 300 bits quebrantadas con un Clúster de 3 nodos de Raspberry pi 3

Con el grupo de llaves de 300 bits se obtuvo un tiempo máximo de factorización para la llave número 2 de aproximadamente 75 minutos y un tiempo mínimo para la llave número 1 de aproximadamente 63 minutos.

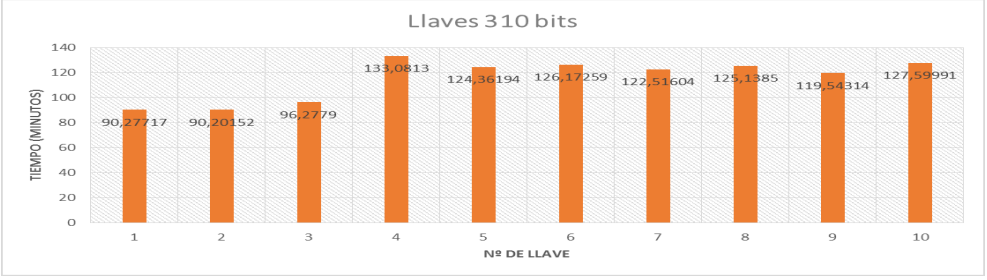


Figura 11 – Llaves de 310 bits quebrantadas con un Clúster de 3 nodos de Raspberry pi 3

Con el grupo de llaves de 310 bits, un tiempo máximo de factorización para la llave número 4 de aproximadamente 133 minutos y un tiempo mínimo para la llave número 2 de aproximadamente 90 minutos.

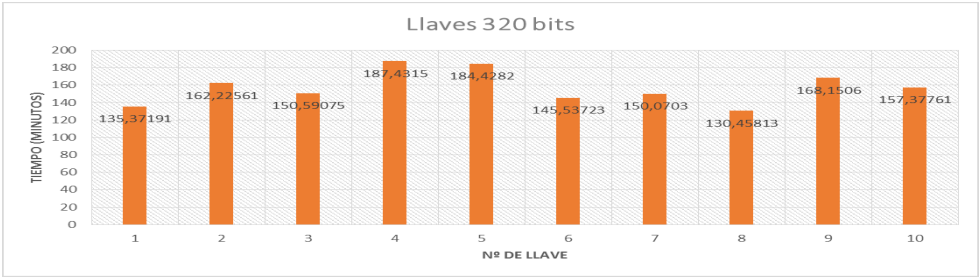


Figura 12 – Llaves de 320 bits quebrantadas con un Clúster de 3 nodos Raspberry pi 3

Con llaves de 320 bits se observa un tiempo máximo de factorización para la llave número 4 de aproximadamente 187 minutos y un tiempo mínimo para la llave número 8 de aproximadamente 130 minutos.

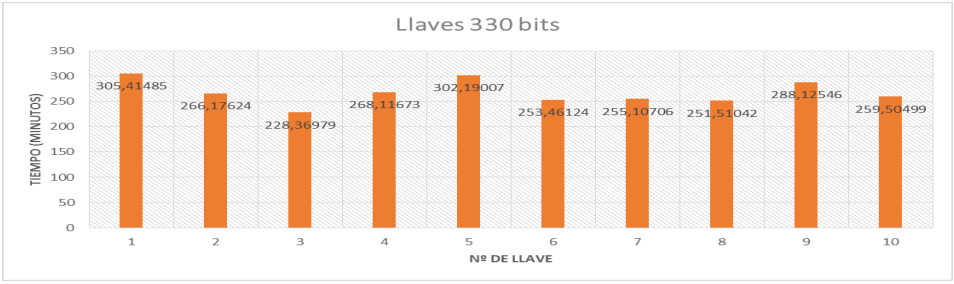


Figura 13 – Llaves de 330 bits quebrantadas con un Clúster de 3 nodos Raspberry pi 3

De igual forma para el grupo de llaves de 330 bits donde todas fueron quebrantadas satisfactoriamente, se obtuvo un tiempo máximo de factorización en la llave número 1 de aproximadamente 305 minutos y un tiempo mínimo en la llave número 3 de 228 minutos.

Como se observa en las ilustraciones donde se distribuyen las llaves en un cluster de raspberry's, el tiempo de factorización es mucho menor que cuando se trabaja computacionalmente con un solo raspberry, esto quiere decir que la teoría implementada del sistema distribuido fue acertada y se logró realizar un criptoanálisis exitoso a llaves de 330 bits generadas por OpenSSL.

Para evidenciar con claridad los tiempos de factorización del cluster versus un solo raspberry pi 3 se generó un comparativo que ilustra el contraste que tiene el comportamiento de una distribución con 3 nodos frente a un sistema unitario con respecto al tiempo empleado para la realización del criptoanálisis, a continuación las ilustraciones correspondientes:

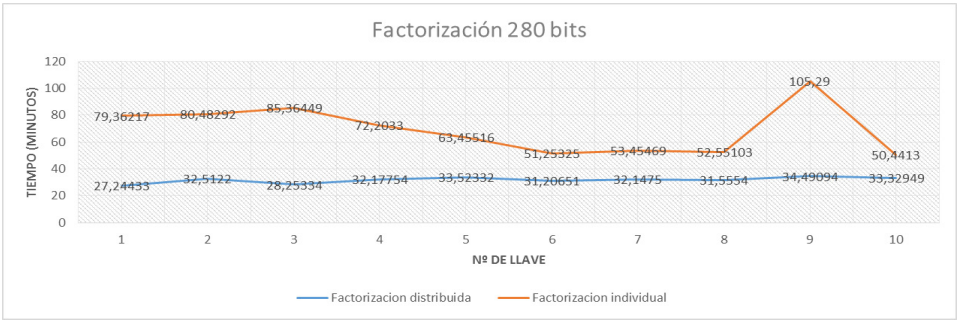


Figura 14 – Factorización distribuida vs individual de 280 bits

Es fácil notar que el clúster de raspberry realiza de forma satisfactoria la factorización de las llaves 280 bits, además se observa que el tiempo efectuado es mucho menor que el de un solo rapsberry pi 3.

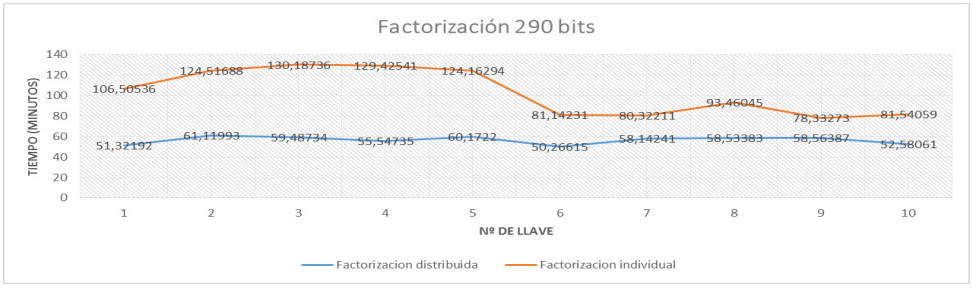


Figura 15 – Factorización distribuida vs individual de 290 bits

La factorización de llaves de 290 bits haciendo uso del clúster resulta eficiente y se nota que los tiempos reducen ampliamente, al igual que el anterior grupo se verifican las llaves una segunda vez.

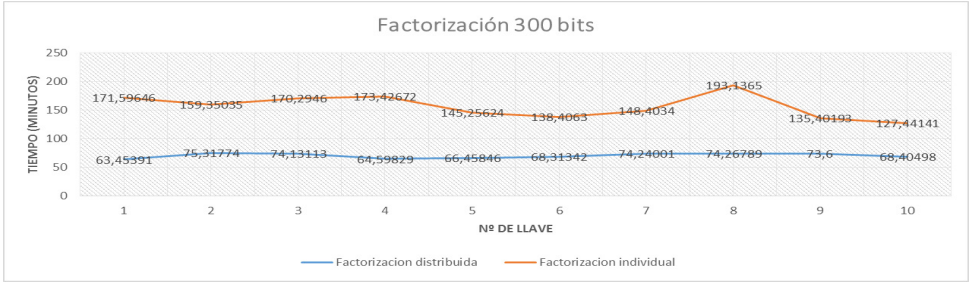


Figura 16 – Factorización distribuida vs individual de 310 bits

En el grupo de llaves de 300 bits el clúster verifica por segunda vez las llaves dando resultados positivos para proceder al siguiente grupo.

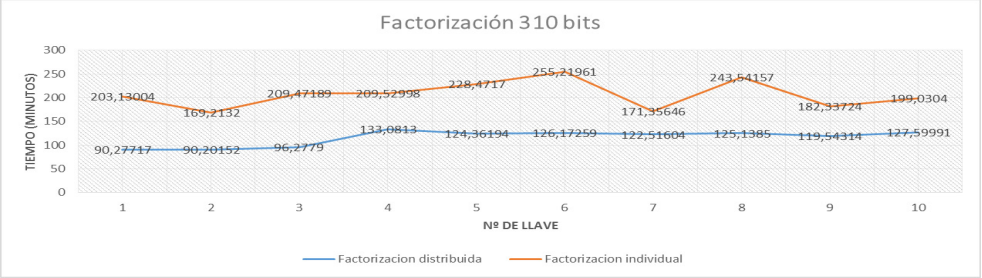


Figura 17 – Factorización distribuida vs individual de 310 bits

Con las llaves de 310 bits se aprecia como el tiempo de factorización en la distribución no depende de la cantidad de nodos, como por ejemplo en la llave número 2, con un raspberry su tiempo fue de 169 minutos aproximadamente y en sistema distribuido de 90 minutos aproximadamente, mientras que la llave número 6 con un solo raspberry su tiempo fue de aproximadamente 255 minutos y distribuida fue de aproximadamente 126 minutos, los tiempos distribuidos de las llaves 2 y 6 son muy diferentes y no significa que si en un sistema de un solo raspberry la llave con tiempo máximo de factorización fue la numero 6, en un sistema distribuido esta misma llave será la de menor tiempo ya que la llave con máximo tiempo en el sistema distribuido no es la número 6 si no la número 4.

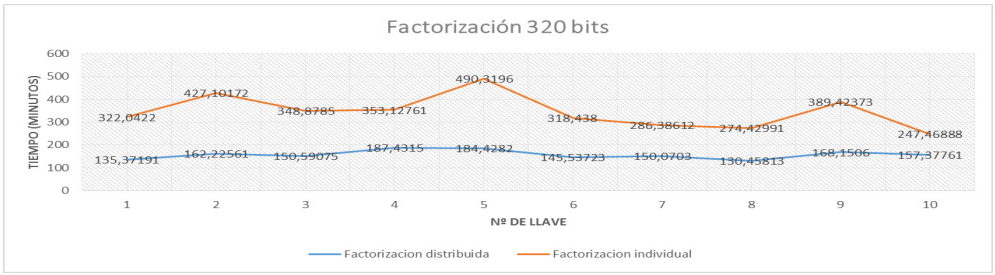


Figura 18 – Factorización distribuida vs individual de 320 bits

Al igual que los anteriores grupos el clúster optimiza de manera eficiente los tiempos de factorización de las llaves de 320 bits y brinda una segunda verificación de que han sido quebrantadas exitosamente.

De la misma forma el clúster realiza su tarea y factoriza exitosamente las llaves de 330 bits, dando una segunda revisión.

A continuación se detallan los tiempos empleados para la ejecución del Data set sobre el clúster de 3 raspberry pi 3:

- Pruebas de 280 bits: Aproximadamente 277 minutos (Aprox. 5 horas)
- Pruebas de 290 bits: Aproximadamente 562 minutos (Aprox. 9 horas)
- Pruebas de 300 bits: Aproximadamente 699 minutos (Aprox. 12 horas)
- Pruebas de 310 bits: Aproximadamente 1152 minutos (Aprox. 20 horas)

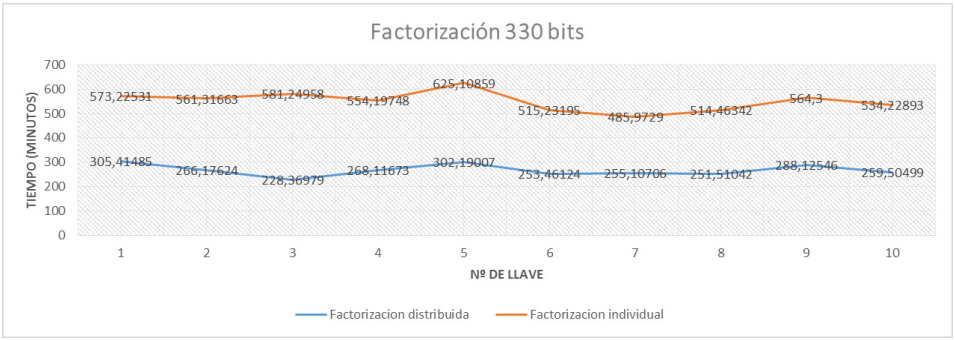


Figura 19 – Factorización distribuida vs individual de 330 bits

- Pruebas de 320 bits: Aproximadamente 1568 minutos (Aprox. 26 horas)
- Pruebas de 330 bits: Aproximadamente 2675 minutos (Aprox. 45 horas)
- Total de minutos de pruebas para el clúster: 6933 (117 horas).

Lo anterior, evidencia como las llaves han sido confirmadas con un nuevo criptoanálisis dentro de un sistema distribuido, y además se muestra como es notoriamente eficiente el clúster para optimizar el tiempo de factorización de las llaves evaluadas

Los resultados presentados demuestran que el objetivo principal del proyecto fue concretado de manera óptima, evidenciando que es posible dar respuesta a la pregunta problema, ¿Qué tan eficiente en tiempo resulta una técnica de criptoanálisis sobre un SBC (Single Board Computer) adecuado para vulnerar llaves RSA de al menos 330 bits?

5. Conclusiones y trabajos futuros

De acuerdo a las pruebas iniciales de configuración del software con el algoritmo NFS para factorizar, se dedujo que la arquitectura del dispositivo es notoriamente influyente para el desempeño del software, pero no impide que se lleve a cabo, todo esto se debe a que la arquitectura de los SBC's Raspberry pi 3 Modelo B es RISC (Reduced Instruction Set Computer), el cual contaba con un procesador ARM Cortex A9, y el software y las librerías necesarias para la ejecución de Cado-nfs debían corresponder a la arquitectura tratada, de haber seleccionado librerías que correspondan a una arquitectura de procesador *x86*, por ejemplo, Cado-nfs no lograba realizar un funcionamiento correcto para la factorización.

Es posible argumentar que sí fue posible optimizar los tiempos de factorización con un clúster de Raspberry pi 3 con arquitectura RISC, también es posible emplear un clúster con ordenadores que posean mejores características hardware, y potenciar aún más los tiempos de factorización de las llaves.

Igualmente se concluye de manera afirmativa la pregunta de investigación previamente planteada, dando por hecho la factibilidad de ejecutar un criptoanálisis a una llave de al menos 330 bits, mediante el uso de un SBC, permitiendo así aprovechar bondades como la de su precio reducido.

Se evidencia el crecimiento en orden exponencial en factores de tiempo al realizar un criptoanálisis a llaves cuyos bits vayan incrementando gradualmente, es decir a medida que las llaves crecen en bits el tiempo de criptoanálisis ira creciendo notoriamente.

Como se observó en las pruebas realizadas los clúster son mucho más eficientes a la hora de tratar problemas complejos y optimizar tiempos, estas pruebas garantizan que abordar problemas matemáticos como el cifrado RSA son preferiblemente tratados con múltiples recursos computacionales o mediante clústeres.

Como trabajos futuros, el primero de ellos es el aumentar la cantidad de nodos al clúster de Raspberry para verificar en qué cantidad de nodos se iguala a un computador convencional, una vez hecho esto verificar si el costo-beneficio del clúster es mejor que el costo-beneficio de un computador convencional.Como segundo trabajo futuro es potenciar el tiempo de criptoanálisis por medio de un clúster de computadores convencionales, incluso el verificar el comportamiento en un servidor o clúster de servidores.El tercer trabajo futuro es ejecutar el mismo proceso de ataque al algoritmo RSA haciendo uso de tarjetas gráficas para abordar el problema y realizar un comparativo de rendimiento de factorización entre CPU y GPU.

Agradecimientos

Agradecimientos a la Universidad del Cauca en especial a su grupo de investigación GTI y al semillero Beta Bit de la Facultad de Ingeniería de la Institución Universitaria Colegio Mayor del Cauca, por el apoyo brindado para el desarrollo del proyecto.

Referencias

Bressoud, D. (1989). Factorization and Primality Testing (1st ed.). New York, NY, USA: Springer-Verlag New York. https://doi.org/10.1007/978-1-4612-4544-5

Dean, J., & Ghemawat, S. (2008). MapReduce: Simplified Data Processing on Large Clusters. Commun. ACM, 51(1), 107–113. https://doi.org/10.1145/1327452.1327492

Erickson, J. (2008). Hacking: The Art of Explotation. San Francisco: Starch Press.

Irwin, D. (1997). The Industrial Electronic Handbook. Boca Raton: CRC Press LLC.

Kleinjung, T., Aoki, K., Franke, J., Lenstra, A. K., Thomé, E., Bos, J. W., … Zimmermann, P. (2010). Factorization of a 768-Bit RSA Modulus. In T. Rabin (Ed.), Advances in Cryptology -- CRYPTO 2010: 30th Annual Cryptology Conference, Santa Barbara, CA, USA, August 15-19, 2010. Proceedings (pp. 333–350). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-14623-7_18

Petit, N. J., Johnson, K., Vo, P., Vo, D., & Grant, C. (2010). Raspberry Pi Computer Cluster. Computer Science, 1–7. Retrieved from http://www.micsymposium.org/mics2015/ProceedingsMICS_2015/Johnson_2A2_40.pdf

Pomerance, C. (1985). The Quadratic Sieve Factoring Algorithm. In T. Beth, N. Cot, & I.

Pointcheval, D. (2002). How to Encrypt Properly with RSA. RSA Laboratories' CryptoBytes, 5(1), 1–10.

Scolnik, H. (1 de Abril de 2004). Departamento de Computación -Facultad de Ciencias Exactas y Naturales - Materias - Documentos - Universidad de Buenos Aires. Obtenido de Departamento de Computación - Universidad de Buenos Aires: http://www-2.dc.uba.ar/materias/crip/docs/rsamath01.pdf

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 12/09/2018
Aceitação/Acceptance: 27/11/2018

Modelo de tópicos para repositorios Git: Un enfoque basado en la estructura de directorios de código fuente

Paul Mendoza del Carpio[1]

p.mendozadc@ulasalle.edu.pe

[1] Universidad La Salle, Departamento de Ingeniería y Matemáticas, Av. Alfonso Ugarte 517 Arequipa, Perú.

Pages: 900–912

Resumen: Las populares plataformas Git albergan proyectos de software de gran escala, conteniendo grandes volúmenes de código fuente que son difíciles de entender en tareas de mantenimiento. El entendimiento de código fuente es influenciado por el vocabulario utilizado para nombrar entidades como directorios y archivos de código. Los modelos de tópicos ayudan a los usuarios a entender grandes colecciones de documentos cuando hay demasiados documentos. Este trabajo presenta un enfoque para la extracción de tópicos en repositorios Git de código fuente, el cual emplea las regularidades frecuentes respecto al nombrado y la estructura de directorios de código fuente sin requerir leer o analizar archivos de código. Se analizaron populares proyectos de código abierto presentes en GitHub, los resultados han sido evaluados atendiendo la coherencia de los tópicos.

Palabras-clave: Modelo de tópicos; repositorio Git.

Topic model for Git repositories: An approach based on source code directory structure

Abstract: Popular Git platforms host large scale software projects, containing huge volumes of source code which are difficult to understand during maintenance tasks. The source code understanding is affected by the vocabulary used for naming entities such as directories and source code files. The topic models support users to understand large collections of documents when there is too much documents. This work presents an approach for extracting topics in Git repositories of source code, using the frequent regularities in respect of the naming and structure of source code directories without necessitate reading or parsing source files. Popular open source software projects in GitHub were analyzed, the results has been evaluated attending the topic coherence.

Keywords*:* Topic model; Git repository.

1. Introducción

Las plataformas Git como GitHub, GitLab, Bitbucket albergan proyectos de software de gran escala, conteniendo grandes volúmenes de código fuente. En el marco de tareas de

mantenimiento y evolución de software, los desarrolladores opinan que generalmente es insuficiente la documentación, lo cual produce que muchas veces ellos se enfoquen sólo en el código fuente (Avidan & Feitelson, 2017), sin embargo su labor de examinación del código puede requerir mucho esfuerzo al abordar repositorios Git con considerable cantidad de archivos de código fuente.

El entendimiento de código fuente es influenciado por el vocabulario utilizado para nombrar entidades como directorios (i.e.; paquetes), archivos de código (e.g.; clases), y bloques de construcción (e.g.; métodos), debido a que estos nombres revelan su propósito (Alidra et al, 2017). Asimismo, los nombres de identificadores constituyen una fuente rica para obtener información ausente (Liu et al, 2016).

Los modelos de tópicos ayudan a los usuarios a entender grandes colecciones de documentos cuando hay demasiados documentos, o poco tiempo para revisar cada documento individualmente (Smith et al, 2018). Estos modelos comprenden un análisis especializado (Loukachevitch et al, 2018), de los documentos para descubrir temas ocultos recurrentes, denominados tópicos (Badenes-Olmedo, 2017; Fan et al, 2018; Madhavan & Gopakumar, 2018; Mamakos & Chalkiadakis, 2018; Shi et al, 2018), los cuales constituyen grupos de términos que apoyan el entendimiento del contenido de los documentos (Silge & Robinson, 2017). Luego, los tópicos descubiertos permitirían anotar y organizar la información (Badenes-Olmedo, 2017), de tal forma que los usuarios pueden realizar una fácil navegación basándose en los tópicos de su interés. Si bien el modelo de tópicos fue desarrollado inicialmente para procesar de texto, éste puede ser aplicado sobre data en general (Santhiappan et al, 2018; Zhang et al, 2018).

El aplicar un análisis de tópicos en plataformas Git que requiera leer los archivos de código, puede presentar un problema por la cantidad de consultas a realizar, debido a las tasas límite del API de la plataforma en uso. Si se realizan cuantiosas invocaciones (e.g.; una invocación por cada archivo de código fuente en un repositorio de volumen), ello puede producir una demora considerable en la adquisición de data (Gousios & Spinellis, 2017) y/o fallos de los programas de consulta por problemas del servidor o la red (Rusk & Coady, 2014).

El presente trabajo presenta un enfoque para la extracción de tópicos en repositorios Git de código fuente, el cual emplea las regularidades frecuentes respecto al nombrado y la estructura de directorios de código fuente en un repositorio. Este trabajo no requiere leer las líneas de código fuente presentes en cada archivo del repositorio, reduciéndose la cantidad de consultas necesarias sobre el API; en cambio considera las palabras frecuentes presentes en los nombres de archivos, frecuencia examinada de acuerdo a la estructura de directorios de código. La contribución de este trabajo se resume como sigue:

1. Proporcionar un enfoque de extracción de tópicos para un repositorio de código fuente en Git.
2. Proporcionar un mecanismo que a partir de la estructura de directorios de código fuente de un repositorio, obtenga documentos aptos de ser procesados por modelos de tópicos.
3. Mostrar la factibilidad de utilizar mecanismos de modelo de tópicos sobre repositorios de código fuente en Git.

2. Modelo de tópicos para repositorios Git

Un tópico es un grupo de palabras significativas (Zhang et al, 2018), descrito mediante un conjunto de palabras (Badenes-Olmedo, 2017; Madhavan & Gopakumar, 2018). Un modelo de tópicos como objeto, consiste en k tópicos, cada uno representado por una lista de palabras (Belford et al, 2018). El modelo de tópicos como método, identifica automáticamente tópicos que ocurren dentro de una colección de documentos (Smith et al, 2018). La obtención de documentos para modelo de tópicos, se basa en tres criterios de organización de código:

1. Criterio de directorio singular (SIN): Un directorio D puede contener archivos de código fuente que guardan una relación entre sí.
2. Criterio de nombre de directorio (NOM): Directorios con un mismo nombre, pero ubicados en diferentes directorios contenedores, pueden contener archivos que guardan una relación entre sí por pertenecer a un contenedor con el mismo nombre.
3. Criterio de super directorio (SUP): Directorios con distinto nombre, dentro de un mismo super directorio D, pueden contener archivos que guardan una relación entre sí por pertenecer a un mismo super contenedor D.

Es conocida la importancia de emplear nombres de identificadores que sean significativos (Avidan & Feitelson, 2017), para facilitar el entendimiento del código (Jaffe et al, 2018), proporcionar contexto a las abstracciones en lenguajes de programación de alto nivel (Jaffe et al, 2018), atender la calidad del código (Schankin et al, 2018). Por ello, los desarrolladores emplean esfuerzo y consideración en la selección de nombres significativos (Jaffe et al, 2018; Alidra et al, 2017), y a menudo expresan conceptos en los nombres de clases (Alidra et al, 2017). Adicionalmente, en el desarrollo del software se pueden establecer convenciones de nombrado para mejorar la consistencia dentro de un proyecto y entre proyectos (Schankin et al, 2018). Este trabajo usa las palabras empleadas en nombres de archivos de código, los cuales en las ampliamente desarrolladas, aplicaciones orientadas a objetos, corresponden a clases. Tal vocabulario de palabras, es aquel que contiene una mayor cantidad de términos en común con otros vocabularios de bloques de construcción internos (i.e.; atributos, funciones, parámetros y comentarios) (Abebe et al, 2009). Adicionalmente, cabe resaltar que este vocabulario es el primero en ser considerado por un desarrollador cuando examina un sistema, antes que otros correspondientes a bloques de construcción internos (Alidra et al, 2017).

Las siguientes definiciones formulan la propuesta de documentos de modelos de tópicos para un repositorio de código fuente:

1. Listado de palabras en archivo de código fuente (*P(a)*): Listado de palabras contenidas en el nombre del archivo de código fuente *a*.
2. Conjunto de archivos de código fuente de un directorio (*A(d)*): Conjunto de archivos de código fuente presentes en el directorio *d*, sin considerar archivos pertenecientes a subdirectorios. $|A(d)|$ es la cantidad de archivos de código fuente presentes en el directorio *d*.
3. Listado de palabras de un directorio singular (*PSIN(d)*): Listado de *P(a)* para $\{a: A(d) \ni a\}$, donde *d* es un directorio.

4. Documento de directorio singular (*DSIN(d)*): Documento de texto que contiene las palabras de *PSIN(d)*, separadas por un espacio. |*DSIN(d)*| es la cantidad de palabras que contiene el documento.
5. Nombre de un directorio (*M(d)*): Nombre de un directorio *d*.
6. Conjunto de directorios de mismo nombre (*DN(m)*): Conjunto de directorios que tienen nombre *m*. |*DN(m)*| es la cantidad de directorios que tienen el nombre *m*.
7. Listado de palabras de un nombre de directorio (*PNOM(m)*): Listado de *P(a)*, para *{a: A(d) ϶ a}* y *{d: DN(m) ϶ d}*, donde m es un nombre de directorio.
8. Documento de directorios de mismo nombre (*DNOM(m)*): Documento de texto que contiene las palabras de *PNOM(m)*, separadas por un espacio. |*DNOM(m)*| es la cantidad de palabras que contiene el documento.
9. Conjunto de subdirectorios (*SD(d)*): Conjunto de los subdirectorios inmediatos contenidos en el directorio *d*. |*SD(d)*| es la cantidad de subdirectorios de *d*.
10. Listado de palabras de un super directorio (*PSUP(d)*): Listado de *P(a)*, para *{a: A(d) ϶ a}* y *{d: SD(d) ϶ d}*, donde *d* es un directorio.
11. Documento de super directorio (*DSUP(d)*): Documento de texto que contiene las palabras de *PSUP(d)*, separadas por un espacio. |*DSUP(d)*| es la cantidad de palabras que contiene el documento.
12. Conjunto de palabras de un documento (*PD(o)*): Conjunto de palabras presentes en el documento *o*.
13. Conjunto de directorios de repositorio (*CDI(r)*): Conjunto de los directorios de código fuente del repositorio *r*.
14. Documentos de repositorio de código para modelo de tópicos (*DMT(r)*): *DMT(r) = {DSIN(d):CDI(r)϶d, |DSIN(d)|>o} ∪ {DNOM(d):CDI(r)϶d, |DNOM(d)|>o} ∪ { DSUP(d):CDI(r)϶d, |DSUP(d)|>o}, r* es un repositorio.

El proceso de obtención de tópicos consiste en dos etapas. La etapa 1: generación de documentos de repositorio de código fuente, inicia con la invocación del API de la plataforma Git, para obtener la estructura de directorios del repositorio y los nombres de los archivos de código. Seguido se realizaría el tratamiento correspondiente, siguiendo las definiciones anteriormente dadas para obtener los documentos del repositorio REP. Acerca de la etapa 2: ejecución de un modelo de tópicos, considerando que los modelos de tópicos tienen una posición agnóstica acerca de cómo están definidos los documentos y las palabras (Zou et al, 2017), ésta se podría llevar a cabo ejecutando un modelo de tópicos convencional sobre los documentos generados en la fase 1.

3. Evaluación

Se implementaron aplicaciones para la generación de documentos de texto a partir de la estructura de directorios de un determinado repositorio en GitHub. Las palabras son extraídas de nombres de archivos de código según el estilo CamelCase, estilo predominante por su facilidad de escritura y su adopción (Guerrouj, 2013). Para la evaluación propiamente se utilizaron dos métricas aplicadas sobre 38 instancias correspondientes a populares proyectos de código abierto disponibles en la plataforma GitHub. Las tablas 1 y 2 muestran datos de cada proyecto.

Para la realización de la métrica 1: porcentaje de palabras significativas presentes en los tópicos, se han utilizado criterios de minería de patrones frecuentes de (Aggarwal & Han, 2014); se realizan las siguientes definiciones:

1. Soporte de una palabra en un directorio (*SUPD(p,d)*): número de archivos de *{a: A(d) ϶ a}* en que ocurre la palabra *p: P(a) ϶ p*; donde *d* es un directorio. *∀p∀d, 0 >= SUPD(p,d) <= |A(d)|*
2. Frecuencia de una palabra en un directorio (*FRD(p,d)*):
   FRD(p, d) = SUPD(p,d) / |A(d)|
3. Soporte de una palabra en un conjunto de directorios (*SUPC(p,m)*): Número de directorios de *{d: DN(m) ϶ d}* en que ocurre la palabra *p: P(a) ϶ p*, *a: A(d) ϶ a*; donde *m* es un nombre de directorio.
   ∀p∀m, 0 >= SUPC(p,m) <= |DN(p,m)|
4. Soporte de una palabra en un conjunto de directorios (*SUPC(p,Q)*): Número de directorios de Q en que ocurre la palabra *p: P(a) ϶ p*, *a: A(d) ϶ a*, *d: Q ϶ d*; donde *Q* es un conjunto de directorios y |Q| es la cantidad de directorios. *∀p∀Q, 0 >= SUPC(p,Q) <= |Q|*
5. Frecuencia de una palabra en un conjunto de directorios (*FRC(p,Q)*):
   FRC(p,Q) = SUPC(p,Q) / |Q|

PR	NA	ND	NPS	NT	PST	%PST
1 apache-accumulo-core	644	73	27	27	25	92.59
2 apache-camel-core	1779	99	66	66	62	93.94
3 apache-cassandra	1727	109	43	43	42	97.67
4 apache-flink	633	39	15	20	15	100
5 apache-groovy	680	63	26	26	24	92.31
6 apache-hadoop-common	973	77	37	37	33	89.19
7 apache-hbase-common	251	22	5	20	5	100
8 apache-ignite-core	3488	355	160	160	152	95
9 apache-kafka-core	204	24	8	20	7	87.5
10 apache-jmeter-core	532	55	20	20	19	95
11 apache-maven-core	334	54	29	29	29	100
12 apache-myfaces-impl	907	76	42	42	41	97.62
13 apache-nutch	298	41	12	20	12	100
14 apache-openjpa-persistence	235	7	0	20	0	0
15 apache-shiro-core	232	28	12	20	11	91.67
16 apache-tomcat	1672	140	54	54	50	92.59
17 apache-zookeeper	268	19	4	20	4	100
18 eclipse-eclipselink.runtime-jpa	564	63	20	20	20	100
19 eclipse-hudson	770	55	27	27	26	96.3

Tabla 1 – Proyectos GitHub – Parte 1. PR: identificador del proyecto, NA: número de archivos de código fuente, ND: número de directorios, NPS: número de palabras significativas, NT: número de tópicos, PST: número de palabras significativas presentes en tópicos, %PST: porcentaje de palabras significativas presentes en tópicos.

Seguido, las palabras significativas son obtenidas de la unión de los siguientes conjuntos de palabras, considérese a *minsup* como un umbral mínimo:

1. Palabras *p: FRP(p,d)* ≥ *minsup*, *d:CDI(r)* ∋ *d*, |*A(d)*| >= *3*; donde *r* es un repositorio. Este conjunto de palabras se denominará *FSIN(r)*. Véase la sección a) de la figura 1 que muestra la palabra significativa *Metrics*, presente en un directorio de *Apache Hadoop*.
2. Palabras *p: FRC(p,DN(m))* ≥ *minsup*, |*DN(m)*| >= *3*, *m=M(d)*, *d:CDI(r)* ∋ *d*; donde *r* es un repositorio. Este conjunto de palabras se denominará *FNOM(r)*. Véase la sección b) de la figura 1 que muestra la palabra significativa *Scope*, presente en los directorios de *Spring Framework Web*.
3. Palabras *p: FRC(p,SD(d))* ≥ *minsup*, *d:CDI(r)* ∋ *d*, |*SD(d)*| >= *3*; donde *r* es un repositorio. Este conjunto de palabras se denominará *FSUP(r)*. Véase la sección c) de la figura 1 que muestra la palabra significativa *Factory*, presente en los directorios de *Mule*.

Luego, la métrica 1 es calculada obteniendo el porcentaje de palabras significativas presentes en el conjunto de tópicos generado.

PR	NA	ND	NPS	NT	PST	%PST
20 eclipse-m2e-core	222	28	9	20	9	100
21 eclipse-vert.x	447	58	13	20	13	100
22 elasticsearch-server	2889	303	158	158	154	97.47
23 hibernate-orm-core	3266	288	96	96	93	96.88
24 mule-core	1038	175	77	77	74	96.1
25 mybatis-3	381	54	28	28	24	85.71
26 pentaho-kettle-core	377	38	18	20	17	94.44
27 pentaho-mondrian	714	29	14	20	12	85.71
28 pentaho-reporting-report-designer	722	70	36	36	36	100
29 primefaces	1046	239	188	188	186	98.94
30 rapidminer-studio	3589	333	139	139	137	98.56
31 redhat-drools	1206	71	41	41	41	100
32 redhat-jbpm-flow	276	48	23	23	23	100
33 redhat-optaplanner-core	569	190	64	64	62	96.88
34 redhat-wildfly-ee	233	20	12	20	11	91.67
35 spring-batch-core	438	47	18	20	17	94.44
36 spring-framework-web	616	54	38	38	35	92.11
37 spring-security-oauth2	249	42	28	28	26	92.86
38 vaadin-framework	551	26	10	20	9	90

Tabla 2 – Proyectos GitHub – Parte 2. Véase leyendas de tabla 1.

La métrica 2: coherencia de los tópicos, ha sido calculada utilizando lo definido por Umass (Mimno et al, 2011), métrica que correlaciona adecuadamente con el criterio humano (Yao et al, 2017). A fin de tener una evaluación comparativa se utilizaron tres diferentes conjuntos de datos (CD):

1. CD1: conjunto de documentos $\{DSIN(d): CDI(r) \ni d, |DSIN(d)|>0\}$, donde r es un repositorio.
2. CD2: $DMT(r)$, donde r es un repositorio. Conjunto de la propuesta.
3. CD3: $\{ DSIN(d): CDI(r) \ni d, |DSIN(d)|>0, \exists p: PSIN(d) \ni p, FSIN(r) \ni p \} \cup \{DNOM(d): CDI(r) \ni d, |DNOM(d)|>0, \exists p: PNOM(d) \ni p, FNOM(r) \ni p \} \cup \{ DSUP(d): CDI(r) \ni d, |DSUP(d)|>0, \exists p: PSUP(d) \ni p, FSUP(r) \ni p \}$. Donde r es un repositorio.

Para la obtención de tópicos se utilizó el modelo de tópicos LDA (*Latent Dirichlet Allocation*) implementado en Mallet, implementación ampliamente utilizada para los modelos de tópicos (Belford et al, 2018), realizándose cien iteraciones sobre cada instancia en tratamiento. La cantidad de tópicos a obtener fue establecida con el mismo número de palabras significativas presentes en la instancia, y un valor predeterminado de veinte para aquellas que tengan menos de veinte palabras significativas (es común modelar decenas de tópicos en un corpus (Le & Lauw, 2017)), ello a fin que el modelo de tópicos pueda descubrir palabras frecuentes que no serían extraídas con el mecanismo presentado anteriormente. Asimismo, el valor *minsup* utilizado para obtener palabras significativas, fue establecido en 0.75.

4. Resultados

Para la métrica 1, CD1, CD2, y CD3 alcanzaron el mismo porcentaje de palabras significativas en los tópicos, obteniéndose en promedio más del 95% de palabras significativas. Las palabras significativas se presentan en una medida considerable, teniéndose un promedio mayor a 42, cabe señalar que sólo los proyectos más pequeños son los que presentan algunos casos con presencia ínfima de palabras significativas. Los resultados de la métrica 1 son mostrados en las tablas 1 y 2.

Para la métrica 2, la propuesta de este trabajo, plasmada en CD2, obtuvo los mejores resultados promedio, tanto para los mejores niveles de coherencia (i.e.; es 2MI menor a otros) como para la coherencia promedio de los tópicos (i.e.; 2AV menor a otros). Si bien otros trabajos han utilizado la minería de patrones frecuentes para modelos de tópicos (Peng et al, 2018), este trabajo sólo lo empleo para tratar de formar una alternativa pre-procesada que pueda superar a la propuesta, CD3 es un subconjunto de CD2, con un recorte basado en palabras significativas. CD2 sólo fue superado por CD3 al comparar sus tópicos de bajo nivel de coherencia en promedio. Los resultados de la métrica 2 son mostrados en las tablas 3 y 4.

El vocabulario de palabras empleados en este trabajo ha sido obtenido de nombres de archivos, vocabulario que en el predominante mundo de software orientado a objetos, es aquel con mayor cantidad de términos en común con otros vocabularios (Abebe et al, 2009). Siendo el modelo de tópicos uno de los procesos más útiles para reducir la representación de grandes volúmenes de data (Potha & Stamatatos, 2018), las palabras presentes en los tópicos obtenidos permiten recortar la cantidad de conceptos a abordar a la hora de analizar un repositorio Git de código fuente.

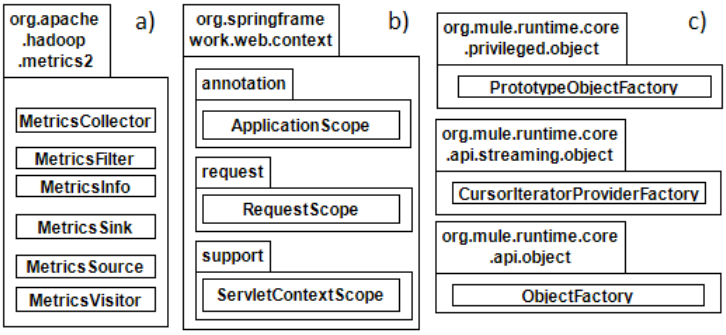


Figura 1 – Ejemplos de palabras significativas

Usualmente los enfoques de modelo de tópicos consideran a cada documento como exclusivamente un "bolso de palabras", sin considerar relaciones (Loukachevitch et al, 2018); asunción bastante criticada en el modelo de tópicos (Kochedykov et al, 2017). El enfoque presentado no ha considerado a los directorios simplemente como "bolsos de nombres" simplemente (e.g.; DSIN), en cambio se han definido otras formas como DNOM y DSUP que toman en consideración la estructura de directorios presente.

Si bien los enfoques convencionales de modelo de tópicos proporcionan un análisis potente sobre textos normales, éstos pierden efectividad en textos cortos (Li et al, 2018; Shi et al, 2018). Este enfoque evade el generar documentos demasiado cortos, agrupando palabras de acuerdo a la estructura de directorios. Los criterios de organización de código aplicados fueron de ayuda para evitar disponer de pocas palabras significativas, lo cual hubiera dificultado la captura de información de coocurrencia utilizada en el modelo convencional (Shi et al, 2018).

Se tiene el inconveniente de que el número de tópicos debe ser establecido manualmente en modelos de tópicos basados en LDA (Alkhodair et al, 2018). Incluso los modelos de tópicos tradicionales pueden proporcionar tópicos que son de baja calidad, o que están desalineados al entendimiento del experto de dominio del corpus de palabras (Smith et al, 2018). Sin embargo, el mecanismo propuesto para la obtención de documentos para modelo de tópicos desde código fuente, puede ser empleado con otros modelos de tópicos y variantes asociadas.

5. Trabajos relacionados.

El modelo de tópicos es un tema que recibe mucha atención en investigación (Tang et al, 2018), y ha sido aplicado en diversas áreas fuera de la minería de texto sobre lenguaje natural (Badenes-Olmedo, 2017; Bukhari & Liu, 2018; Fan et al, 2018; Madhavan & Gopakumar, 2018; Mamakos & Chalkiadakis, 2018).

Se realizó una búsqueda de trabajos similares en las bibliotecas digitales ACM, IEEE Xplore y ScienceDirect. En ACM se obtuvieron tres resultados, (Sharma et al, 2017) trata con archivos *readme*. En IEEE Xplore se encontraron nueve resultados, (Kononenko

PR	1MI	1MA	1AV	2MI	2MA	2AV	3MI	3MA	3AV
1	-828.9	-381.1	-641.1	-840.0	-38.1	-405.2	-676.4	-12.1	-444.6
2	-759.9	-33.5	-520.3	-786.0	-7.7	-282.8	-650.1	-22.2	-475.8
3	-868.6	-150.2	-550.5	-708.8	0.0	-265.7	-736.6	-73.9	-510.8
4	-825.4	-49.5	-508.7	-736.3	-23.7	-346.8	-620.2	-142.8	-474.8
5	-848.0	-290.4	-609.7	-869.5	-50.8	-354.2	-740.2	-13.0	-513.3
6	-823.2	-27.5	-534.6	-771.3	-17.5	-301.9	-688.3	-87.7	-431.8
7	-697.0	-102.5	-489.0	-629.2	-7.7	-373.7	-235.4	0.0	-101.7
8	-857.4	-91.4	-601.0	-821.4	-11.4	-452.2	-727.4	-52.1	-487.1
9	-743.1	-366.7	-562.7	-643.5	-22.8	-386.9	-309.2	0.0	-153.6
10	-824.1	-33.1	-479.7	-810.1	-27.5	-235.0	-701.8	-27.7	-450.0
11	-802.1	-29.2	-492.4	-762.1	-16.9	-395.1	-509.2	-17.6	-362.9
12	-779.8	-157.2	-581.0	-820.7	-16.8	-393.7	-679.0	-155.5	-499.4
13	-782.2	-309.6	-638.8	-733.8	-30.4	-493.6	-572.5	-96.1	-332.2
14	-605.6	-26.1	-430.0	-607.2	0.0	-391.6	0.0	0.0	0.0
15	-716.3	-28.2	-515.8	-758.7	-353.2	-604.2	-624.3	-128.1	-461.1
16	-872.8	-26.2	-651.7	-821.5	-12.1	-381.0	-688.9	-11.0	-489.9
17	-716.4	-114.8	-528.4	-728.4	0.0	-414.5	-475.4	0.0	-247.1
18	-727.2	-68.8	-533.4	-745.5	-44.8	-383.6	-682.5	-23.0	-434.2
19	-869.3	-106.8	-461.1	-744.7	-27.1	-266.7	-617.7	-70.4	-421.1

Tabla 3 – Resultados de métrica de coherencia – Parte 1. PR: identificador del proyecto, 1MI: coherencia mínima de CD1, 1MA: coherencia máxima de CD1, 1AV: coherencia promedio de CD1, 2MI: coherencia mínima de CD2, 2MA: coherencia máxima de CD2, 2AV: coherencia promedio de CD2, 3MI: coherencia mínima de CD3, 3MA: coherencia máxima de CD3, 3AV: coherencia promedio de CD3.

et al, 2018) y (Yu et al, 2018) examinan *pull requests*. ScienceDirect no obtuvo ningún resultado. Otros resultados no eran concernientes. Ningún resultado encontrado trata el modelo de tópicos aplicado sobre el código fuente de un repositorio Git. Las consultas empleadas fueron las siguientes:

ACM: recordAbstract:(“topic model” “topic models”) AND recordAbstract:(git github gitlab bitbucket)

IEEE Xplore: (“Abstract”:topic model OR “Abstract”:topic models) AND (“Abstract”:git OR “Abstract”:github OR “Abstract”:gitlab OR “Abstract”:bitbucket)

ScienceDirect: ABS(“topic model” OR “topic models”) AND ABS(git OR github OR gitlab OR bitbucket)

PR	1MI	1MA	1AV	2MI	2MA	2AV	3MI	3MA	3AV
20	-715.1	-465.0	-610.5	-692.7	-203.2	-556.1	-400.7	0.0	-219.1
21	-853.5	-229.0	-606.8	-805.9	-122.8	-523.9	-628.6	-109.4	-461.8
22	-813.3	-57.8	-532.5	-812.3	-5.2	-430.1	-743.6	0.0	-435.5
23	-878.8	-52.4	-585.9	-785.5	-14.3	-318.1	-680.2	-65.3	-457.7
24	-847.1	-209.5	-602.6	-788.3	-44.7	-564.7	-666.7	-25.4	-479.0
25	-770.0	-401.8	-569.7	-697.0	-31.3	-475.1	-615.7	-76.3	-457.8
26	-768.3	-139.3	-554.3	-721.2	-44.3	-360.8	-664.2	-103.8	-474.7
27	-783.4	-22.7	-395.7	-573.1	-42.7	-195.8	-528.2	-92.5	-451.5
28	-814.9	-280.7	-641.9	-769.0	-40.2	-319.8	-669.9	-26.2	-447.5
29	-655.3	0.0	-276.4	-667.7	0.0	-372.2	-599.7	0.0	-287.5
30	-854.0	-66.1	-605.1	-816.8	-8.5	-336.6	-761.9	-7.7	-460.3
31	-821.3	-160.4	-524.2	-684.0	-22.9	-300.5	-721.7	-111.1	-482.3
32	-771.5	-230.5	-536.0	-732.0	-40.5	-516.7	-663.8	-27.1	-440.9
33	-727.5	-191.6	-465.8	-762.0	-87.3	-518.1	-616.3	-32.5	-414.9
34	-667.8	-151.3	-506.9	-692.7	-14.5	-485.1	-545.1	-0.7	-351.6
35	-769.5	-318.9	-613.8	-799.8	-71.8	-462.7	-721.9	-31.9	-436.1
36	-735.8	-119.7	-550.7	-764.2	-22.6	-430.8	-672.0	-204.2	-529.3
37	-626.2	-157.7	-474.4	-693.3	-20.5	-556.0	-524.3	-0.2	-370.5
38	-692.3	-88.1	-412.2	-758.8	-20.7	-315.2	-572.9	0.0	-319.4

Tabla 4 – Resultados de métrica de coherencia – Parte 2. Véase leyendas de tabla 3.

6. Conclusiones

Este trabajo ha presentado un enfoque de extracción de tópicos para un repositorio de código fuente en Git. La aplicación de este enfoque no requiere leer o realizar *parsing* sobre los archivos de código fuente, por tanto, utiliza una cantidad reducida de consultas necesarias a realizar sobre el API, aminorando el riesgo de sobrecargar al servidor de la plataforma Git, y evitando la transferencia de bytes de archivos sobre la red; característica relevante en repositorios Git, que son manipulados en forma global y manejan grandes volúmenes de archivos.

Asimismo se ha presentado un mecanismo que genera documentos de texto convencionales, que son aptos de ser procesados por diversos modelos de tópicos; tal mecanismo, en forma particular, emplea la estructura de directorios de código fuente del repositorio, aprovechando utilizar relaciones subyacentes, producto de criterios comunes que se aplican en la organización de archivos de código fuente.

Finalmente, mediante la aplicación de la propuesta sobre populares proyectos alojados en GitHub, se ha mostrado la factibilidad de utilizar mecanismos de modelo de tópicos sobre repositorios de código fuente en Git.

Referencias

Abebe, S.L., Haiduc, S., Marcus, A., Tonella, P., & Antoniol, G. (2009). *Analyzing the Evolution of the Source Code Vocabulary*. Paper presented at 2009 European Conference on Software Maintenance and Reengineering. https://doi.org/10.1109/CSMR.2009.61

Aggarwal, C.C., & Han, J. (2014). *Frequent Pattern Mining*. Springer Publishing Company, Incorporated.

Alidra, A., Saker, M., Anquetil, N. & Ducasse, S. (2017). *Identifying class name inconsistency in hierarchy: a first simple heuristic*. Paper presented at 12th edition of the International Workshop on Smalltalk Technologies. https://doi.org/10.1145/3139903.3139920

Alkhodair, S.A., Fung, B.C.M., Rahman, O., & Hung, P.C.K. (2018). Improving interpretations of topic modeling in microblogs. J. *Assoc. Inf. Sci. Technol.* 69(4), 528-540. https://doi.org/10.1002/asi.23980

Avidan, E. & Feitelson, D.G. (2017). *Effects of variable names on comprehension an empirical study*. Paper presented at 25th International Conference on Program Comprehension. https://doi.org/10.1109/ICPC.2017.27

Badenes-Olmedo, C., Redondo-García, J.L., & Corcho, O. (2017). *Efficient Clustering from Distributions over Topics*. Paper presented at Knowledge Capture Conference. https://doi.org/10.1145/3148011.3148019

Belford, M., Mac Namee, B., & Greene, D. (2018). Stability of topic modeling via matrix factorization. *Expert Syst. Appl.* 91(C). https://doi.org/10.1016/j.eswa.2017.08.047

Bukhari, A. & Liu, X. (2018). A Web service search engine for large-scale Web service discovery based on the probabilistic topic modeling and clustering. *Serv. Oriented Comput. Appl.* 12(2), 169-182. https://doi.org/10.1007/s11761-018-0232-6

Fan, Y., Zhou, Q., Yue, W., & Zhu, W. (2018). A dynamic causal topic model for mining activities from complex videos. *Multimedia Tools Appl.* 77(9), 10669-10684. https://doi.org/10.1007/s11042-017-4760-4

Guerrouj, L. (2013). *Normalizing source code vocabulary to support program comprehension and software quality*. Paper presented at 2013 International Conference on Software Engineering.

Gousios, G. & Spinellis, D. (2017). *Mining software engineering data from GitHub*. Paper presented at 39th International Conference on Software Engineering Companion. https://doi.org/10.1109/ICSE-C.2017.164

Jaffe, A., Lacomis, J., Schwartz, E.J., Le Goues, C., & Vasilescu, B. (2018). *Meaningful variable names for decompiled code: a machine translation approach*. Paper presented at 26th Conference on Program Comprehension. https://doi.org/10.1145/3196321.3196330

Kochedykov, D., Apishev, M., Golitsyn, L., & Vorontsov, K. (2017). *Fast and Modular Regularized Topic Modelling*. Paper presented at 21st Conference of Open Innovations Association FRUCT. https://doi.org/10.23919/FRUCT.2017.8250181

Kononenko, O., Rose, T., Baysal, O., Godfrey, M., Theisen, D. & de Water, D. (2018). *Studying Pull Request Merges: A Case Study of Shopify's Active Merchant*. Paper presented at 2018 IEEE/ACM 40th International Conference on Software Engineering: Software Engineering in Practice Track.

Le, T.M.V., & Lauw, H.W. (2017). *SemVis: Semantic Visualization for Interactive Topical Analysis*. Paper presented at 2017 ACM on Conference on Information and Knowledge Management. https://doi.org/10.1145/3132847.3133181

Li, X., Li, C., Chi, J., & Ouyang, J. (2018). Short text topic modeling by exploring original documents. *Knowl. Inf. Syst*. 56(2), 443-462. https://doi.org/10.1007/s10115-017-1099-0

Liu, H., Liu, Q., Staicu, C., Pradel, M., & Luo, Y. (2016). *Nomen est omen: exploring and exploiting similarities between argument and parameter names*. Paper presented at 38th International Conference on Software Engineering https://doi.org/10.1145/2884781.2884841

Loukachevitch, N., Ivanov, K., & Dobrov, B. (2018). *Thesaurus-Based Topic Models and Their Evaluation*. Paper presented at 8th International Conference on Web Intelligence, Mining and Semantics. https://doi.org/10.1145/3227609.3227659

Madhavan, M., & Gopakumar, G. (2018). *A tf-idf based topic model for identifying lncRNAs from genomic background*. Paper presented at 33rd Annual ACM Symposium on Applied Computing. https://doi.org/10.1145/3167132.3167133

Mamakos, M., & Chalkiadakis, G. (2018). *Overlapping Coalition Formation via Probabilistic Topic Modeling*. Paper presented at 17th International Conference on Autonomous Agents and MultiAgent Systems.

Mimno, D., Wallach, H.M., Talley, E., Leenders, M., & McCallum, A. (2011). *Optimizing semantic coherence in topic models*. Paper presented at Conference on Empirical Methods in Natural Language Processing.

Peng, M., Zhu, J., Wang, H., Li, X., Zhang, Y., Zhang, X., & Tian, G. (2018). Mining Event-Oriented Topics in Microblog Stream with Unsupervised Multi-View Hierarchical Embedding. *ACM Trans. Knowl. Discov. Data*. 12(3). https://doi.org/10.1145/3173044

Potha, N., & Stamatatos, E. (2018). *Intrinsic Author Verification Using Topic Modeling*. Paper presented at 10th Hellenic Conference on Artificial Intelligence. https://doi.org/10.1145/3200947.3201013

Rusk, D. & Coady, Y. (2014). *Location-Based Analysis of Developers and Technologies on GitHub*. Paper presented at 28th International Conference on Advanced Information Networking and Applications Workshops. https://doi.org/10.1109/WAINA.2014.110

Santhiappan, S., Chelladurai, J., & Ravindran, B. (2018). *A novel topic modeling based weighting framework for class imbalance learning*. Paper presented at ACM India Joint International Conference on Data Science and Management of Data. https://doi.org/10.1145/3152494.3152496

Schankin, A., Berger, A., Holt, D.V., Hofmeister, J.C., Riedel, T., & Beigl, M. (2018). *Descriptive compound identifier names improve source code comprehension*. Paper presented at 26th Conference on Program Comprehension. https://doi.org/10.1145/3196321.3196332

Sharma, A., Thung, F., Kochhar, P.S., Sulistya, A., & Lo, D. (2017). *Cataloging GitHub Repositories*. Paper presented at 21st International Conference on Evaluation and Assessment in Software Engineering. https://doi.org/10.1145/3084226.3084287

Shi, T., Kang, K., Choo, J., & Reddy, C.K. (2018). *Short-Text Topic Modeling via Non-negative Matrix Factorization Enriched with Local Word-Context Correlations*. Paper presented at 2018 World Wide Web Conference. https://doi.org/10.1145/3178876.3186009

Silge, J., & Robinson, D. (2017). *Text Mining with R: A Tidy Approach (1st ed.)*. O'Reilly Media, Inc.

Smith, A., Kumar, V., Boyd-Graber, J., Seppi, K., & Findlater, L. (2018). *Closing the Loop: User-Centered Design and Evaluation of a Human-in-the-Loop Topic Modeling System*. Paper presented at 23rd International Conference on Intelligent User Interfaces. https://doi.org/10.1145/3172944.3172965

Tang, Y., Mao, X., Huang, H., Shi, X., & Wen, G. (2018). Conceptualization topic modeling. *Multimedia Tools*. 77(3).

Yao, L., Zhang, Y., Chen, Q., Qian, H., Wei, B., & Hu, Z. (2017). Mining coherent topics in documents using word embeddings and large-scale text data. *Eng. Appl. Artif. Intell*. 64(C).

Yu, S., Xu, L., Zhang, Y., Wu, J., Liao, Z., and Li, Y. (2018). *NBSL: A Supervised Classification Model of Pull Request in Github*. Paper presented at IEEE International Conference on Communications. https://doi.org/10.1109/ICC.2018.8422103

Zhang, Y., Ma, J., Wang, Z., Chen, B., & Yu, Y. (2018). Collective topical PageRank: a model to evaluate the topic-dependent academic impact of scientific papers. *Scientometrics*. 114(3).

Zou, W., Zhao, W., Chen, J.J., & Perkins, R. (2017). *Best Setting of Model Parameters in Applying Topic Modeling on Textual Documents*. Paper presented at 8th ACM International Conference on Bioinformatics, Computational Biology,and Health Informatics. https://doi.org/10.1145/3107411.3108195

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 14/10/2018
Aceitação/Acceptance: 04/12/2018

Prototipo de monitoreo basado en inferencia difusa para la prevención de incendios en habitaciones urbanas mediante tecnologías IoT

Kevin Riaño-Vargas[1], Jorge Alonso-Echeverri[1], Paulo Gaona-García[1], Carlos Montenegro-Marín[1]

karianov@correo.udistrital.edu.co, jialonsoe@correo.udistrital.edu.co, pagaonag@udistrital.edu.co, cemontenegrom@udistrital.edu.co

[1] Facultad de Ingeniería, Universidad Distrital Francisco José de Caldas, Bogotá D.C., Colombia.

Pages: 913–922

Resumen: Las habitaciones estructurales de ciudades urbanas están expuestas a la ocurrencia de incendios en la medida que dentro de estas se encuentren electrodomésticos o materiales inflamables, pero no dispositivos que posibiliten prevenir el desastre. Analizar las propiedades ambientales correspondientes a medición de temperatura y presencia de humo dentro de una habitación es un mecanismo que permite monitorear el comportamiento de la estructura con relación a los incendios estructurales, de manera que el usuario obtenga esta información y evite la materialización de los daños que pueden ocasionarse. Este artículo presenta el desarrollo de un prototipo de monitoreo usando Internet de las Cosas (IoT) para prevenir incendios estructurales sobre habitaciones urbanas por medio de un modelo de inferencia difusa, partiendo de un ensamblaje físico dedicado a la captura de datos ambientales hasta el envío de recomendaciones hacia los dispositivos móviles de los ciudadanos.

Palabras-clave: Prototipos; Montajes físicos; Internet de las Cosas; Analítica de datos; Algoritmos heurísticos.

Prototype of monitoring based on fuzzy inference for the prevention of fires in urban rooms using IoT technologies

Abstract: Structural rooms of urban cities are exposed to the occurrence of fires insofar as electrical appliances or flammable materials are found inside them, but not devices that make possible to prevent the disaster. Analyze the environmental properties corresponding to temperature measurement and presence of smoke inside a room is a mechanism that allows to monitor the behavior of the structure in relation to structural fires, so that the user obtains this information and prevents the materialization of the damage that can be caused. This article presents the development of a monitoring prototype using the Internet of Things (IoT) to prevent structural fires in urban rooms through a fuzzy inference model, starting from a physical assembly dedicated to the capture of environmental data until sending recommendations to mobile devices of citizens.

Keywords*:* Prototypes; Physical assemblies; Internet of Things; Data analytics; Heuristic algorithms.

1. Introducción

Las instituciones o entidades dedicadas al cuidado de la vida humana en el ámbito civil entienden que el diseño y construcción de viviendas deben contemplar la ocurrencia de desastres que afecten la integridad de los habitantes (Cábová, Blesák & Wald, 2016). Los incendios estructurales son un desastre que ocurre en viviendas urbanas, cuyas causas incluyen la anomalía o alteración de las propiedades ambientales dentro de la habitación en la que se generan (Asociación Nacional de Sistemas de Rociadores Automáticos Contra Incendios [ANRACI], 2018).

Este artículo presenta el desarrollo de un prototipo de monitoreo sobre habitaciones para la vivienda en sectores urbanos, usando IoT para el diseño y ensamblaje del montaje físico dedicado a la captura de datos y envío de acciones recomendadas hacia los dispositivos móviles de los usuarios correspondientes (Chow & Ma, 2017), con el fin de generar mecanismos de pronóstico y prevención del riesgo asociado a los incendios estructurales, soportado en un modelo basado en inferencia difusa.

En las siguientes secciones se describen los trabajos relacionados con esta propuesta de investigación, seguido de la descripción del modelo que contiene el diseño y detalle sobre el prototipo basado en inferencia difusa, luego el caso de estudio para la evaluación del desempeño de esta propuesta, y finalmente la discusión de los resultados junto a los trabajos futuros de la investigación en la sección de conclusiones.

2. Trabajos Relacionados

Existen varias propuestas de investigación y desarrollo alrededor del mundo enfocadas en darle tratamiento a la problemática de incendios en habitaciones estructurales dentro de zonas urbanas, asumiendo una postura de prevención, control y reacción oportuna de dichos desastres. Los modelos independientes de las propiedades apuntan a un elemento biológico similar, antecedentes de incendios, elementos actuales o la generación de bases de conocimiento para la creación de sistemas predictores cada vez más acertados y con el mínimo porcentaje de error posible.

Dentro de ese universo de investigaciones, existen los montajes multisensor, ampliamente utilizados, y que se suman al uso de la inteligencia artificial, para originar propuestas que involucran la implementación de machine learning permitiendo predecir la ocurrencia de incendios (Jiang, Liu, Li & Ma, 2017). En este trabajo, se contempla el tratamiento de cuatro elementos: temperatura, concentración de monóxido de carbono (CO), humo y longitud de onda de una señal infrarroja, factores que se convierten en los insumos necesarios para obtener predicciones de ocurrencia de incendios, usando Extreme Learning Machine (ELM) y Online Sequential Extreme Learning Machine (OS-ELM) cuyas técnicas hacen uso de neuronas para un buen aprendizaje del sistema (Huang, Liang, Rong, Saratchandran & Sundararajan, 2005), esta propuesta concluye que en términos de tiempo y ocurrencia del mismo la mejor es OS-ELM (Jiang et al., 2017).

La domótica representa otra línea de investigación útil e importante para abordar la problemática; optando por un conjunto de sensores de temperatura, humedad y detectores de gases trabajando paralelamente, creando la posibilidad de analizar comportamientos o escenarios de incendios, alarmado a sus habitantes de manera oportuna (Jose & Malekian, 2017).

Otra vía trabajo, es la definición de una red de sensores inalámbricos (WSN) en un espacio de automatización en donde se ubica una población de personas mayores (Ransing & Rajput, 2015) sumado a protocolos de comunicación de bajo costo (ZigBee) para informar y efectuar una acción oportunamente en la ocurrencia de un incendio.

A nivel de tecnología, una propuesta de transmisión de información en escenarios de determinados, está definida por la creación de un esquema similar a una trama (como las usadas en redes de comunicación) de 29 bits (Kyung & Hong-Hee, 2004), incluyendo todos los elementos señalados como importantes y que deban ser compartidos.

Adicional a los trabajos antes mencionados, otras áreas de investigación fundamental inmersa en ambientes de predicción y análisis de registros, es la Lógica Difusa y el internet de las cosas (IoT) de manera que, la concepción de ensamblajes que involucren dispositivos físicos equipados con software que aplique y configure técnicas de lógica difusa, así como, implementación de modelos IoT para la comunicación y transmisión de los datos (Riaño, Salcedo & Correa, 2018), elevan las posibilidades de prevenir escenarios de incendios o minimizar los daños de éstos. Por otro lado, se debe complementar toda la gestión de recolección, análisis y estudio de los registros con una estrategia IFTTT (If This, Then That) para el envío masivo o controlado de alertas y mensajes que sean necesarios (Yoddumnern, Yooyativong & Chaisricharoen, 2017).

Así pues, se recogen los diferentes caminos que han sido ejecutados para darle tratamiento a la situación problema que se ha descrito en secciones anteriores, objeto del actual trabajo de investigación.

3. Descripción del Prototipo

3.1. Metodología de trabajo

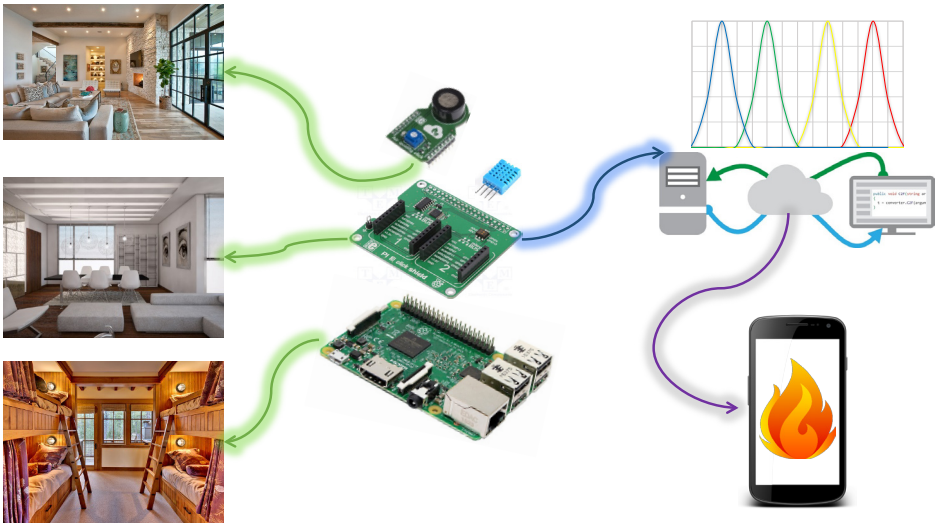


Figura 1 – Escenario de trabajo

Como lo muestra la Figura 1, el escenario de trabajo para este prototipo consiste en la ubicación de varias habitaciones estructurales en zonas urbanas de la ciudad, dentro de las cuales se hallará un ensamblaje físico dedicado al envío de los datos capturados sobre la medición de temperatura y presencia de humo en las estaciones de monitoreo, de forma que el servidor pueda enviar las acciones recomendadas desde el modelo de inferencia difusa a los dispositivos móviles que estén vinculados al servicio de acuerdo a la habitación correspondiente.

Cada habitación es tratada como una estación de monitoreo y sus datos se envían en una estructura que permite identificar su procedencia y los valores capturados en forma de parámetros diferenciados. Los datos entran al servidor y éste se encarga de procesar las propiedades ambientales recibidas en el modelo de inferencia difusa para pronosticar las capturas en pulsaciones futuras de acuerdo al comportamiento de la habitación, y para definir las acciones recomendadas que serán enviadas a los dispositivos móviles asociados al proceso de monitorización de incendios estructurales.

3.2. Ensamblaje Físico

Volviendo a la Figura 1 que describe el escenario de trabajo, en el centro de la misma se muestra una organización de elementos electrónicos cuya conectividad permite pasar desde las habitaciones estructurales para el monitoreo hacia el servicio de análisis de datos para el envío de notificaciones sobre el riesgo de ocurrir un incendio estructural.

Dicha organización comprende desde Raspberry Pi 3 Modelo B hasta dos sensores periféricos: uno de calidad de aire integrado con un MQ-135 para detectar gases venenosos en interiores; y uno digital de temperatura y humedad en interiores, ayudados por una placa extensible intermediaria. En la Tabla 1 se estructuran los dispositivos utilizados para el ensamblaje con sus nombres y características físicas asociadas.

Dispositivo	**Características**
Raspberry Pi 3 Modelo B	Computadora de placa reducida con LAN inalámbrica y bluetooth.
Pi 3 click shield	Escudo adaptable a los 40 pines de la Raspberry Pi 3 Modelo B con dos buses para la conexión de periféricos específicos, integrado con un convertidor análogo a digital MCP3204 de 12 bits.
Air quality click	Sensor que reacciona a gases amoníacos (NH3), óxidos de nitrógeno (NOx), bencenos, humo, dióxidos de carbono (CO2), entre otros.
DHT11	Sensor capacitivo de humedad y termistor para medir el aire circundante
Resistencia	10Ω (Ohmios)
Cable puente para prototipos	Tres (3) cables con terminales macho–macho.
Placa de inserción	Tarjeta de 400 pines sin soldadura para prototipos.

Tabla 1 – Descripción de los dispositivos (Satoshi, 2017; MikroElektronika, 2014)

3.3. Interfaz de Comunicación de Datos

A partir de la disposición física establecida para el trabajo de investigación, la siguiente fase comprende la captura, registro y uso de esas señales desde una perspectiva lógica

también conocida como la dimensión de aplicación informática. Es así que, se propone elaborar una interfaz de comunicación de datos, dado que permite la conexión entre el ensamblaje físico, el modelo de predicción y los dispositivos móviles vinculados.

Para llevar a cabo una ejemplificación de este proceso, la Figura 2 presenta la interfaz que se compone de tres elementos fundamentales: capturador, encargado de la comunicación entre el ensamblaje físico y el registro en la base de datos con un intervalo entre pulsaciones de diez (10) segundos; proceso de inferencia, delegado para consultar los registros capturados y generar respuestas en forma de acciones recomendadas para cada estación de monitoreo; y envío de las notificaciones, que construye el mensaje correspondiente al resultado del modelo y lo envía a los dispositivos móviles configurados a través del sistema FCM (Firebase Cloud Messaging) que provee Google.

Figura 2 – Proceso de captura de datos

3.4. Proceso de Inferencia

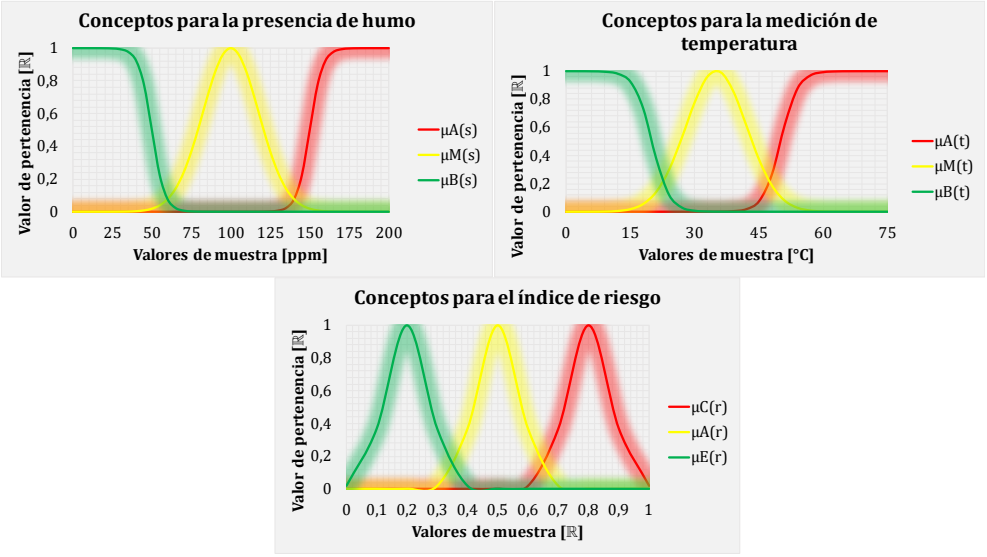


Figura 3 – Conceptos propuestos para el modelo de inferencia

La idea de formular un modelo basado en un proceso de inferencia, consiste en trabajar con conceptos y reglas lógicas más flexibles que las usuales o convencionales, de forma que puedan procesarse los datos con variables de configuración que admitan una

formulación abierta a la lingüística y no restringida a respuestas binarias (Topaloğlu & Pehlıvan, 2018). Por consiguiente, como se muestra en la Figura 3, se proponen los conceptos que traducirán la configuración lingüística hacia el contexto matemático para poder tratar los datos capturados durante el proceso de monitorización.

De acuerdo con Marta (2016), la medición de temperatura, el registro ideal a la que se debe mantener la calefacción en un interior es de 20 o 21 grados Celsius (°C), para salas estructurales, con lo que se tomaron en cuenta estas cifras para las estaciones de monitoreo y se configuraron las funciones de alta, media y baja para el concepto de temperatura propuesto.

Con la relación a la presencia de humo en habitaciones estructurales, los niveles promedio en hogares sin fugas de gas varían de 0.5 a 5 partes por millón (ppm), teniendo presente que los niveles cerca de estufas de gas adecuadamente ajustadas suelen ser de 5 a 15 ppm y las que están cerca de estufas mal ajustadas pueden ser de 30 ppm o más (Environmental Protection Agency, 2017), de lo que, análogamente a la formulación de las funciones para el concepto de temperatura, estas cifras hicieron parte de la configuración de las funciones de alto, medio y bajo para el concepto de humo propuesto.

Finalmente, dada la definición de conceptos para el antecedente de las reglas lógicas, es necesario definir un consecuente que interprete el comportamiento del modelo para cada combinación de reglas (Duţu et al., 2018), y la propuesta para esto es extender las nociones de riesgo crítico, promedio y estabilizado hacia funciones de pertenencia que arrojen el resultado de asociar los valores capturados para medición de temperatura y presencia de humo, de forma que se pueda estructurar en un mensaje entendido como una acción recomendada desde el monitoreo.

4. Caso de Estudio

Acorde a la disposición física y lógica, se recolectan un total de 80 mediciones representativas para el funcionamiento del modelo, y así, bajo la implementación basada en lógica difusa, se estimen valores inferenciales que validen la ocurrencia de escenarios de incendio a partir de la base de conocimiento y los nuevos valores registrados. La siguiente ecuación representa la actuación del compresor basado en relaciones lógicas, para las definiciones y funciones de pertenencia correspondientes con la base de conocimiento, es decir, valores anteriores (X(t-1)), valor actual (X(t)) y valor futuro o de predicción (X(t+1)), logrando que el modelo de inferencia pronostique un comportamiento dentro de la habitación.

$$X(t+1) = K_1 min\left(1-\mu_{X(t-1)}, 1-\mu_{X(t)}\right) + K_2 min\left(\mu_{X(t-1)}, 1-\mu_{X(t)}\right)$$
$$+K_3 min\left(1-\mu_{X(t-1)}, \mu_{X(t)}\right) + K_4 min\left(\mu_{X(t-1)}, \mu_{X(t)}\right)$$

El contraste entre los valores existentes (X(t)) y los valores futuros (X(t+1)) de presencia de humo que se evidencia en la Figura 4, manifiesta actuaciones semejantes, atribuyendo un alto grado de veracidad a los siguientes registros que se obtengan, en función de establecer la situación actual de la habitación en cuanto a dicho factor.

De manera idéntica se refleja en la Figura 4, la tendencia de los valores existentes y valores futuros de la temperatura dentro de la habitación estructural que se está estudiando.

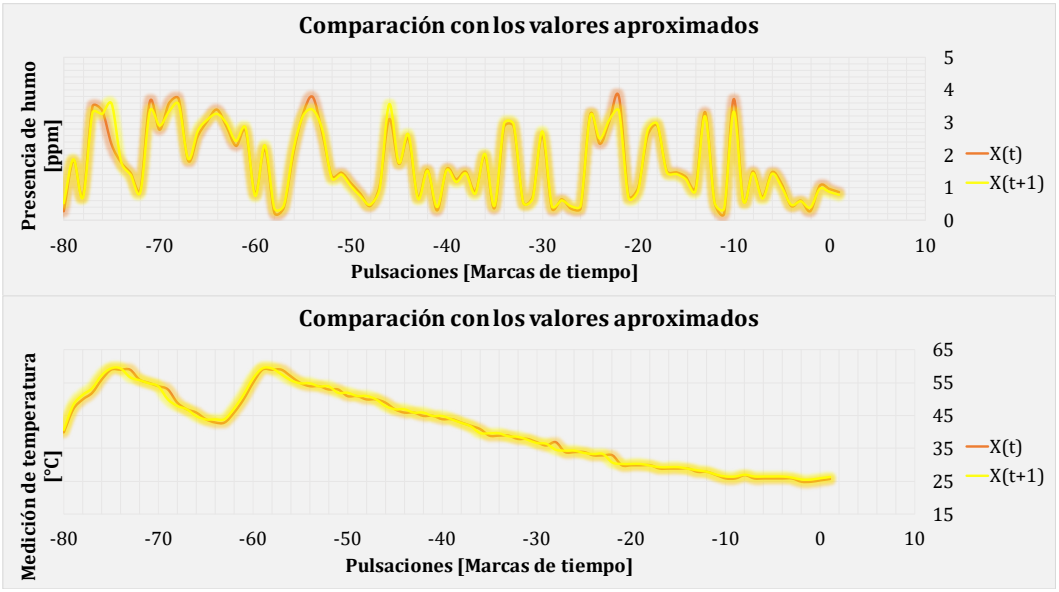


Figura 4 – Comportamiento de los valores predichos

De este modo, el siguiente registro capturado y que representa la temperatura y cantidad de humo presente en la habitación, se procesa y analiza dentro del modelo predictor basado en lógica difusa, revelando los valores pronosticados conforme el comportamiento conocido. Así pues, en el caso de estudio arroja una temperatura 25,84ºC y una cantidad de 0,869771 ppm de presencia de humo, concluyendo que el estado de la habitación es normal y que cuenta con un indicador muy bajo para un escenario de incendio.

5. Conclusiones

La creación de un prototipo ubicado en habitaciones estructurales, que tiene inmersos diferentes componentes tanto físicos como lógicos, en donde se concibe la obtención y transmisión de registros presentes en el ambiente y posterior estudio a través de algoritmos que representan un modelo basado en lógica difusa, dotan a la red de redes (Internet) de una capacidad de inmersión y propagación a gran escala, que resulta en la interacción de: dispositivos electrónicos; aplicaciones destinadas al consumo y análisis de datos capturados en determinada zona; y el sistema involucrado e interesado por ese servicio en particular (Zanella, Bui, Castellani, Vangelista & Zorzi, 2014).

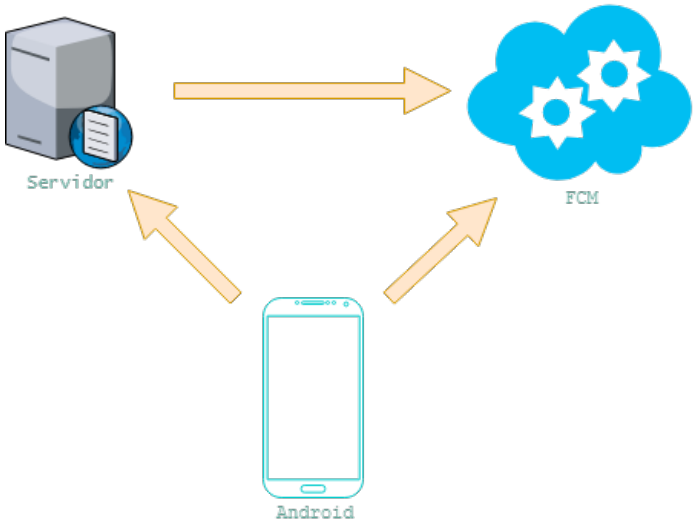


Figura 5 – Visualización de la comunicación en dispositivos móviles

Este modelo mejora la calidad de su inferencia en términos del error cuadrático y su modificación en tiempo de ejecución, para ser minimizado a través de un algoritmo de optimización (como GRG Nonlinear (Cheusheva, 2018)), y que su configuración de valores para la predicción se ajuste más precisamente a la función original. La demostración de esto, es que con la predicción de la presencia de humo se obtuvo un error cuadrático de 3,01478, y con la medición de la temperatura se obtuvo un error cuadrático de 58,40689, lo que proporciona una mejor similitud con el comportamiento de los registros para generar la predicción, y hace que el proceso de inferencia se base en datos confiables.

Esta investigación se complementará con la extensión y mejora de las características de ensamblaje físico, utilizando sensores con un margen de error menor que los dispositivos utilizados, así como la posibilidad de trabajar con más variables ambientales (como partículas de monóxido de carbono o porcentaje de humedad) y que el estudio ambiental brinde una descripción más detallada del estado de la habitación en relación con los incendios estructurales. Uno de los trabajos futuros es extender el modelo de inferencia a otros contextos, donde no solo se interpretan conceptos de variables lingüísticas, sino también un consumo de datos abiertos vinculados (Gaona-García, Herrera-Cubides, Alonso-Echeverri, Riaño-Vargas & Gómez-Acosta, 2018) que permita la interacción con otras estaciones en estudio de incendios estructurales, y de esta manera mejorar los valores de configuración con un mayor margen de captura de datos.

Si bien todas las propuestas estudiadas comparten el objetivo de detectar y comunicar oportunamente la ocurrencia de posible incendio estructural, no todas las estrategias tienen una construcción o funcionamiento similar. Por ejemplo (Kyung & Hong-Hee, 2004) propone un esquema específico para comunicar la información capturada por el dispositivo, funcionalidad que si bien genera un estándar podría ser un problema al

querer extender el proyecto e implementar otras tecnologías, dificultad que es suavizada y manejada al optar por la transferencia de datos por medio del sistema REST y sus operaciones, tal como se visualiza en la Figura 2.

Del mismo modo, una transferencia de datos vía Global System for Mobile communications (GSM) tal como proponen (Ransing & Rajput, 2015), cumple su labor, pero si existe la necesidad de compartir y distribuir información mucho más estructurada y acompañada de elementos tales como, gráficas, informes, historiales, estadísticas, predicciones y muchas otras, la estrategia GSM se queda corta y limitada para lograr ese objetivo. Así pues, construir e implementar un modelo de comunicación IoT tal como se muestra en la figura 5, representa un camino adecuado y correcto para explotar al máximo el uso de los datos desde su recolección y tratamiento, hasta su transmisión vía dispositivos móviles y de manera paralela, se abarcan temáticas a nivel de software de experiencia de usuario que mejoran de manera impactante la propuesta de investigación.

Finalmente, el modelo propuesto, se hace fuerte y robusto en función del uso de la lógica difusa como estrategia clave en el análisis y evaluación de eventos y la inferencia de los mismos, gracias al conocimiento que él va adquiriendo y perfeccionando en medio de su funcionamiento.

Referencias

Asociación Nacional de Sistemas de Rociadores Automáticos Contra Incendios [ANRACI] (2018). Misión y Visión. [En línea]. Recuperado el 30 de mayo de 2018 en el sitio web: http://anraci.org/anraci-colombia-laasociacion/mision-vision/.

Cábová, K., Blesák, L., & Wald, F. (2016). Advanced prediction methods in structural fire safety engineering. In 2016 Smart Cities Symposium Prague (SCSP), (pp. 1–5).

Cheusheva, Svetlana (2018). How to use solver in excel with examples. [En línea]. Recuperado el 28 de abril de 2018 en el sitio web: https://www.ablebits.com/office-addins-blog/2016/06/22/how-touse-solver-in-excel-with-examples/.

Chow, T. C.-L. & Ma, W. W. K. (2017). Do we really know what people are using their smartphone for? In 2017 International Symposium on Educational Technology (ISET), (pp. 34–38).

Duţu, L., Mauris, G., & Bolon, P. (2018). A fast and accurate rule-base generation method for Mamdani fuzzy systems. IEEE Transactions on Fuzzy Systems, 26 (2), 715–733.

Environmental Protection Agency (2017). Carbon monoxide's impact on indoor air quality. [En línea]. Recuperado el 28 de abril de 2018 en el sitio web: https://www.epa.gov/indoor-air-quality-iaq/carbon-monoxidesimpact-indoor-air-quality.

Gaona-García, P. A., Herrera-Cubides, J. F., Alonso-Echeverri, J. I., RiañoVargas, K. A., & Gómez-Acosta, A. C. (2018). A fuzzy logic system to evaluate levels of trust on linked open data resources. Revista Facultad de Ingeniería, (86), 40–53.

Huang, G.-B., Liang, N.-Y., Rong, H.-J., Saratchandran, P., & Sundararajan, N. (2005). On-line sequential extreme learning machine. In IASTED International Conference on Computational Intelligence, (pp. 232–237).

Jiang, L., Liu, Y., Li, Y., & Ma, W. (2017). Fire prediction based on online sequence extreme learning machine. In 2017 7th IEEE International Conference on Electronics Information and Emergency Communication (ICEIEC), (pp. 568–571).

Jose, A. C. & Malekian, R. (2017). Improving smart home security: Integrating logical sensing into smart home. IEEE Sensors Journal, 17, 4269–4286.

Kyung, C. L. & Hong-Hee, L. (2004). Network-based fire-detection system via controller area network for smart home automation. IEEE Transactions on Consumer Electronics, 50 (4), 1093–1100.

Marta (2016). ¿cuál es la temperatura adecuada de una casa? [En línea]. Recuperado el 28 de abril de 2018 en el sitio web: http://www.casasdigitales.com/la-temperatura-adecuada-una-casa/.

MikroElektronika (2014). Air quality click. [En línea]. Recuperado el 18 de septiembre de 2018 en el sitio web: https://download.mikroe.com/documents/add-on-boards/click/air-quality/air-quality-click-manual-v100.pdf.

Ransing, R. S. & Rajput, M. (2015). Smart home for elderly care, based on wireless sensor network. In 2015 International Conference on Nascent Technologies in the Engineering Field (ICNTE), (pp. 1–5).

Riaño Vargas, K. A., Salcedo Parra, O., & Correa, L. (2018). Temperature monitoring and forecast system in remote areas with 4.0G LTE mobile technologies. In The Fourth International Symposium on Ubiquitous Networking (UNet 2018) (UNet'18), Hammamet, Tunisia.

Satoshi, P. (2017). Raspberry Pi and Dht11 humidity Sensor. [En línea]. Recuperado el 18 de septiembre de 2018 en el sitio web: https://medium.com/dyi-electronics/raspberry-pi-and-dht11-humidity-sensor-cccf6b3072ad.

Topaloğlu, F. & Pehlıvan, H. (2018). Comparison of Mamdani type and Sugeno type fuzzy inference systems in wind power plant installations. In 2018 6th International Symposium on Digital Forensic and Security (ISDFS), (pp. 1–4).

Yoddumnern, A., Yooyativong, T., & Chaisricharoen, R. (2017). The wifi multisensor network for fire detection mechanism using fuzzy logic with ifttt process based on cloud. In 2017 14th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), (pp. 785–789).

Zanella, A., Bui, N., Castellani, A., Vangelista, L., & Zorzi, M. (2014). Internet of things for smart cities. IEEE Internet of Things Journal, 1 (1), 22–32.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 13/10/2018
Aceitação/Acceptance: 24/12/2018

Implementación de un Laboratorio Remoto (LR), como recurso de apoyo en un sistema de Educación a Distancia

Jairo Antonio Contreras-Mendieta[1], Celia Paola Sarango-Lapo [1], Dunia Inés Jara-Roa[1], Martha Vanessa Agila-Palacios[1].

jacontreras@utpl.edu.ec, cpsarango@utpl.edu.ec, dijara@utpl.edu.ec, mvagila@utpl.edu.ec.

[1] Universidad Técnica Particular de Loja, Dirección de Tecnologías para la Educación, 1101608, Loja Ecuador.

Pages: 923–935

Resumen: Actualmente el uso y apropiación de las Tecnologías de la Información y de la Comunicación (TIC) en la educación superior, ha originado el surgimiento de nuevas tendencias de aprendizaje que fomentan en los estudiantes la práctica y experimentación confrontando así la teoría con la práctica. En el presente trabajo, se expone la metodología que se utilizó para la implementación de un LR en modalidad Abierta y a Distancia, con el fin de proporcionar a los estudiantes una alternativa para realizar sus prácticas experimentales, al cual se accede remotamente desde cualquier parte. Así mismo, se presentan los resultados obtenidos respecto a la percepción del estudiante frente a la funcionalidad del sistema y a la estrategia utilizada para la ejecución de la práctica del péndulo simple.

Palabras-clave: laboratorio remoto; educación a distancia; estudiantes; física.

Implementation of a Remote Laboratory (LR), as a support resource in an education system at distance

Abstract: Currently the use and appropriation of Information and Communication Technologies (ICT) in higher education has led to the emergence of new learning trends that encourage students to practice and experiment thus confronting theory with practice. In the present work, the methodology that was used for the implementation of a Remote Laboratory (LR) in open and distance programs is exposed, in order to provide students with an alternative to perform their experimental practices, which is accessed remotely from anywhere. Likewise, the results obtained are presented regarding the student's perception of the functionality of the system and the strategy used to execute the simple pendulum practice.

Keywords*:* remote laboratory; distance education programs; students; physics.

1. Introducción

La educación a distancia se presenta como la solución idónea para un sin número de estudiantes que exigen y necesitan disponer de sistemas de enseñanza más flexibles,

accesibles y adaptativos sin limitaciones de tiempo ni espacio. Actualmente el modelo educativo de educación a distancia con ayuda de las TIC presentan alternativas para mejorar la interacción (profesor-alumno, alumno-contenido, alumno-alumno) y el proceso de enseñanza-aprendizaje (Dormido, Sánchez y Morilla, 2016). Es por esto, que los docentes deben incorporar en sus prácticas educativas, estrategias innovadoras tomando en cuenta las TIC como elementos integrales del contexto educativo (González, 2008).

Una de las estrategias innovadoras más importantes que se realiza en la formación de los estudiantes universitarios es la integración de prácticas de laboratorio con el fin de desarrollar hábitos y habilidades profesionales en el área de estudio de cada uno de ellos (Chanfón, Gómez y Borroto, 2016). Además, dado que las prácticas de laboratorio permiten llevar a cabo actividades experimentales se han constituido en uno los aspectos clave en el proceso de enseñanza-aprendizaje (López y Tamayo, 2012; Cardoso et al., 2018).

Gracias a la integración de las TIC en el aula, la práctica en el laboratorio tradicional propone alternativas para la experimentación a través de la incorporación de LR y virtuales (Lima, Viegas, Alves y García-Peñalvo, 2017), es por ello que, muchas universidades realizan experimentos reales a distancia por medio de los LR (Matarrita y Concari, 2015).

Un LR según Zamora (2012) es un sistema basado en un entorno cliente – servidor, donde los clientes desde un sitio remoto, contactan a un servidor (equipo de cómputo, software de acceso, dispositivos a manipular), a través de distintos medios de interconexión (Intranet, Internet), además, con el apoyo de sensores e instrumentación, el usuario realiza una interactividad con equipamientos reales disponibles en un laboratorio (Grisolio, Zanetti, Bellotti y Castiñeiras, 2018). Mientras que, para Ibarra y Rodríguez (2013) los LR son recursos tele operados, es decir el usuario por medio de una interfaz de presentación que se conecta mediante una red, puede controlar equipos reales en ubicaciones remotas, además de recibir mediciones y visualizar en tiempo real lo que ocurre. (López & Tamayo, 2012).

De acuerdo a Sartorius et al. (2005, citado por Santana, Ferre, Hernández, Aracil, Rodríguez y Pinto , 2010) las características que son comunes en varios LR a distancia son:

- Disponibilidad del sistema las 24 horas del día, con su adecuada autoprotección,
- Debe ser accedido desde cualquier parte del mundo
- Los usuarios del sistema deben tener los conocimientos básicos de la disciplina objeto de prácticas
- El sistema permita atender múltiples pedidos de forma paralela administrando de forma centralizada dispositivos similares que se encuentren geográficamente separados pero unidos por redes de área extensa (WAN).
- Contar con un sistema de reservas que permite la correcta utilización de las prácticas.

Para solventar problemas como el costo y tiempo de desplazamiento para asistir a un laboratorio presencial, se presentan las siguientes ventajas que tienen los LR (Matarrita, Elizondo y Villalobos, 2016):

- El estudiante o usuario trabaja con instrumentos reales, no con programas de simulación.

- Las prácticas con equipos reales tienen menor margen de error en cuanto a la obtención de resultados.
- La interacción directa con el equipamiento del laboratorio aporta una experiencia difícil de igualar dado que, además de las variables medidas, los alumnos perciben los experimentos con los cinco sentidos (vista, tacto, oído, olfato e incluso, a veces, gusto).
- Los estudiantes pueden acceder al laboratorio desde cualquier lugar sin restricciones de horarios.

En la tabla 1 se presentan algunas propuestas de LR llevadas a cabo en algunos países de Latinoamérica y Europa (Matarrita y Concari (2014); García, Gustavsson, Hernánde, Orduña, Angulo y Ruiz de Garibay (2016)).

Proyecto de LR	**País**	**Número de prácticas de física**	**Nivel educativo**	**Áreas temáticas**	**Apoyos educativos**
WebLab-Deusto	España	5	Universitario	Mecánica. Circuitos eléctricos.	Guías de usuario. Manuales técnicos. Videos explicativos. Acceso con dispositivos móviles.
Grupo Galileo	Argentina	3	Universitario	Mecánica. Circuitos eléctricos. Magnetismo.	Descripción de experiencias. Guías didácticas.
Physil@b	Colombia	4	Universitario	Mecánica	Un libro integrador entre teoría y experimentación, simulación y experimentación remota.
Laboratorio Remoto FCEIA	Argentina	3	Universitario	Circuitos eléctricos. Circuitos electrónicos. Energía solar	Apoyo con la plataforma educativa. Acceso con dispositivos móviles.
UNEDLabs	España	8	Universitario	Mecánica Ópticas	Manuales de uso y guías didácticas
Proyecto e-Laboratory Project	República Checa	18	Universitario		Cada práctica se incluye la motivación, la fundamentación física, una ilustración del experimento, la actividad a realizar, videos donde se muestra el diseño experimental.
VISIR del BTH	Suecia		Universitario	Electrónicos Analógicos Básicos	
LabsLand	Europa	10	Universitario	Biología, electrónica, física, robótica.	Protocolo de la práctica, documento con la parte teórica, videos

Tabla 1 – Propuestas de laboratorios remotos

Como se observa en la tabla 1 y en coherencia con la investigación de Gravier, Fayolle, Bayard, Ates, y Lardon (2008) los campos disciplinares de la Ingeniería, como la Electrónica, la Robótica, la Automática y la Física por ser ciencias de carácter experimental son los que cuentan con mayor número de desarrollos en LR. Los laboratorios se constituyen en espacios esenciales en la enseñanza de estas disciplinas, en la educación a distancia el tiempo para realizar prácticas de laboratorio queda reducido a las instancias presenciales, por lo que se deben buscar opciones para fortalecer el trabajo experimental.

Con lo antes expuesto el presente trabajo tiene como objetivo describir aspectos técnicos y pedagógicos considerados para la implementación de un LR como recurso de apoyo en un sistema de educación a distancia, de una universidad ubicada al sur del Ecuador integrada por 86 centros universitarios distribuidos en el país y tres centros universitarios internacionales: New York, Madrid y Roma

Así mismo se presentan resultados referentes al diseño del sistema y actividades planteadas para la ejecución de la práctica.

2. Metodología de implementación de un LR

De acuerdo a Zamora (2011) antes de implementar un LR es importante llevar un proceso de planificación, con elementos claves para su diseño como:

a. Análisis y evaluación de las características, técnicas como pedagógicas que den como resultado laboratorios óptimos, pertinentes.
b. Relación costo-beneficio para que la inversión no se convierta en un gasto oneroso, con problemas de adecuaciones, infraestructura y obsolescencia de equipos.
c. Revisar diferentes alternativas de implementación para los LR con el fin de lograr que estos cumplan con los objetivos académicos propuestos.

Diferentes autores tomaron como base para la implementación de un LR metodologías de desarrollo de software las que a continuación se detallan:

a. Arguedas (2017), utilizó una metodología basada en RUP (Rational Unified Process), que permite mantener todo el proceso documentado, para facilitar el seguimiento de las diferentes actividades. Esta metodología divide el proceso en cuatro fases: (a) Inicio, (b) Elaboración, (c) Construcción y (d) Transición; en cada una de estas fases debe por lo menos existir una iteración entre cuatro etapas: análisis de requerimientos, diseño, implementación, y pruebas.
b. Arrieta y Piedra (2015), inician un proceso investigativo sobre el diseño de laboratorios virtuales y remotos, donde se identifican los entornos de programación, herramientas de simulación, creación y distribución de páginas Web, herramientas para el cálculo numérico, técnicas de control de procesos, continuando con el análisis del diseño y arquitectura a ser implementadas en el proyecto, y para la implementación se debe conocer la interacción y manejo adecuado de los componentes físicos (hardware), junto con los entornos de simulación y programación (software).

Para la selección de la metodología que permitió la implementación del LR del presente estudio, se tomó como base ciertas etapas que los autores consideraron en sus trabajos, que a continuación se detallan.

3. Desarrollo de la metodología

Tomando como base los estudios expuestos por Argueda (2017) y Arrieta y Piedra (2015), se elaboró la metodología para la implementación del LR que comprende cuatro fases que se muestran en la tabla 2.

Fases	Descripción
Análisis	Recolección de requerimientos e identificación de usuarios.
Diseño	Diseño de prototipos y arquitectura del sistema.
Implementación	Definición de los instrumentos tanto técnicos como pedagógicos.
Pruebas	Pruebas de campo.

Tabla 2 – Fases de la metodología

A continuación se detallan las fases de la metodología aplicada en la implementación del LR.

3.1 Análisis

En esta fase se trabajó conjuntamente con el docente de la asignatura, con el fin de elaborar e implementar prácticas de física que ayuden al fortalecimiento en el proceso de enseñanza aprendizaje, también se tomó en cuenta aquellas prácticas que presentan mayor dificulta para los estudiantes, ya sea en el acceso a laboratorios físicos o en la construcción de las prácticas con sus propios recursos.

En base a esto se definieron los siguientes requerimientos funcionales y no funcionales que debe contar el LR que se presentan en la tabla 3.

Rol	Requerimientos funcionales	Requerimientos no funcionales
Estudiante	Contar con medios que permitan realizar el proceso educativo en horarios fijados en tiempo real las 24 horas del día, los 7 días de la semana. Contar con una comunicación bidireccional y multidireccional que permita al estudiante compartir ideas, actividades grupales y resolver inquietudes con ayuda de docentes o compañeros. Podrá acceder vía web sin necesidad de descargarse un aplicativo en sus equipos.	Los tiempos de respuesta del sistema sean cortos (máximo 5 minutos). Garantizar que los resultados obtenidos de las prácticas sean reales y verídicos. La interfaz de usabilidad del LR, sea intuitiva y fácil de usar. Control y reglas de acceso para los usuarios del sistema
Docente	Permitir al estudiante realizar sus actividades de aprendizaje en un ambiente que estimule la vista, el tacto, en un espacio físico remoto en tiempo real, que logre la inmersión como como si lo realizará presencialmente, asimismo debe brindar la seguridad de que funcionará con altas concurrencias de usuarios. Permitir al docente definir los parámetros de las prácticas como son actividades, teoría, preguntas de análisis y preguntas de cuestionario, en un intervalo de tiempo, aspectos considerados desde el punto de vista pedagógico.	
Técnico	Integrar el LR al Sistema de Gestión de Aprendizajes (LMS) a través de herramientas externas (LTI) para administrar y distribuir las actividades de formación no presencial.	

Tabla 3 – Requerimientos funcionales y no funcionales

3.2. Diseño

En esta fase se inició con la elaboración de prototipos: (a) interfaz del sistema, (b) diseño electrónico de la maqueta, una vez aprobado se procedió a realizar el diseño de la arquitectura del LR:

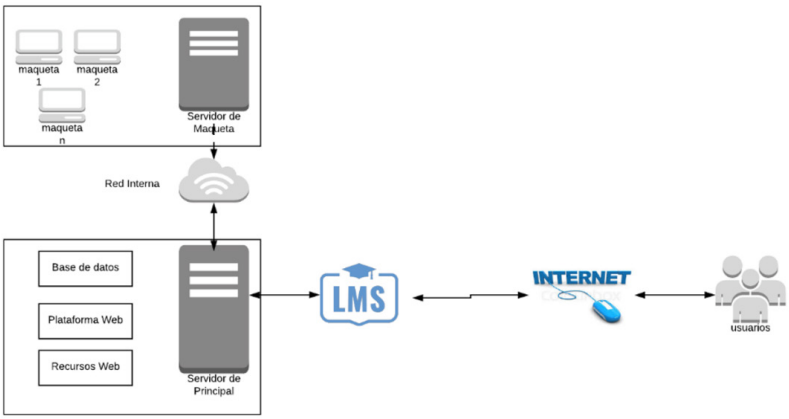


Figura 1 – Arquitectura del LR

Como se indica en la figura 1, la arquitectura software se encuentra divida en dos servidores y en la red interna, cada uno con diferentes funcionalidades y servicios:

Servidor	Funcionalidades	Servicios
Principal	Gestionar los servicios necesarios para proveer al resto de componentes de una comunicación estable. Proporcionar el acceso a los usuarios. Almacenar la información de los distintos usuarios acceden a la plataforma.	Servicio Web (Apache) para montar una aplicación web Servicio de Recursos Web (Node JS) permite levantar los distintos recursos Web (Restful) y comunicación cliente-servidor (Web Sockets), que serán consumidos por la aplicación web y el servidor de maquetas para las distintas peticiones que puedan llegar a realizar. Base de Datos (Mariadb) para levantar el Sistema Gestor de Base de Datos Relacional, el que permite almacenar la información que se transmite por medio del servicio de recursos web, esta información contiene registros con respecto a usuarios, perfiles, horarios, experimentos, resultados e informes.
Maquetas	Se encarga de levantar los servicios que permiten la comunicación entre las maquetas y el servidor principal.	Servicio de Control (LabView) para enviar y recibir información de y a las maquetas. Servicio de Comunicación (Java) permite establecer una comunicación cliente-servidor. Servicio de Streaming (Java) captura la transmisión de la cámara y la envía al servidor principal que puede ser compartida a todos los usuarios que estén realizando la práctica y almacenar el respectivo video.
Red interna	Establecer una correcta y segura comunicación entre los recursos que conforman el sistema del LR	

Tabla 4 – Descripción de los elementos que conforman la arquitectura.

3.3. Implementación

El sistema de LR fue concebido para 10 prácticas de física y para la realización de las distintas prácticas los estudiantes podrán conectarse vía web previo a una reserva.

En la implementación del LR se utilizó un modelo tecno-pedagógico con el propósito de definir un proceso dinámico y robusto que contemple tanto la parte técnica como la parte pedagogía, ya que el sistema de LR no se podría utilizar sin la ayuda pedagógica del docente.

*3.3.1. **Implementación técnica:** la parte técnica está conformada por software y hardware.*

Hardware: a continuación se detalla las características recomendadas para los servidores utilizados para el correcto funcionamiento del sistema.

Características	**Descripción**	
	Servidor Principal:	**Servidor de Maquetas**
Sistema operativo	CentOS 7 minimal (64 bits)	Windows 10 de 64 bits
Memoria RAM	8GB3	8GB
Disco Duro	1TB 24/7	1TB
Procesador	Intel Cori 7	Intel Cori 7
IP(Protocolo de Internet)	Publica	Publica

Tabla 5 – Características físicas de servidores

Software: para la implementación de la solución se utilizó el modelo cliente-servidor, en la figura 2 se presenta dicho modelo.

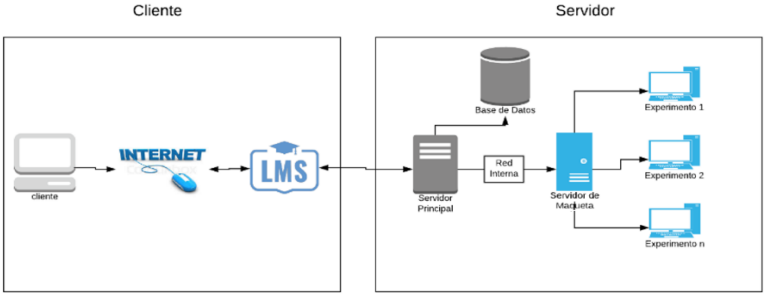


Figura 2 – Arquitectura cliente-servidor del sistema

Cliente: el estudiante por medio de un equipo Pc o portátil, mediante Internet accede al LMS específicamente a la actividad “herramientas externas” la que está configurada para acceder al servidor principal del sistema de LR.

Servidor: Los elementos del LR se encuentran en una red privada tras un cortafuego, para cada experimento o práctica existe una maqueta, una cámara y un servidor, el cual, a través de una tarjeta de adquisición de datos, permite llevar a cabo el experimento y el almacenamiento de los datos. El sistema se completa con un servidor principal de prácticas y un servidor de bases de datos, el servidor principal de prácticas se encarga de acceder al servidor de maquetas para gestionar los recursos necesarios para realizar la práctica (manuales, videos tutoriales, protocolo de la práctica, etc.) y de hacer cumplir su protocolo (leer la documentación, presentar cálculos previos, pasar un test, etc.); también proporciona la interfaz para preparar el experimento y analizar los datos. En el servidor de bases de datos se almacena datos sobre los alumnos (información de identificación, reservas de prácticas), sobre sus prácticas (datos de la práctica) y estadísticos (número de accesos para completar una práctica, distribución temporal de los acceso).

3.3.2. Implementación pedagógica: Con el fin de lograr que el aprendizaje teórico sea práctico se procedió a implementar el aprendizaje experiencial, para esto se adoptó el modelo de tres fases en base a lo que señala (Dewey, 1997) que contempla la ejecución secuencial cíclica de tres fases tal como lo indica la figura 3, se inicia con la experimentación hasta llegar a la evaluación, las cuales se explican a continuación.

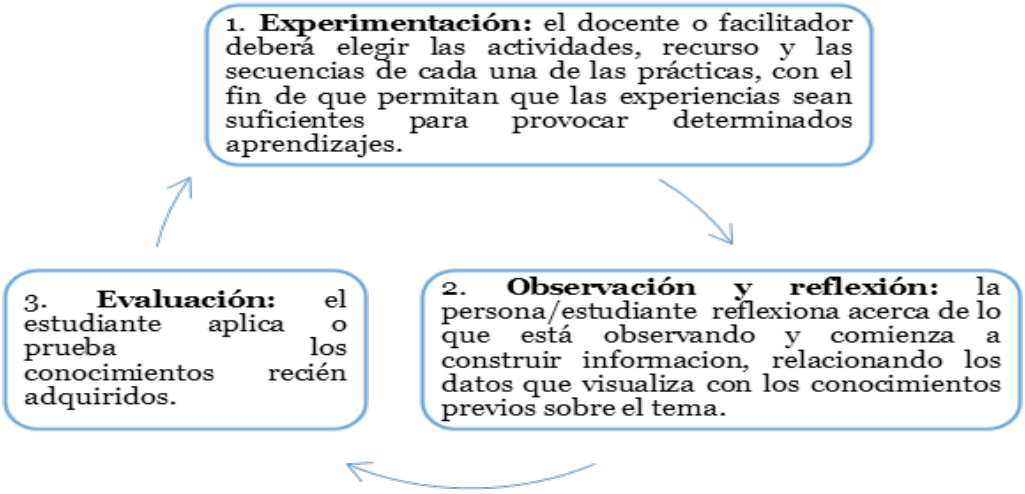


Figura 3 – Adaptación al modelo de tres fases basado en Dewey´s.

Además manifestar que el sistema soporta prácticas grupales, por lo que se implementó un sistema de chat, para la comunicación de los integrantes del grupo durante la realización de la práctica.

Conjuntamente con docentes de la asignatura de física se trabajó en los elementos que deberían tener el protocolo de la práctica y la forma de llevar a cabo el aprendizaje experiencial.

3.4. Pruebas

Las pruebas sobre las que se hace énfasis son: las de conexión, de desempeño y de aceptación.

Para las pruebas de conexión se trabajó bajo tres escenarios, primeramente se probaron diferentes alternativas de conexión al servidor principal del LR, de la cual se obtuvo, a LTI como la alternativa idónea para esta conexión. La segunda prueba involucró, la conexión del servidor principal con el servidor de las maquetas, para lo cual se realizó varios testeos hasta que se pudo acceder a la interfaz gráfica de cada una de las prácticas. Y por último se probó todo el flujo de conexión, desde que el usuario accede a través del LMS, hasta tener respuesta de la maqueta (parte electrónica).

Para las pruebas de desempeño se trabajó en un ambiente de pruebas creado en el LMS de la institución de educación superior, donde se revisó periódicamente la interfaz de los módulos que conforman el sistema para luego probar la funcionalidad de cada uno de estos. Una vez comprobado el funcionamiento de cada uno de los módulos se realizó las pruebas de manera integral al sistema.

Las pruebas de aceptación se enmarcaron en la revisión de la arquitectura, documentación y experimentación en el que estuvo presentes el equipo interdisciplinar conformado por técnicos y pedagogos.

4. Resultados

A fin de medir el nivel de satisfacción de estudiantes sobre el uso del LR, se elaboró un cuestionario utilizando dos dimensiones, funcionalidad (implementación técnica) y ejecución (implementación pedagógica); la funcionalidad está compuesta por cinco ítems y la ejecución por cuatro ítems bajo la escala de Likert, donde 1 significa muy en desacuerdo y 4 muy de acuerdo, con un nivel de confiabilidad de Alpha de Cronbach = 0.95. Los ítems evaluados se muestran en la tabla 6

El cuestionario se aplicó a estudiantes que realizaron la práctica del Péndulo Simple en el año 2017. De los 70 estudiantes convocados, 52 (74%) realizaron la experimentación. Culminada la práctica se procedió a aplicar el cuestionario que fue respondido por 25 (48%) estudiantes. Los resultados se muestran en la tabla 6.

A fin de simplificar el análisis de los datos de la encuesta, se procedió a combinar las cuatro categorías de las respuestas, los datos de Muy de acuerdo y De acuerdo se agruparon como De acuerdo, y los datos de Muy en desacuerdo y En desacuerdo se agruparon como Desacuerdo.

En cuanto a la evaluación de la implementación técnica del LR (figura 4), se puede evidenciar que la mayoría de ítems evaluados tienen una alta aceptación ya que están sobre el 71%, a excepción del ítem relacionado al tiempo de acceso ya que casi la mitad de ellos indica que el tiempo de acceso fue más de 3 minutos, estos porcentajes posiblemente pueden estar relacionados a la conectividad del Internet de los usuarios del sistema, ya que la estabilidad en la conexión es diferente dependiendo del lugar en donde se encuentren.

Tema	Pregunta
Funcionalidad	1. El sistema de LR tiene un sistema intuitivo de reservas para realizar la práctica.
	2. Los botones de ayuda del sistema de LR le orientaron para el desarrollo de las actividades de la práctica.
	3. Las alertas del sistema del LR le ayudaron a conocer el estado de las actividades que realizaba.
	4. La interfaz del sistema de LR es intuitiva y de fácil navegación
	5. El tiempo de acceso al sistema de LR fue adecuado (no se tuvo que esperar más de tres minutos)
Ejecución	6. Los productos de aprendizaje solicitados en las actividades ejecutadas en el LR le permitieron relacionar la teoría aprendida con la práctica.
	7. Las instrucciones diseñadas para realizar las actividades propias de la práctica de LR le fueron de fácil comprensión.
	8. Las prácticas del LR le permitieron responder de manera adecuada las preguntas de análisis
	9. Las preguntas del cuestionario le ayudaron a alcanzar el resultado de aprendizaje .

Tabla 6 – Tabla de resultados

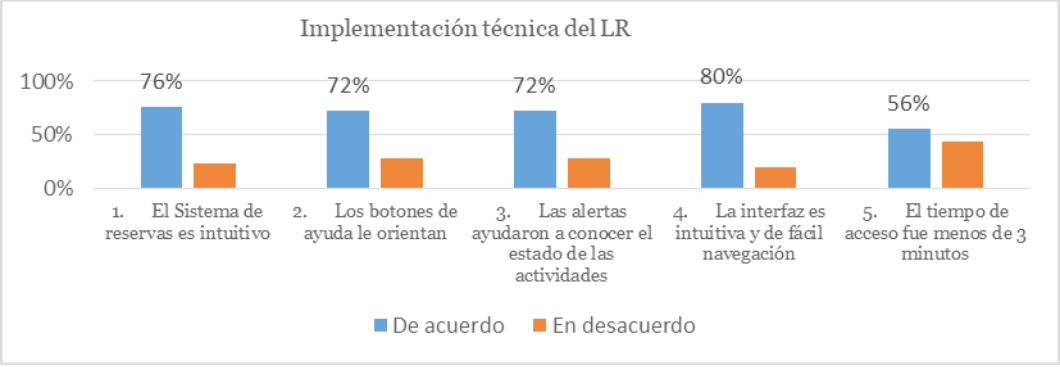


Figura 4 – Implementación técnica del LR

Referente a la ejecución de la práctica (figura 5) en promedio el 82% de estudiantes están de acuerdo ya que consideran que la parte pedagógica contribuyó de manera positiva a su proceso de aprendizaje, no obstante, en promedio el 18% de estudiantes están en desacuerdo lo que nos invita a continuar mejorando en las estrategias de enseñanza.

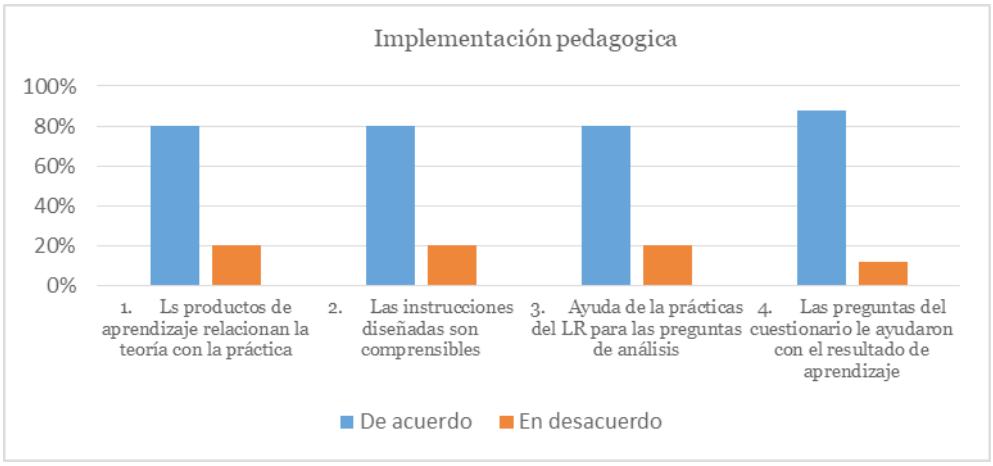


Figura 5 – Implementación pedagógica del LR

5. Conclusiones

Para la implementación de un sistema de LR, es indispensable utilizar una metodología que permita analizar, diseñar, implementar y probar el proceso de desarrollo (Zamora, 2011).

En sintonía con Arguedas (2017) para la implementación del LR es importante vincular la parte técnica con la parte pedagógica a fin de estructurar sistemas que faciliten la experimentación del estudiante y asegurar el logro de resultados de aprendizaje.

La calidad de respuesta cuando se utiliza un LR vía web está directamente relacionado con el ancho de banda del servicio de Internet, para poder controlar equipos reales en ubicaciones remotas (Ibarra y Rodríguez, 2013 ; Grisolio, Zanetti, Bellotti y Castiñeiras, 2018).

La aplicación de LR permite a los estudiantes la ejecución de prácticas experimentales desde diferentes lugares en cualquier momento, por lo que es importante que estos sistemas se diseñen para estar activos las 24/7 (Santana et al., 2005).

El impacto económico que representa la implementación de un LR puede ser mitigado con la simulación de LR, de tal manera que las actualizaciones o ampliaciones sean fáciles de implementar sin generar mayores costos en contraste con lo señalado por Matarrita, Elizondo y Villalobos (2016).

Referencias

Arguedas, C. A. (2017). Diseño y desarrollo de un Laboratorio Remoto para la enseñanza de la física en la UNED de Costa Rica, 1–6.

Arrieta, R. A. y Piedra, O. A. (2015). Diseño e implementacion de un laboratorio virtual y remoto para el desarrollo de prácticas de control de procesos., 207.

Cardoso, M., de Castro, A. V., & Rocha, A. (2018). Integration of virtual programming lab in a process of teaching programming EduScrum based. In *2018 13th Iberian Conference on Information Systems and Technologies (CISTI)* (pp. 1-6). IEEE.

Chanfón, J. G., Gómez, M. R. y Borroto, G. (2016). Impacto de la introducción de los

laboratorios virtuales en la educación superior. *Congreso Universidad, ISSN-E: 2306-918*, *5*(5), 45–59. Recuperado de http://www.congresouniversidad.cu/revista/index.php/congresouniversidad/index

Dewey, J. (1997). EXPERIENCE AND EDUCATION : Retrieved July 11, 2018, from https://archive.org/details/ExperienceAndEducation

Dormido, S., Sánchez, J. y Morilla, F. (2016). LABORATORIOS VIRTUALES Y REMOTOS PARA LA PRÁCTICA A DISTANCIA DE LA AUTOMÁTICA.

García, J., Gustavsson, I., Hernández, U., Orduña, P., Angulo, I. y Ruiz de Garibay, J. (2016). LABORATORIO REMOTO PARA ELECTRÓNICA BÁSICA.

González, J. C. (2008). ICT and the transformation of educational practice in the context of knowledge societies. *RUSC. Revista de Universidad Y Sociedad Del Conocimiento*, *5*(2), 1–8. https://doi.org/10.7238/rusc.v5i2.330

Gravier, C., Fayolle, J., Bayard, B., Ates, M., & Lardon, J. (2008). State of the Art About Remote Laboratories Paradigms – Foundations of Ongoing Mutations. *International Journal of Online Engineering (iJOE)*, *4*(1), 19–25. Retrieved from http://www.online-journals.org/index.php/i-joe/article/view/480

Grisolio, M., Zanetti, P., Bellotti, M. A. y Castiñeiras, J. (2018). Laboratorios Remotos para Prácticas de Ingeniería en Carreras de Grado. *La Conferencia Latinoamericana de Informática (CLEI)*, 56–66.

Ibarra, Ó. y Rodríguez, C. (2013). Laboratorios remotos como innovación educativa. Retrieved February 20, 2018, from https://es.slideshare.net/oscardanielibarra/laboratorios-remotos-como-innovacin-educativa

Lima, N., Viegas, C., Alves, G., & García-Peñalvo, F. (2017). A utilização do VISIR como um recurso educativo : uma revisão da literatura, (April).

López, A. y Tamayo, Ó. (2012). Las Prácticas De Laboratorio En La Enseñanza De Las Ciencias Naturales Laboratory Sessions in Natural Sciences Teaching. *Revista Latinoamericana de Estudios Educativos*, *8*(81), 145–166. Recuperado de http://www.redalyc.org/html/1341/134129256008/

Matarrita, C. A. y Concari, S. B. (2015). Hacia un estado del arte de los laboratorios remotos en la enseñanza de la física. *Revista de Enseñanza de La Física*, *27*(2), 133–139. Recuperado de https://revistas.unc.edu.ar/index.php/revistaEF/article/view/12596/12872

Matarrita, C. A., Elizondo, F. U. y Villalobos, M. C. (2016). Laboratorios remotos : Herramientas para fomentar el aprendizaje experimental de la Física en educación a distancia, 2–6.

Santana, I., Ferre, M., Hernández, L., Aracil, R., Rodríguez, Y. y Pinto, E. (2010). Aplicación del Sistema de Laboratorios a Distancia en Asignaturas de Regulación Automática. *Revista Iberoamericana de Automática E Informática Industrial RIAI*, *7*(1), 46–53. https://doi.org/10.1016/S1697-7912(10)70007-8

Zamora, R. (2011). Implementación de un laboratorio remoto (LR) en el área de microprocesadores en la corporacion universitaria de la Costa, CUC, 60.

Zamora, R. (2012). Laboratorios Remotos: Actualidad y tendencias futuras. *Scientia et Technica Año XVII*, *N° 51*(51), 113–118.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 25/09/2018
Aceitação/Acceptance: 15/11/2018

Modelo de serious game para mejorar la atención en niños con trastorno por déficit de atención e hiperactividad (TDAH).

Mayra Albán Taipe[1]

mayra.alban@utc.edu.ec

[1] Unidad Académica de Ciencias de la Ingeniería y Aplicadas, Universidad Técnica de Cotopaxi, 050101, Latacunga, Ecuador.

Pages: 936–946

Resumen: El trastorno por por déficit de atención e hiperactividad (TDAH) es un problema neurológico que implica dificultades de atención y deterioro en el rendimiento académico de los niños, especialmente en edades tempranas. La revisión de la literatura permitió identificar investigaciones relacionadas con los serious game encaminadas a solucionar los problemas que genera el TDAH en los niños en la etapa escolar. Sin embargo en general los trabajos analizados no han considerado la integración de aspectos psicológicos, pedagógicos y terapéuticos que permitan el diseño y construcción del serious game orientado a satisfacer las necesidades de los niños, los terapeutas, padres de familia y docentes. Por tal razón, la presente investigación propone un modelo para el diseño de serious game que permita mejorar la atención de los estudiantes que presentan este problema. Como resultado se establece que los aspectos pedagógicos y terapéuticos son estadísticamente significativos y pueden ser considerados por los diseñadores de juegos educativos en la fase de análisis y diseño del producto, Así como también, se determina la incorporación de elementos de la ingeniería de ingeniería de software y la integración de la realidad aumentada como tecnología emergente para el desarrollo de una herramienta tecnológica de apoyo en los procesos terapéuticos y de enseñanza-aprendizaje de los niños.

Palabras-clave: serious game, trastorno por déficit de atención e hiperactividad, realidad aumentada.

Serious game model to improve attention in children with attention deficit hyperactivity disorder (ADHD).

Abstract: Attention deficit hyperactivity disorder (TDAH) is a neurological problem that involves attention difficulties and deterioration in the student's academic performance, especially at early ages. The review of the literature made it possible to identify investigations related to serious games aimed at solving the problems generated by ADHD in children in the school stage. However, in general the analyzed works have not considered the integration of psychological, pedagogical and therapeutic aspects that allow the design and construction of a serious game oriented to satisfy the needs of children, therapists, parents, and

teachers. For this reason, this research proposes a model for the design of a serious game that allows improving the attention of students who present this problem. As a result it is established that the pedagogical and therapeutic aspects are statistically significant and can be considered by the educational game designers in the analysis and design phase of the product, as well as the incorporation of elements of software engineering and the integration of augmented reality as an emerging technology for the development of a technological tool to support the therapeutic and teaching-learning processes of children.

***Keywords**:* Serious game, attention deficit hyperactivity disorder, augmented reality.

1. Introducción

El déficit de atención e hiperactividad es un problema biológico que afecta en muchas ocasiones a los niños, se expresa en el contexto psicosocial y se relaciona con problemas de rendimiento escolar (De la Peña Olvera, Ortiz, & Pérez, 2010). Los niños que experimentan este trastorno tienen efectos de impulsividad que modifica la conducta del niño, el desarrollo de habilidades sociales y las relaciones del trabajo grupo (Presentación Herrero, Siegenthaler Hierro, Jara Jiménez, & Miranda Casas, 2010).

El TDAH se presenta entre el 3 y 6% de la población escolar que genera impacto sobre la conducta y el rendimiento académico (Villar, 2004). Para De la Peña Olvera et al. (2010) el TDAH se manifiesta en la infancia y prevalece en un 60% en la adultez si no se obtiene apoyo terapéutico psicológico y tratamiento farmacológico adecuado. Mientras que Rodríguez Molinero, et al. (2009) manifiestan que los niños que presentan TDAH corresponden a un porcentaje mayor del 4% de la población escolar.

En el contexto escolar este problema ha sido tratado mediante la incorporación de herramientas e instrumentos que permiten canalizar la aplicación de metodologías de aprendizaje. Su aplicación permite que el niño pueda obtener una educación de calidad, la cual no debe estar imposibilitada por la conducta que asocia el niño con el TDAH. Una de las estrategias propuestas en la revisión de la literatura es la incorporación de terapias psicopedagógicas, neuropsicológicas, cognitivas, conductuales y familiares que permitan la disminución de los síntomas y la posible integración en el ámbito de la educación primaria (De la Peña Olvera, 2000).

Por otro lado, el uso de metodologías activas para el diseño de serious game, es ampliamente utilizado para el desarrollo de competencias de aprendizaje, que permitan al usuario adquirir habilidades y destrezas basadas en el juego para el desarrollo motor, social, afectivo e intelectual de los escolares (Céspedes, 2015). Debido al avance significativo de la tecnología se han incrementado la implementación de serious game en el contexto educativo escolar, con la incorporación de realidad aumentada se incrementa la productividad del juego que permite la visualización de elementos del mundo real combinados con elementos virtuales (Martínez, Olivencia, & Terrón, 2016), lo cual, genera interacción de los niños con escenarios activos en tiempo real para la estimulación del aprendizaje por descubrimiento.

Cabe señalar que un aspecto importante en el diseño de serious game es la conformación de un equipo multidisciplinario formado por: ingenieros de software, desarrolladores,

diseñadores, pedagogos, psicólogos, terapistas, incluyendo al usuario final que en nuestro caso son los maestros, los padres de familia y los niños; esto permite compartir conocimientos, necesidades y objetivos que garanticen un desarrollo del juego significativo y exitoso, estas afirmaciones guardan relación con lo que sostiene Cano, Aslan y Saavedra (Aslan & Balci, 2015; Cano, Arteaga, Collazos, Gonzalez, & Zapata, 2016; Saavedra, Rodríguez, Arteaga, Salgado, & Ordoñez, 2014)

Por otro lado, dado que en la última década ha existido un amplio desarrollo de herramientas tecnológicas que permiten el mejoramiento de los procesos en la educación escolar, se propone un modelo de juegos serios con realidad aumentada para mejorar la atención en niños con TDAH.

Este artículo está organizado en cinco secciones. En la sección dos se describen los antecedentes del proceso de investigación, en la sección tres se presenta la metodología para el proceso de investigación. La discusión de resultados se presenta en la sección cuatro. Y finalmente la sección cinco presenta las conclusiones.

2. Antecedentes

La revisión de la literatura permitió identificar la existencia de varios estudios que relacionan los objetivos pedagógicos como factor importante en el diseño y construcción de un serious game. Etxeberria (2006), citado por Imaz (2012) menciona que para "*promover una cultura de innovación en el ámbito educativo se requiere fomentar el uso de videojuegos seleccionándolos en función de la edad, las capacidades y los valores*".

En el trabajo realizado por López (2016) se señala que el interés del serious game diseñado en base a objetivos pedagógicos no se centra como objetivo principal en la consecución de avances o victorias sino más bien en la resolución de problemas y asimilación del conocimiento que el docente desea impartir. El uso de objetivos, retos y reglas en un serious game determina el orden, los derechos y las responsabilidades del jugador para descubrir información y nuevo conocimiento Graus et.al (2015).

Por otro lado, establecer programas de intervención para contrarrestar esta problemática requiere de un juicio clínico que condicione las características personales del niño, las intervenciones terapéuticas en formato interactivo digital podría ayudar a que el niño generalice los aprendizajes y los aplique en el ámbito educativo, Molina y Martínez (2015).

Los serious game como herramientas terapéuticas han sido utilizados por los profesionales en el contexto educativo, mismo que debe cumplir con el objetivo de terapia de juego muy utilizada para tratar problemas en niños con trastornos por déficit de atención, Enriques (2017).

3. Metodo

3.1. Recopilación de información

Los datos presentados corresponden a los resultados preliminares de un proyecto de investigación sobre diseño de serious game para mejorar la atención en niños con TDAH. Para la el diseño del modelo se construye un data set que contiene información de los

resultados de una encuesta aplicada a expertos en el tema objeto de estudio, la misma que se llevó a cabo entre los meses de diciembre del 2017 a marzo del 2018.

La encuesta estuvo diseñada en GoogleForm y se encontraba organizada en tres secciones. En la primera sección se determinó preguntas relacionadas con la experiencia de los expertos en el tema objeto de estudio. La sección dos relacionada con preguntas sobre los aspectos que influyen negativamente en la atención en niños con TDAH. La sección tres relacionada con aspectos para el diseño de serious game para niños con problemas de atención.

3.2. Análisis de resultados

La metodología aplicada para la construcción del modelo propuesto comprende tres etapas:

1. Integración de los elementos psicológicos, pedagógicos y terapéuticos aplicados para el diseño de serious game.
2. Estadística descriptiva para analizar y caracterizar las variables de ingresos al modelo propuesto.
3. Diseño del modelo del serious game para determinar la estructura y relaciones de consistencia en el proceso de diseño del juego con fines terapéuticos y educativos.

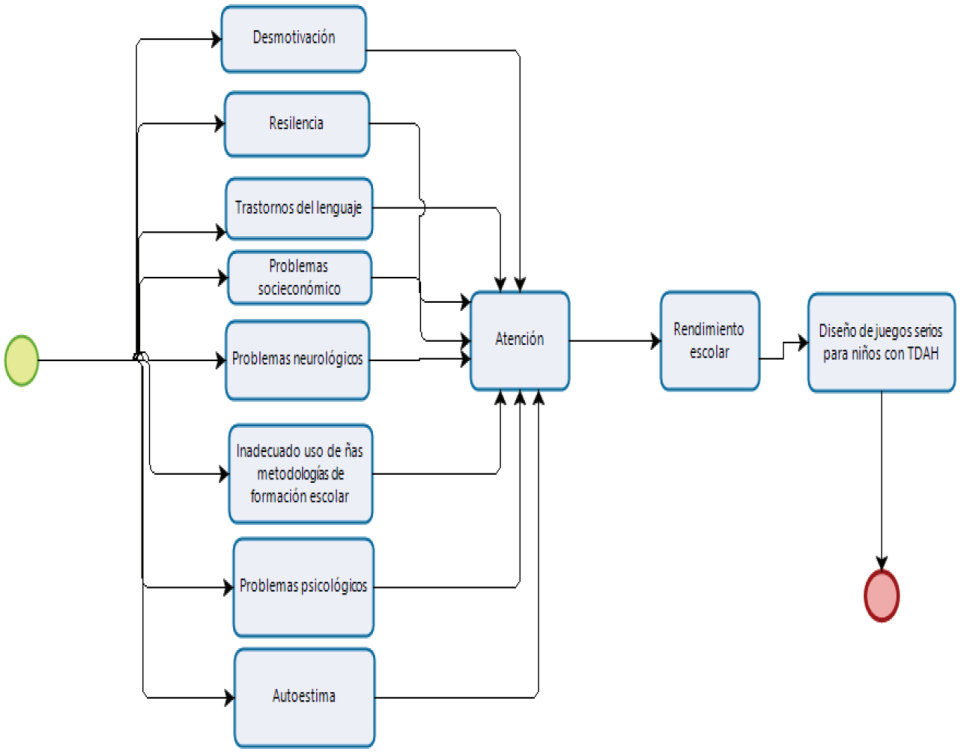


Figura 1 – Aspectos que influyen en la atención en niños con problemas de TDAH

4. Resultados y Discusión

4.1. Integración de aspectos psicológicos, pedagógicos y terapeuticos

Para un adecuado diseño de serious game se debe considerar los factores que inciden negativamente en la atención de los niños en el proceso de aprendizaje y que a su vez influyen en el rendimiento escolar especialmente en aquellos niños que presentan déficit de atención e hiperactividad. La figura 1 presenta los aspectos identificados que afectan negativamente en la atención en niños con el problema descrito.

La incorporación de herramientas tecnológicas en el aprendizaje deben ser canalizados a través del análisis de los causales que impiden la asimilación, comprensión y retención de los contenidos educacionales que conlleven a la consecución adecuada de los objetivos del aprendizaje de los niños.

4.2. Estadistica descriptiva

La Tabla 1 presenta la descripción de los aspectos psicológicos, pedagógicos y terapéuticos que influyen en la atención en niños con problemas de trastorno por déficit de atención por hiperactividad, mientras que la Tabla 2 presenta los resultados de la estadística descriptiva correspondiente a los aspectos identificados en base a la encuesta realizada a los expertos.

ID	Aspecto
V1	Desmotivación
V2	Resilencia
V3	Trastornos de lenguaje
V4	Problemas socioeconómicos
V5	Problemas neurológicos
V6	Inadecuada aplicación de las metodologías de formación escolar
V7	Problemas psicológicos
V8	Autoestima

Tabla 1 – Aspectos que influyen en la atención en niños con TDAH

Atención	Desviación típica	No.
V1	1,123	49
V2	0,407	49
V3	0,487	49
V4	0,474	49
V5	0,606	49
V6	0,000	49
V7	0,624	49
V8	0, 466	49

Tabla 2 – Estadística descriptiva de los aspectos identificados

Por otro lado, se aplica el histograma de residuos para determinar las características generales de la información. La Figura 2 permite evidenciar que existe normalidad en los datos analizados, es decir, se evidencia que provienen de una población distribuida normalmente. Se consideró los valores atípicos como anomalías en los datos, estos pueden tener una gran influencia sobre el ajuste realizado sobre la pendiente y el intercepto.

Es importante realizar el análisis de los valores atípicos debido a que tienden a incrementar la varianza residual, disminuyendo la probabilidad de rechazar la hipótesis nula, esto puede deberse a errores de registro, los cuales se pueden corregir, o también debido a que no todos los valores de la variable dependiente están siendo muestreados de la misma población.

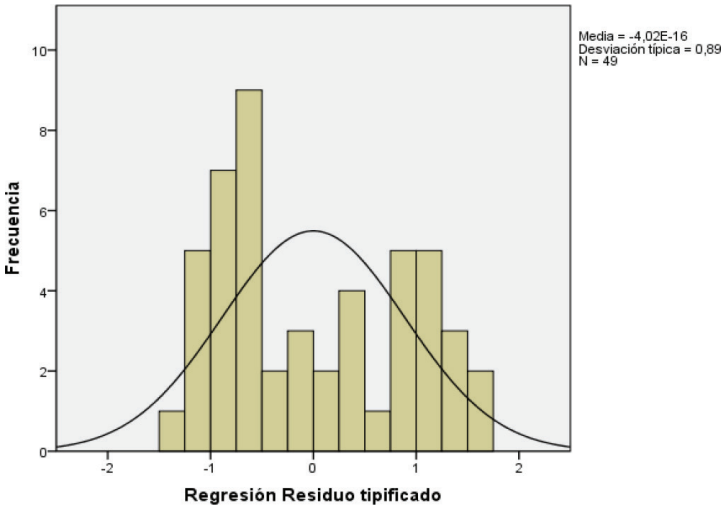


Figura 2 – Histograma del componente atención

Se utiliza la técnica de regresión lineal, para determinar la relación de causalidad entre las variables explicativas con la atención y el rendimiento en niños con TDAH. Sin embargo cuando se utiliza número pequeño de datos, la regresión lineal ofrece menor protección contra las violaciones de supuestos. Los resultados obtenidos muestran que el coeficiente de determinación (R-Cuadrado) fue estadísticamente significativo con una probabilidad (Prob) < 0.05. Los datos resultantes del proceso experimental se presentan en la Tabla 3.

En base a los resultados, se determina que las variables identificadas influyen negativamente en los problemas de atención. Estas a su vez generan dificultades en el rendimiento escolar en niños con trastorno por déficit de atención e hiperactividad. Las variables analizadas fueron consideradas como variables de ingreso al modelo para mejorar el diseño del serious game, que tiene como finalidad la disminución de problemas de aprendizaje por efectos de la inadecuada atención en los niños. Además, en la fase de análisis del proceso de ingeniería del software se debe establecer claramente los requisitos funcionales y no funcionales del producto los cuales deben ser establecidos con el trabajo del equipo multidisciplinar.

Dependent Variable: TDAH **Method: Least Squares** **Included observations: 49**				
ID	**Coef**	**Std. Error**	**t-Statistic**	**Prob.**
V1	0.049141	0.016731	2.937	0.0033
V2	0.043142	0.017348	2.486	0.0129
V4	0.066384	0.017119	3.877	0.0001
V4	0.090469	0.020386	4.437	0.0000
V6	0.046126	0.018144	2.542	0.0111
V7	0.056792	0.016708	3.399	0.0007
V9	0.049018	0.016921	2.896	0.0038
V10	0.087841	0.016681	5.266	0.0000
C	0.306859	0.011523	2.662	0.0000
R-squared	0.002723	Mean dependent var		0.330171
Adjusted R-squared	0.002407	S.D. dependent var		0.470349
S.E. of regression	0.469782	Akaike info criterion		1.327.538
Sum squared resid	6.973.978	Schwarz criterion		1.331.370
Log likelihood	2.096.838	Hannan-Quinn criter.		1.328.913
F-statistic	8.626.830	Durbin-Watson stat		1.927.855
Prob(F-statistic)	0.003336			

Tabla 3 – Estadística descriptiva de los aspectos identificados

4.2. Modelo de serious game para mejorar la atención en niños con TDAH

La Figura 3 presenta el modelo de serious game propuesto que permitirá mejorar la atención en niños con TDAH. El diseño del modelo contempla 3 fases. La primera integra la determinación de los objetivos pedagógicos y terapéuticos, así como también, la integración de las variables que influyen negativamente en la atención en los niños con TDAH. El contexto de la fase 1 permitirá ampliar la concepción de la etapa de análisis para el diseño de un juego educativo desarrollado en base a las necesidades de los niños que presentan este problema.

Como elementos base del modelo se considera el desarrollo del juego con fines educativos que integre la ingeniería del software a través de su amplia game de metodologías de desarrollo y la determinación de su ciclo de vida. El esquema general de los aspectos considerados para el diseño del juego educativo a través de las fases establecidas en el proceso de ingeniería se presenta en la Figura 4.

Por otro lado, las tecnologías emergentes generan perspectivas y son consideradas como tecnologías de gran impacto en el área de la educación en la determinación de procesos terapéuticos y de enseñanza-aprendizaje (Adell & Castañeda, 2012). Dentro de la las tecnologías emergentes utilizadas para este propósito se encuentra la realidad aumentada que complementa la percepción y comprensión del mundo real con entornos virtuales que contempla modelos virtuales gráficos en 2D y 3 D (Basogain, Olabe, Espinosa, Rouèche, & Olabe, 2007).

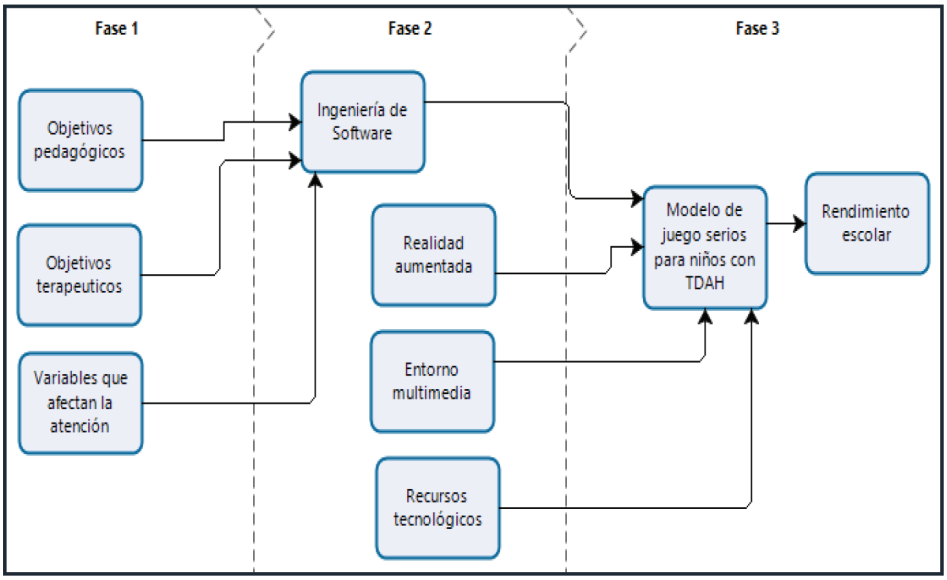


Figura 3 – Modelo de serious game para mejorar la atención en niños con TDAH

Los elementos que determina la realidad aumentada son varios, entre ellos se encuentra la innovación emergente de la educación personalizada que permite la interactividad intuitiva, la rápida asimilación de procedimientos de aprendizaje, aumenta la motivación de los niños, el desarrollo cognitivo divergente y genera resultados positivos sobre procedimientos pedagógicos en los niños (Olivencia & Martínez, 2015). Otra de las bondades de la realidad aumentada es la visualización directa del mundo real que se combina con elementos virtuales (Olivencia & Martínez, 2015). Por tal razón, se considera este aspecto en el modelo para el diseño de serious game que permita mejorar la atención en niños con TDAH.

Para lograr el equilibrio en el diseño de los serious game se debe incorporar procesos de desarrollo que contemplen etapas de especificación de requerimientos y diseño los cuales deben estar enmarcados dentro del diseño centrado en el usuario y que incorpore la usabilidad pedagógica (Spinelli & Massa, 2017).

De igual manera (Pressman & Troya, 1988) en su investigación señalan que se requiere de un proceso adecuado de levantamiento de requerimientos que permita la determinación de la calidad del producto en base a las actividades de recopilación, análisis y verificación de las necesidades (Gil, 2002) para la construcción de un producto exitoso y de calidad. El objetivo fundamental de determinar esta etapa dentro de la construcción del modelo se basa en la creación de prototipos y modelos que aporten al producto final (Ayala, Ramírez, & Ocampo, 2012) que en nuestro caso es el serious game, etapa que será valorada en la fase de evaluación. Los principios de diseño para la construcción del modelo propuesto fueron elaborados en base a la función del proceso de aprendizaje de los niños con problemas de atención e hiperactividad.

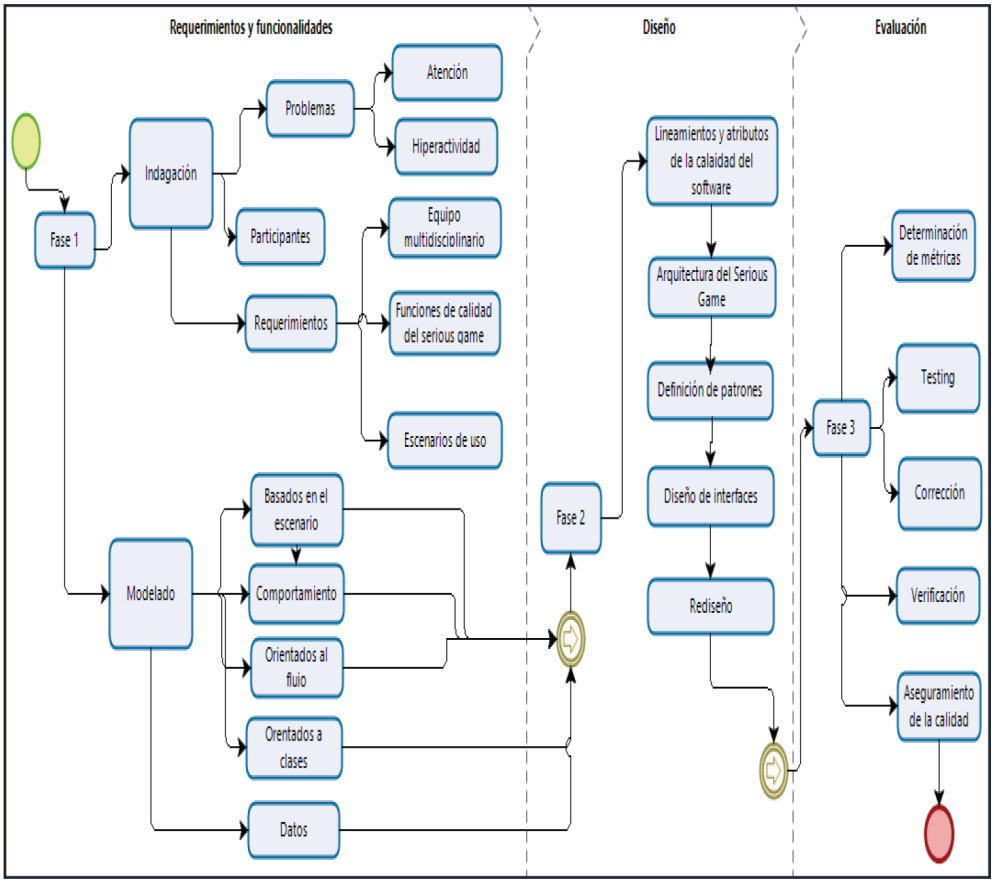


Figura 4 – Esquema general de aspectos considerados en el diseño de Serious Game

Es importante señalar que para la construcción del serious game se considerará los elementos fundamentales para la especificación de requerimientos propuesto por Pressman (1988) y Spinell (2017).

5. Conclusiones

Los serious game han sido utilizados ampliamente en la educación especialmente en el contexto escolar. Su uso como herramientas de apoyo en los procesos terapéuticos y de enseñanza-aprendizaje ha generado impacto significativo que permite un desenvolvimiento motivado en la aplicación de herramientas tecnológicas y tecnologías emergentes que generan el interés de los niños de interactuar en el aula. El modelo de serious game propuesto contempla 8 variables de ingreso que determinan la influencia que tienen estas variables sobre los trastornos de atención e hiperactividad de los niños. Además se establece la incorporación de los objetivos pedagógicos y terapéuticos que

deben ser considerados en el proceso de análisis y diseño del serious game. Se incorpora la ingeniería de software como un enfoque sistemático para la construcción del juego en base a las características de los usuarios, determinándose 3 etapas: requerimientos y funcionalidades, diseño y evaluación del juego académico y terapéutico. Por otro lado, el uso de tecnologías emergentes a través de realidad aumentada permitirá la interacción del mundo real mediante la aplicación de entornos virtuales basados en movimiento. Por tal razón se considera que el modelo propuesto puede generar impacto positivo en el proceso de mejoramiento de la atención y por consecuencia en el rendimiento académico de los niños con trastornos de atención e hiperactividad.

Referencias

Adell, J., & Castañeda, L. (2012). Tecnologías emergentes, ¿pedagogías emergentes. *Tendencias emergentes en educación con TIC*, 13–32.

Aslan, S., & Balci, O. (2015). GAMED: digital educational game development methodology. *Simulation, 91*(4), 307–319.

Ayala, B., Ramírez, C. M., & Ocampo, L. M. (2012). La Ingeniería de Requerimientos aplicada al desarrollo de sistema de información.

Basogain, X., Olabe, M., Espinosa, K., Rouèche, C., & Olabe, J. (2007). Realidad Aumentada en la Educación: una tecnología emergente. *Escuela Superior de Ingeniería de Bilbao, EHU. Recuperado de http://bit. ly/2hpZokY*.

Cano, S., Arteaga, J. M., Collazos, C. A., Gonzalez, C. S., & Zapata, S. (2016). Toward a methodology for serious games design for children with auditory impairments. *IEEE Latin America Transactions, 14*(5), 2511–2521.

De la Peña Olvera, F. (2000). El trastorno por déficit de atención con hiperactividad (TDAH). *Rev Fac Med UNAM, 43*(6), 243–244.

De la Peña Olvera, F., Ortiz, J. D. P., & Pérez, E. B. (2010). Declaración de Cartagena para el Trastorno por Déficit de Atención con Hiperactividad (TDAH): rompiendo el estigma. *Revista Ciencias de la Salud, 8*(1).

Enríquez, A. L., Rodas-Osollo, J., & Ochoa-Zezatti, (2017) A. Using a Serious Game to Help with the Socialization of Children Identified with Asperger Syndrome.

Etxeberria, F. (2006). ¿ Son violentos los videojuegos violentos?. *Primeras Noticias: Comunicación y Pedagogía*, (216), 26–31.

Gil, G. D. (2002). *Herramienta para implementar LEL y escenarios (TILS)*. Facultad de Informática.

Gudiño Morales, C. L., & Herrera Jácome, P. J. (2013). Elaboración e implementación de un cd interactivo sobre juegos visuales para fortalecer la atención en los niños/as de 5 a 6 años, en el proceso enseñanza-aprendizaje, en el área de parvularia en la sala de la ludoteca de la universidad técnica Cotopaxi en la parroquia Eloy Alfaro, cantón Latacunga, provincia de Cotopaxi durante el periodo 2011-2012.

López Raventós, C. (2016). El videojuego como herramienta educativa. Posibilidades y problemáticas acerca de los serious games. *Apertura (Guadalajara, Jal.)*, *8*(1), 0–0.

Molina, J., & Martínez-González, A. E. (2015). Eficacia de una intervención computerizada para mejorar la atención en un niño con TDAH. *Revista de Psicología Clínica con Niños y Adolescentes*, *2*(2).

Martínez, N. M. M., Olivencia, J. J. L., & Terrón, A. M. (2016). Mobile learning, Gamificación y Realidad Aumentada para la enseñanza-aprendizaje de idiomas. *IJERI: International Journal of Educational Research and Innovation*(6), 16–34.

Oliva-Feria, A. Y., Mancebo-Ayala, R., & Betancourt-Aldana, G. (2017). La orientación familiar para mejorar las manifestaciones del trastorno por déficit de atención con hiperactividad en los escolares. *Santiago*(144), 638–651.

Olivencia, J. J. L., & Martínez, N. M. M. (2015). Tecnologías de geolocalización y realidad aumentada en contextos educativos: experiencias y herramientas didácticas. *DIM: Didáctica, Innovación y Multimedia*(31), 1–18.

Presentación Herrero, M. J., Siegenthaler Hierro, R., Jara Jiménez, P., & Miranda Casas, A. (2010). Seguimiento de los efectos de una intervención psicosocial sobre la adaptación académica, emocional y social de niños con TDAH. *Psicothema, 22*(4).

Pressman, R. S., & Troya, J. M. (1988). Ingeniería del software.

Rodríguez Molinero, L., López Villalobos, J., Garrido Redondo, M., Sacristán Martín, A., Martínez Rivera, M., & Ruiz Sanz, F. (2009). Estudio psicométrico-clínico de prevalencia y comorbilidad del trastorno por déficit de atención con hiperactividad en Castilla y León (España). *Pediatría Atención Primaria, 11*(42), 251–270.

Saavedra, A. B., Rodríguez, F. J. Á., Arteaga, J. M., Salgado, R. S., & Ordoñez, C. A. C. (2014). *A serious game development process using competency approach: Case Study: Elementary School Math*. Paper presented at the Proceedings of the XV International Conference on Human Computer Interaction.

Spinelli, A. T., & Massa, S. M. (2017). *Elicitación de Requerimientos, Centrada en el Usuario, para el Desarrollo de un Serious Game*. Paper presented at the XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017, ITBA, Buenos Aires).

Villar, I. O. (2004). Impacto y detección de niños con trastorno por déficit de atención con hiperactividad. *Educación y futuro: revista de investigación aplicada y experiencias educativas*(10), 11–20.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 25/10/2018
Aceitação/Acceptance: 09/12/2018

Explorando la relación entre Gestión del Conocimiento y el Rendimiento Organizativo en Instituciones de Educación Superior Universitaria

Sergio Araya-Guzmán[1], Carolina Henríquez Bravo[1], Patricio Ramírez-Correa[2], Ana Barra Salazar[1]

saraya@ubiobio.cl, carohenriquez04@gmail.com, patricio.ramirez@ucn.cl, abarra@ubiobio.cl

[1] Universidad del Bío-Bío, avenida Collao 1202, 4051381, Concepción, Chile.

[2] Universidad Católica del Norte, Larrondo 1281, 1781421, Coquimbo, Chile.

Pages: 947–959

Resumen: El objetivo de este estudio es explorar si la gestión del conocimiento tiene un impacto positivo en el rendimiento de las instituciones de educación superior universitaria. Teniendo en cuenta que la creación, el almacenamiento, la transmisión y el uso del conocimiento están involucrados en la gestión del conocimiento, se han aplicado encuestas a 50 profesionales no académicos de las universidades. Los resultados del análisis con ecuaciones estructurales indican que efectivamente dentro de estas entidades, hay un desarrollo de la gestión del conocimiento involucrado en su gestión y operación diaria, lo cual es posible que tenga un impacto positivo en su desempeño, principalmente en aspectos relacionados con el cumplimiento de los objetivos establecidos y la capacidad de adaptarse a los cambios.

Palabras-clave: Gestión del Conocimiento; Instituciones de Educación Superior Universitaria; Rendimiento Organizativo; Chile.

Exploring the relationship between knowledge management and organizational performance in university institutions of higher education

Abstract: The objective of this study is to explore whether knowledge management has a positive impact on the performance of university higher education institutions. Taking into account that the creation, storage, transmission and use of knowledge are involved in knowledge management, surveys have been applied to 50 non-academic professionals from universities. The results of the analysis with structural equations indicate that effectively within these entities, there is a development of knowledge management involved in its management and daily operation, which is likely to have a positive impact on its performance, mainly in aspects related to the compliance with established objectives and the ability to adapt to changes.

Keywords*:* Knowledge Management; University Higher Education Institutions; Organizational Performance; Chile.

1. Introducción

El conocimiento constituye una mezcla de experiencias, valores, información y saber hacer (Davenport & Prusak, 1998; Hernández, 2016), donde su gestión puede colaborar en la generación de nuevos conocimientos (Arboníes, 2001) y el desarrollo y mejora de los niveles de competitividad (Drucker, 1999), desempeño o rendimiento de la organización. Esta gestión del conocimiento contempla un conjunto de etapas a través de un proceso continuo, considerando por ejemplo la creación, adquisición, transmisión, utilización del conocimiento (Davenport & Prusak, 2001; Pallarès, 2014), para incrementar el aprendizaje de la organización (Lakomski, 2011), su valor (Ríos, 2012), su competitividad, desempeño, su rendimiento (Nonaka & Takeuchi, 1996).

Entre los estudios sobre este fenómeno, la investigación de Joglar (2010) sobre un modelo de gestión del conocimiento se destaca por su análisis holístico. En ese estudio se presenta un modelo en organizaciones militares que centra su atención en la identificación de aspectos asociados a la gestión del conocimiento, involucrando la creación, almacenamiento, transferencia y utilización del conocimiento.

Un tipo particular de organizaciones que permanentemente se ha preocupado por mejorar sus niveles de competitividad y rendimiento, son las Instituciones de Educación Superior Universitaria, las que siendo entidades generadoras de conocimiento a través de la ejecución de proyectos e investigaciones en distintas disciplinas, utilizan conocimiento para su funcionamiento y desarrollo de sus actividades habituales, donde, según Ríos (2012) se presenta un bajo nivel de desarrollo de la gestión del conocimiento.

En este contexto, a partir de lo establecido por Joglar (2010), este trabajo busca explorar si la gestión del conocimiento puede tener un efecto en el rendimiento de las Instituciones de Educación Superior Universitarias, entendiendo que en dicha gestión intervienen la creación, almacenamiento, transmisión y utilización del conocimiento.

2. Antecentes

El conocimiento es un concepto que a través del tiempo ha alcanzado mayor relevancia en las organizaciones, en el contexto de la denominada sociedad de conocimiento, surgiendo en la década de los noventa lo que se ha denominado gestión del conocimiento, de la mano de Nonaka & Takeuchi (1996). Según Davenport & Prusak (1998), el conocimiento es una mezcla de experiencia, valores, información y saber hacer, que sirve como marco para la incorporación de nuevas experiencias e información, siendo útil para la acción, se origina y aplica en la mente de las personas, encontrándose dentro de las organizaciones por medio de documentos, bases de datos, procesos, prácticas, normas y rutinas organizativas. En esta línea, la gestión del conocimiento puede ser utilizada por las organizaciones para generar valor y desarrollar su competitividad, constituyendo según Drucker (1999) "la manera de administrar los activos intangibles que aportan valor a la organización para conseguir capacidades o competencias esenciales distintivas como son los procesos, prácticas, saberes y destrezas de los individuos".

La gestión del conocimiento, según Arboníes (2001), se relaciona con la capacidad de organizar para crear nuevos conocimientos, diseminarlos y encapsularlos en productos, servicios y sistemas. Diversos autores (Rodríguez, 2006; Sallán, 2006; Lakomski, 2011;

Pallarès, 2014; entre otros) plantean la gestión del conocimiento como un conjunto de etapas a través de un proceso continuo, contemplándose la creación, adquisición, captura, transmisión y utilización del conocimiento, independientemente de donde se encuentre, con el fin de incrementar el aprendizaje y los resultados de las organizaciones, lo que coincide con lo señalado por autores como Davenport, De Long & Beers (1997) en el sentido que "la gestión del conocimiento es el proceso de captura, distribución y utilización eficaz de los conocimientos", o lo expresado por Davenport & Prusak (2001), que definen la gestión del conocimiento como un proceso lógico, organizado y sistemático para producir, transferir y aplicar en situaciones concretas una combinación armónica de saberes, "experiencias, valores, información contextual y apreciaciones expertas que proporcionan un marco para su evaluación e incorporación de nuevas experiencias e información". Ríos (2012) conceptualiza la gestión del conocimiento como "el conjunto de procesos que dirigen el análisis, diseminación, utilización y traspaso de experiencias, información y conocimientos entre todos los miembros de una organización para generar valor", incidiendo, en consecuencia, en el rendimiento de la organización. En este sentido, Nonaka & Takeuchi (1996), analizan el motivo del éxito de las empresas japonesas, llegando a la conclusión de que el uso de la gestión del conocimiento incrementa el desempeño de la organización, ya que permite la innovación, la creatividad y la imaginación. Estos autores no sólo señalan que hay un aumento en el desempeño o rendimiento organizativo, sino que además se incrementa la competitividad.

Por otra parte, el rendimiento organizativo es un aspecto que cada vez ha cobrado más relevancia en todo tipo de organizaciones. Un tipo particular de organizaciones, que también se han preocupado constantemente por mejorar sus niveles de rendimiento, son las universidades (Araya et al., 2007), las que permanentemente se ven enfrentadas a un entorno educacional cada vez más exigente (Araya & Chaparro, 2005; Gonçalves et al. 2016), y donde la gestión del conocimiento es un fenómeno que presenta un bajo nivel de desarrollo (Ríos, 2012). En este sentido, García (2016) plantea que el papel de la Universidad debe ser visto dentro de un sistema de educación superior, como el centro de una red que interconecta todo el conocimiento, considerando así a las personas que la integran, los sistemas de información con los que trabaja y la cultura organizacional con que cuenta la institución. Luego, resulta interesante analizar si la gestión del conocimiento puede tener un efecto en el rendimiento de las Instituciones de Educación Superior Universitarias, entendiendo que en dicha gestión intervienen la creación, almacenamiento, transmisión y utilización del conocimiento.

Este trabajo, considerando la investigación de Joglar (2010) sobre una propuesta de modelo de gestión del conocimiento en organizaciones militares, presenta un modelo de investigación que se encuentra en proceso de validación, que centra su atención en la identificación de aspectos asociados a la gestión del conocimiento, involucrando la creación, almacenamiento, transferencia y utilización del conocimiento, en universidades, y pretende verificar si esta gestión del conocimiento presenta alguna incidencia en el rendimiento organizativo de estas entidades.

3. Hipótesis de Investigación

El conocimiento generado en la organización debe ser de utilidad para generar nuevo conocimiento, que permita que la organización aprenda, lo que motiva que

este conocimiento sea almacenado correctamente, para facilitar su retención y posterior recuperación (Fainholc, 2005), lo anterior es base para la primera hipótesis de investigación:

H1: La creación del conocimiento presenta una incidencia positiva en la memoria organizativa de las instituciones de educación superior universitarias.

El conocimiento existente, almacenado en una memoria organizativa, puede ser de utilidad para el desarrollo de nuevas ideas o iniciativas para mejorar la calidad del trabajo, reducir la cantidad de errores o mejorar la forma de hacer las cosas, entendiéndose que esto significa que el conocimiento disponible podría tener un efecto sobre la efectividad de la organización (Joglar, 2010), lo que sustenta la segunda hipótesis de investigación:

H2: La memoria organizativa presenta una incidencia positiva en la efectividad organizativa en las instituciones de educación superior universitarias.

Sin embargo, no basta con que este conocimiento se encuentre almacenado en la organización, es necesario que se difunda y transmita entre sus miembros, basando su integración en el logro de los objetivos definidos (Barker & Camarata, 1998), lo que apoya la tercera hipótesis de investigación:

H3: La memoria organizativa presenta una incidencia positiva en la efectividad organizativa, a través de la transmisión interna del conocimiento, en las instituciones de educación superior universitarias.

La aplicación del conocimiento existente en la organización puede ser de utilidad para la creación de nuevas ideas e iniciativas que permitan mejorar la calidad del trabajo, reducir la cantidad de errores producidos o mejorar la forma de hacer las cosas (efectividad organizativa), lo que en consecuencia puede afectar el rendimiento de la organización, entendiendo que este refiere al cumplimiento, ejecución o logro, que denota un resultado cuantificado o una serie de resultados obtenidos (Gilley, Dean & Bierema, 2001). En lo anterior se basa la cuarta hipótesis de investigación:

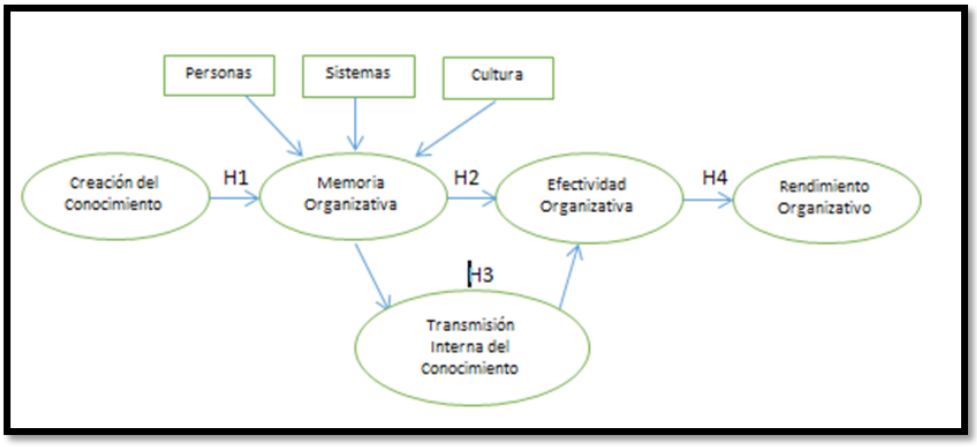


Figura 1 – Modelo de Investigación

H4: La efectividad organizativa presenta una incidencia positiva en el rendimiento organizativo de las instituciones de educación superior universitarias.

Las hipótesis definidas dan origen al modelo de investigación que se presenta en la figura 1.

4. Metodología

Este estudio es de carácter exploratorio, donde el estudio empírico considera como unidad de análisis instituciones de Educación Superior Universitaria en el contexto chileno y como unidad de observación profesionales no académicos con por lo menos un año de antigüedad en la institución. Los instrumentos de medición utilizados corresponden a instrumentos ya definidos y utilizados en otros estudios, utilizando una escala Likert de cinco puntos. El análisis estadístico de los datos contempla la aplicación de modelo de ecuaciones estructurales basado en PLS y la utilización del software SmartPLS.

5. Análsisis de Datos y Resultados

5.1. Resultados de la aplicación de encuesta

En este trabajo se presentan los resultados para 50 casos de encuestas válidas obtenidas. La Tabla 1 entrega la descripción muestra.

Categoría		**Nº**	**%**
Género	Femenino	34	68%
	Masculino	16	32%
Edad	Menor de 30 años	7	14%
	Entre 30 y 39 años	22	44%
	Entre 40 y 50 años	14	28%
	Mayor de 50 años	7	14%
Antigüedad en la Organización	Entre 1 y 5 años	15	30%
	Entre 6 y 10 años	11	22%
	Mayor de 10 años	24	48%
Años de Vida Laboral	Entre 1 y 5 años	7	14%
	Entre 6 y 10 años	9	18%
	Mayor de 10 años	34	68%

Tabla 1 – Descripción de la muestra

5.2. Resultados obtenidos

El modelo presenta un constructo de segundo orden (Memoria Organizativa, formado por los constructos de primer orden Personas, Sistemas y Cultura Organizativa) y cuatro constructos de primer orden (Creación ce Conocimiento, Transmisión Interna de Conocimiento, Efectividad Organizativa, Rendimiento Organizativo). Los datos

obtenidos de las encuestas fueron analizados por medio del software para análisis de ecuaciones estructurales SmartPLS (versión 3.2.7).

5.3. Evaluación Modelo de Medida

Dado que la Memoria Organizativa es un constructo de segundo orden, no se evalúa para éste el Modelo de Medida siendo necesario en estos casos evaluar la presencia de multicolinealidad entre los indicadores. Memoria Organizativa fue tratado considerando el procedimiento utilizado por el "método de construcción por medio de latent variables score (Cepeda & Roldán, 2004)", donde las cargas y pesos finales obtenidos se presentan en la Tabla 2. La presencia de multicolinealidad entre los indicadores formativos se evalúa por medio de la prueba de factor de influencia de la varianza (VIF). Los resultados de esta medición para el constructo Memoria Organizativa se presentan en la Tabla 3, donde es posible apreciar que todos los valores del indicador VIF son menores que 5 (valor exigido), lo que indica que no existen problemas asociados a multicolinealidad entre los indicadores del constructo.

Constructo	Carga	Peso
Cultura Organizativa (CO)		0.3702
MOCO1	0,7621	
MOCO11	0,5356	
MOCO12	0,6849	
MOCO2	0,7721	
MOCO3	0,8096	
MOCO8	0,6636	
Personas (P)		0.1699
MOP2	0,6285	
MOP4	0,8307	
MOP5	0,7919	
MOP6	0,8496	
MOP7	0,6466	
Sistemas (SI)		0.5340
MOSI1	0,7210	
MOSI2	0,7872	
MOSI3	0,7359	
MOSI4	0,7951	
MOSI5	0,8369	
MOSI6	0,7231	
MOSI7	0,8531	

Tabla 2 – Cargas y Pesos finales obtenidas para los constructos asociados a la Memoria Organizativa

Constructo	VIF
Cultura Organizativa (CO)	1,8709
Personas (P)	1,7031
Sistemas (SI)	1,6264

Tabla 3 – Evaluación de Multicolinealidad (VIF) de Memoria Organizativa

Los datos obtenidos, una vez finalizado el proceso de aplicación de PLS se presentan en la Tabla 4 y Tabla 5.

Constructo/Ítem	Cargas	Fiabilidad Compuesta (P_c)	Alfa de Cronbach	Varianza Extraída Media (AVE)
Creación Conocimiento (CC)		0,9136	0,8738	0,7259
CC2	0,8522			
CC3	0,7948			
CC4	0,9091			
CC5	0,8481			
Transmisión Interna del Conocimiento (TIC)		0,9031	0,8559	0,7008
TIC1	0,9076			
TIC2	0,7678			
TIC3	0,8824			
TIC4	0,7818			
Efectividad Organizativa (EO)		0,9121	0,8836	0,6348
EO2	0,7031			
EO4	0,7526			
EO5	0,8453			
EO6	0,8488			
EO7	0,7650			
EO8	0,8532			
Rendimiento Organizativo (RO)		0,9201	0,9012	0,5905
RO1	0,7270			
RO10	0,7843			
RO11	0,8196			
RO12	0,7516			
RO2	0,7806			
RO7	0,7485			
RO8	0,7758			
RO9	0,7563			

Tabla 4 – Cargas, Fiabilidad Compuesta, Alfa de Cronbach y AVE

Constructo	Creación Conocimiento (CC)	Efectividad Organizativa (EO)	Rendimiento Organizativo (RO)	Transmisión Interna Conocimiento (TIC)
Creación Conocimiento (CC)	0,8520			
Efectividad Organizativa (EO)	0,6189	0,7968		
Rendimiento Organizativo (RO)	0,6062	0,7680	0,7684	
Transmisión Interna Conocimiento (TIC)	0,2937	0,5311	0,4987	0,8371

Tabla 5 – Validez Discriminante

La fiabilidad individual del ítem se valora examinando las cargas de los indicadores con su respectivo constructo. El valor establecido como mínimo para que se cumpla esta fiabilidad es de 0.65. La Tabla 4 muestra que todos los ítems asociados a la Creación de Conocimiento, Transmisión Interna de Conocimiento, Efectividad Organizativa y Rendimiento Organizativo satisfacen esta condición, cumpliéndose, en consecuencia, la fiabilidad individual del ítem para cada caso.

La fiabilidad del constructo se valora examinando la fiabilidad compuesta (ρc). El valor asumido como apropiado para la fiabilidad compuesta es de 0.7. La Tabla 4 muestra que el valor de la fiabilidad compuesta de la Creación de Conocimiento, Transmisión Interna de Conocimiento, Efectividad Organizativa y Rendimiento Organizativo supera el mínimo establecido, cumpliéndose, entonces, la fiabilidad del constructo para cada caso.

La validez convergente se valora por medio de la Varianza Extraída Media (AVE). El valor del AVE debe ser superior a 0.5. La Tabla 4 muestra que el AVE de la Creación de Conocimiento, Transmisión Interna de Conocimiento, Efectividad Organizativa y Rendimiento Organizativo satisfacen el valor exigido, por lo que se cumple la validez convergente, para cada caso.

La validez discriminante se valora utilizando la Varianza Extraída Media (AVE). La condición exigida es que el AVE de un constructo debe ser superior a la varianza compartida entre él y los otros constructos del modelo. La Tabla 5 muestra los datos correspondientes a la validación discriminante, donde se puede apreciar que se satisface la condición exigida, cumpliéndose, por lo tanto, la validez discriminante.

Finalmente, dado que se cumple la fiabilidad individual del ítem, la fiabilidad del constructo, la validez convergente y la validez discriminante, se puede concluir que se cumple la evaluación del modelo de medida en forma satisfactoria.

5.2. Evaluación del Modelo Estructural

La evaluación del Modelo Estructural se lleva a cabo por medio de tres indicadores: los coeficientes path (β), la varianza explicada (R2) y los t-estadísticos. Los coeficientes path deben presentar un valor igual o superior a 0.2, la varianza explicada deber ser mayor o igual a 0.1. Los datos obtenidos para la evaluación de dicho modelo se muestran en la Tabla 6 y Tabla 7.

Relación		Coeficiente Path (β)	t-estadísticos/ Valor P	Significancia Estadística
Creación Conocimiento (CC)	Memoria Organizativa (MO)	0,6189	6,9957 / 0,0000	***
Memoria Organizativa (MO)	Efectividad Organizativa (EO)	0,6784	3,9697 / 0,0001	***
Memoria Organizativa (MO)	Transferencia Interna Conocimiento (TIC)	0,6674	7,3008 / 0,0000	***
Transferencia Interna Conocimiento (TIC)	Efectividad Organizativa (EO)	0,0783	0,4377 / 0,6618	ns
Efectividad Organizativa (EO)	Rendimiento Organizativo (RO)	0,7680	14,0742 / 0,0000	***

Para n=500 submuestras: *p<0.05; ** p<0.01; *** p<0.001 (basado en una distribución $t_{(499)}$ de Student de una cola)
$t_{(0.05;\ 499)}$=1.64791345; $t_{(0.01;\ 499)}$=2.333843952; $t_{(0.001;\ 499)}$=3.106644601

Tabla 6 – Coeficiente Path y significancia estadística

Constructo	Varianza Explicada (R^2)
Memoria Organizativa (MO)	0,3830
Efectividad Organizativa (EO)	0,5373
Transferencia Interna Conocimiento (TIC)	0,4455
Rendimiento Organizativo (RO)	0,5898

Tabla 7 – Varianza Explicada (R^2)

Considerando el análisis efectuado, se soportan las hipótesis H1, H2 y H4, y no es posible soportar la hipótesis H3. La Figura 2 muestra esquemáticamente el resultado de análisis del modelo estructural.

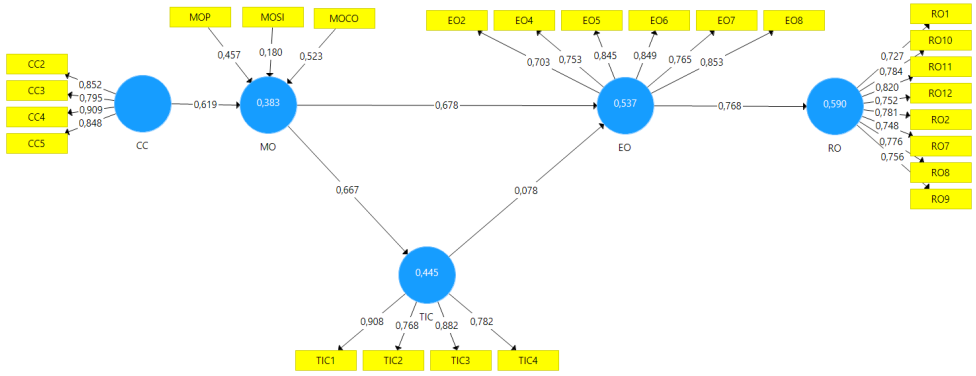


Figura 2 – Modelo Estructural

5.3. Discusión

En relación a la Creación del Conocimiento los tres aspectos donde más se ve reflejada son: (i) Mi organización tiene mecanismos para crear nuevos conocimientos a partir de conocimiento existente, (ii) Mi organización alienta y tiene procesos para el intercambio de ideas y conocimientos entre individuos y grupos, y (iii) Mi organización utiliza las lecciones aprendidas y las mejores prácticas de los proyectos para mejorar los proyectos sucesivos. En cambio, aspectos como "Mi organización tiene mecanismos para crear y adquirir conocimientos de diferentes fuentes, tales como empleados, clientes, socios comerciales y competidores" no aparecen reflejados. Lo anterior puede interpretarse en el sentido que en las Instituciones de Educación Superior se proyectan principalmente las fuentes internas para la creación de conocimiento organizativo, en comparación con fuentes externas, lo que podría estar asociado al aprendizaje que estas instituciones han desarrollado sobre la forma de hacer las cosas en su gestión y desarrollo como organización.

En relación a la Memoria Organizativa, puede apreciarse que de las tres variables involucradas, la que presenta una mayor relevancia es Sistemas, seguida de Cultura Organizativa y Personas. Específicamente, respecto a Sistemas, todos los aspectos medidos son aceptados, reflejándose principalmente en "El sistema de clasificación, codificación y administración de conocimientos facilita el trabajo de los empleados", "Todas las personas en la organización tienen acceso a las bases de datos disponibles de acuerdo a las necesidades de su cargo y trabajo", y "Las bases de datos, en caso que existan, son actualizadas en forma continua", mientras que aspectos como "La organización dispone de bases de datos que le permiten almacenar sus experiencias y conocimientos, para que puedan ser utilizados con posterioridad" no se visualizan. Lo anterior muestra que se otorga más relevancia a aspectos asociados con la disponibilidad y colaboración que otorga un conocimiento actualizado en el desarrollo de las labores habituales.

Respecto a la Cultura Organizativa, los tres aspectos donde más se refleja son "Los valores de mi organización representan con exactitud su esencia", "Los valores amparados por mi institución ejercen una influencia sobre mi comportamiento en la organización" y "Siento que entiendo los valores en los que se ampara mi organización", mientras que no se visualizan aspectos como "A través de relatos divulgados en mi organización he aprendido quienes son las personas importantes", "En relación con las actitudes de algunas personas en mi organización, a veces pienso: "si ellos lograron, yo también puedo hacerlo", "Hay personas en esta organización de las que frecuentemente se escucha hablar". Lo anterior permite percibir que adquieren más proyección aspectos relacionados con los valores institucionales.

Sobre las Personas, los tres aspectos donde más se refleja son "Los trabajadores sienten que son reconocidos por sus logros", "Sus compañeros de trabajo se comportan de una manera que beneficia el cumplimiento de los objetivos de la organización", y "El rendimiento de la unidad en la que trabaja se ve beneficiado por la contribución positiva individual de sus miembros", mientras que aspectos como "Usted siente que su trabajo es útil para la organización" y "Usted siente que sus habilidades personales están bien utilizadas en su trabajo" no se visualizan. Lo anterior muestra que se proyectan principalmente aspectos asociados con el comportamiento y la colaboración entre las personas.

En relación con la Efectividad, los tres aspectos donde más se refleja son "Mi organización ha incrementado los ingresos de productos y servicios que fueron inspirados por nuevos conocimientos", "Mi organización ha experimentado un crecimiento y uso significativo en los recursos de conocimiento (repositorios, patentes, publicaciones), y "Mi organización ha visto un crecimiento significativo en la capacidad de conocimiento de sus empleados", mientras que no se visualizan aspectos como "Mi organización ha creado muchas nuevas oportunidades de negocio", y "Mi organización está constantemente reajustando o re-alineando metas y objetivos para satisfacer cambios en el ambiente". Lo anterior refleja que se proyectan principalmente aspectos relacionados con el crecimiento del conocimiento existente en la organización y su utilización.

En relación al Rendimiento, los tres aspectos donde más se refleja son "Nuestra organización reacciona con rapidez a la toma de decisiones en forma proactiva e informada", "Nuestra organización reacciona con rapidez a los cambios del entorno", y "Los objetivos de la organización son alcanzados plenamente". Mientras que en los aspectos menos visualizados se encuentran "El desempeño futuro de nuestra organización es seguro", "Los recursos de nuestra organización se aprovechan eficientemente". Lo anterior muestra que se da mayor importancia a aspectos relacionados con la capacidad de adaptase a los cambios y al logro de los objetivos establecidos.

6. Conclusiones

Se aprecia una fuerte relación entre las variables estudiadas y con una alta significancia estadística: Creación del Conocimiento y Memoria Organizativa, Memoria Organizativa y Efectividad Organizativa, Memoria Organizativa y Transferencia Interna de Conocimiento, y Efectividad Organizativa y Rendimiento Organizativo, lo que ha motivado que se soporten las hipótesis H1, H2 y H4. Sin embargo, dado que no se cumple la relación entre Transferencia Interna del Conocimiento y la Efectividad Organizativa, no se soporta la hipótesis H3, a pesar que la relación entre Memoria Organizativa y Transferencia Interna del Conocimiento sea relevante.

Si bien no se han validado todas las hipótesis, es posible visualizar que efectivamente al interior de las Instituciones de Educación Superior Universitaria existe un desarrollo de la gestión del conocimiento involucrado en su gestión y funcionamiento cotidiano, donde adquiere relevancia la generación interna del conocimiento, la fácil disponibilidad de acceso a este conocimiento, la colaboración entre las personas para compartir su conocimiento, en un contexto donde interesan los valores institucionales, con lo cual es posible incidir positivamente en el rendimiento organizativo, principalmente en aspectos relacionados con el cumplimiento de los objetivos establecidos y en la capacidad de adaptarse a los cambios.

Este trabajo ha considerado 50 casos de encuestas válidas obtenidas, lo que constituye la principal limitación, siendo necesario realizar un análisis con una mayor cantidad de datos, de manera de validar los resultados ya obtenidos o bien explorar nuevos resultados y conclusiones. De igual forma, resulta interesante desarrollar este estudio en otros ámbitos o contextos, por ejemplo en Establecimientos Educacionales, organizaciones públicas, organizaciones privadas, lo que permitiría efectuar análisis comparativos entre diferentes entidades sobre su comportamiento frente a la gestión del conocimiento disponible y la posible incidencia en su rendimiento.

Referencias

Araya, S., & Chaparro, J. (2005). Una aplicación del Análisis de Recursos y Capacidades en el ámbito de Instituciones de Educación Superior Universitaria. In *IX Congreso Ingeniería de Organización* (p.110).

Araya, S., Chaparro, J., Orero, A., & Joglar, H. (2007). Sistemas y Tecnologías de Información y su relación con el Desempeño de Instituciones Universitarias en el ámbito de Gestión Institucional: un estudio desde la Teoría de Recursos y Capacidades. *In International Conference on Industrial Engineering & Industrial Management, en XI Congreso de Ingeniería de Organización (*pp. 1805–1816).

Arboníes, A. L. (2001). Las dificultades para construir la empresa del conocimiento. *Harvard Deusto Business Review*, (104), 46–51.

Barker, R. T., & Camarata, M. R. (1998). The role of communication in creating and maintaining a learning organization: Preconditions, indicators, and disciplines. *Journal of Business Communication*, 35 (4), 443–467.

Cepeda, G., & Roldán, J.L. (2004). Aplicando en la práctica la técnica PLS en la administración de empresas. *Congreso de la Asociación Científica de Economía y Dirección de la Empresa*.

Davenport, T. H., & Prusak, L. (1998). *Working knowledge: How organizations manage what they know*. Harvard Business Press.

Davenport, T. H., & Prusak, L. (2001). *Conocimiento en acción como las organizaciones manejan lo que saben*. Prentice Hall.

Davenport, T. H., De Long, D. W., & Beers, M. C. (1997). Building successful knowledge management projects. *Center for business innovation working paper*, *4*.

Drucker, P. F. (1999). *Los desafíos de la gerencia para el siglo XXI*. Editorial Norma.

Fainholc, B. (2006). Rasgos de las universidades y de las organizaciones de educación superior para una sociedad del conocimiento, según la gestión del conocimiento. *International Journal of Educational Technology in Higher Education (ETHE)*, *3*(1).

Gonçalves, M. J. A., Rocha, Á., & Cota, M. P. (2016). Information management model for competencies and learning outcomes in an educational context. *Information Systems Frontiers*, 18(6), 1051–1061.

García, B. (2016). Gestión del Conocimiento desde el Contexto Universitario. Red Peruana de Divulgadores Científicos, Universidad Nacional de Tumbes, Perú. *Iberciencia*. https://www.oei.es/historico/divulgacioncientifica/?Gestion-del-conocimiento-desde-el-contexto-universitario.

Gu, J. W., & Jung, H. W. (2013). The effects of IS resources, capabilities, and qualities on organizational performance: An integrated approach. *Information & Management,* 50(2-3), 87–97. DOI: 10.1016/j.im.2013.02.001.

Gilley, J., Dean, P. y Bierema, L. (2001). *Philosophy and practice of organizational learning, performance, and change*. Cambridge: Perseus Publishing.

Joglar, H. (2010). *Elaboración de un Modelo para el Diseño de Estrategias de Gestión del Conocimiento en Organizaciones Militares*. Tesis Doctoral, Universidad Politécnica de Madrid, España.

Lakomski, G. (2011). Saber cómo aprender: Liderazgo, gestión del conocimiento y el reto de crear comunidades de aprendizaje. *Educar*, 47(1), 13–30.

Liu, H., Ke, W., Wei, K. K., & Hua, Z. (2013). The impact of IT capabilities on firm performance: The mediating roles of absorptive capacity and supply chain agility. *Decision Support Systems*, 54(3), 1452–1262. DOI: 10.1016/j.dss.2012.12.016.

Medina, J. M., García, E., & De la Garza, M. (2009). Influence of the implementation factors in the information systems quality for the user satisfaction. *JISTEM-Journal of Information Systems and Technology Management*, 6(1), 25–44.

Nonaka, I., & Takeuchi, H. (1996). The knowledge-creating company: How Japanese companies create the dynamics of innovation. *Long range planning*, *4*(29), 592.

Pallarès, M. (2014). Los medios de comunicación y tecnológicos como ejes de canalización y gestión del conocimiento. *Educar*, *50*(1), 207–229.

Porter, M. (1980). *Competitive Strategy: techniques for analyzing industries and competitors*, New York: Free Press.

Ríos, T. J. (2012). La Gestión del Conocimiento y la Educación Superior Universitaria. Gestión en el Tercer Milenio, 15(30), 43–48.

Rodríguez, D. (2006). Modelos para la creación y gestión del conocimiento: una aproximación teórica. *Educar*, (37), 25–39.

Sallán, J. G. (2006). Las comunidades virtuales de aprendizaje. *Educar*, (37), 41–64.

Sánchez, D. G., Reyes, R. G. R., Sekeres, M. F. A., & Ortiz, J. H. G. (2013). Satisfacción laboral y compromiso en las organizaciones de Rio Verde, S.L.P. *Revista de Psicología y Ciencias del Comportamiento de la UACJS: RPCC-UACJS*, 4(1), 59–76.

Tian, J., Wang, K., Chen, Y., & Johansson, B. (2010). From IT deployment capabilities to competitive advantage: An exploratory study in China. *Information Systems Frontiers*, *12*(3), 239–255.

Turban, E., Volonino, L., & Wood, G. R. (2015). *Information Technology for Management: Digital Strategies for Insight, Action, and Sustainable Performance*, Hoboken: John Wiley & Sons.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 27/09/2018
Aceitação/Acceptance: 04/11/2018

Interpretación de lenguaje de señas ecuatoriano empleando visión por computador

Javier P. Pichucho[1], Patricia N. Constante[2], Andrés M. Gordón[3], Darío J. Mendoza[4]

jppichucho@espe.edu.ec, pnconstante@espe.edu.ec, amgordon@espe.edu.ec, djmendoza@espe.edu.ec

[1,2,3,4] Universidad de las Fuerzas Armadas ESPE, Sangolquí – Ecuador.

Pages: 960–971

Resumen: Se propone un sistema de visión artificial eficiente y de fácil acceso para la interpretación del alfabeto y números en lenguaje de señas ecuatorianas (LSEC) para ser ejecutado por diferentes usuarios en ambientes distintos, como una herramienta que contribuya al aprendizaje del LSEC. Se inicia con la adquisición de la imagen a través de una cámara web, esta es la entrada de la etapa de procesamiento, donde la mano del usuario es la región de interés (ROI), se le aplica diferentes técnicas de visión por computador para combatir los efectos de luz y otras anomalías, luego se usan Redes Neuronales Artificiales (RNA) para el aprendizaje del abecedario y números del LSEC mediante el empleo de patrones; se ha usado C++ con librerías de OpenCV y la IDE de Qtcreator. El sistema ha sido probado en tiempo real por diferentes usuarios, identificando satisfactoriamente 9 números y 25 letras.

Palabras-clave: Visión artificial; Redes neuronales artificiales; Lenguaje de señas.

Interpretation of Ecuadorian Sign Language using computer vision.

Abstract: This work proposes an efficient and easily accessible computer vision system for the interpretation of the Ecuadorian alphabet and sign language (LSEC) numbers that can be used regardless of the environment and type of user. It begins with the acquisition of the image in the form of a video through a webcam, this will be the entrance to the processing stage. It is intended to use the hand of the user as a region of interest (ROI) that will then be analyzed and processed to combat the effects of light and other anomalies that may arise, then use Artificial Neural Networks (ANN) for learning the alphabet and LSEC numbers by using patterns obtained from different hands forming the letters of the alphabet and necessary numbers, has been used C + + with libraries OpenCV and the IDE Qtcreator.

Keywords*:* Artificial vision; Artificial neural networks; Sign language.

1. Introducción

La sociedad hace ya varios años ha venido interesándose en los grupos de atención prioritaria y personas con discapacidad. Dentro del grupo de individuos con discapacidad

se encuentran aquellas que poseen discapacidad auditiva, las mismas que en Ecuador y en todo el mundo requieren de la atención y el apoyo de herramientas tecnológicas (Caliwag, Angsanto, & Lim, 2018) que ayuden a mejorar sus capacidades y contribuyan con su desarrollo tanto lingüístico social y psicológico. Para cooperar con el progreso de las personas con discapacidad auditiva (Bueno & Limiñana, 2002) es conveniente que la sociedad tenga la oportunidad de poder comunicarse con ellos y de alguna forma contribuir con la atención de sus inquietudes, requerimientos y seguir fomentando la defensa de sus derechos. La comunicación entre personas sordas y personas oyentes es indispensable para convivir en una sociedad pluricultural y solidaria (Riotordo, 2012).

Mediante la Visión Artificial se busca capturar e interpretar la información visual del entorno que rodea al sistema inteligente con el fin de extraer características cruciales para asimilar y pronosticar (Díez & Gómez, 2001) el alfabeto y números en LSEC. En (Piedra, 2008) (Caliwag, Angsanto, & Lim, 2018) (Ji, Kim, & Lee, 2017) los autores presentan los niveles de visión artificial que permiten obtener la región de interés en el caso de estudio la mano del usuario.

El estudio de la visión artificial ha sido utilizado para la interpretación y reconocimiento de gestos (Wang, Zhang, & Zhuang, 2012) así como del lenguaje de señas (Salian et al., 2017) a nivel mundial dando un paso importante en la inclusión social para la población que sufre de discapacidad auditiva. En (Escudeiro et al, 2015) se presenta un traductor de lenguaje de señas portugués que utiliza dos tecnologías el sensor de movimiento Kinect de Microsoft y los guantes de sensores 5DT los cuales mediante la aplicación de algoritmos realizados en OpenCV y el método Support Vector Machine logran la traducción de las señas en un texto legible. En (Rivas et al., 2017) se usa un sensor Leap Motion para la adquisición de los gestos e interpretación del alfabeto ecuatoriano.

El presente proyecto se basa en el desarrollo de un algoritmo para detectar los datos más importantes de la imagen de la mano del usuario y con esto mediante la utilización de clasificadores reconocer cada uno de los patrones que forman las letras del abecedario y números en LSEC, que funciona en diferentes condiciones de luz y diferentes usuarios. Se presenta una aplicación “Aprendiendo el lenguaje de señas ecuatoriana” de libre difusión que permita el aprendizaje tanto del abecedario como de los números de manera interactiva en forma de juego el mismo que al final será evaluado para medir los conocimientos adquiridos.

2. Desarrollo

2.1. Visión Artificial y tratamiento de imágenes.

El diseño del sistema de visión artificial se apoya en las etapas primordiales para el proceso de visión artificial (Piedra, 2008), tomando en consideración que el proyecto va orientado a satisfacer necesidades del aprendizaje del Lenguaje de Señas Ecuatoriana en entornos no controlados y con beneficiarios distintos. Como se observa en la figura 1, el sistema se basa en la adquisición de las imágenes a través de una cámara web (Noor et al., 2017) tomando énfasis en la mano del usuario, para mediante técnicas apropiadas del tratamiento de imágenes extraer las características más relevantes (Salian et al., 2017) dicha imagen se constituye en una entrada adecuada para una RNA (Constante et al.,

2016) (Badi, Hamza, & Hasan, 2017) que se encarga de la interpretación del alfabeto y de los números del uno al diez en Lenguaje de Señas Ecuatoriana. Se usa una computadora Intel Core i5 480M (2670 MHz) con Memoria RAM de 8.00 GB.

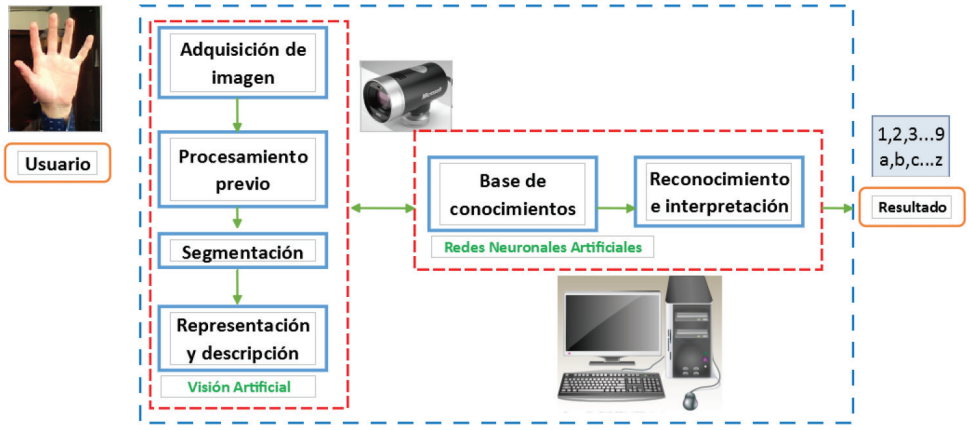


Figura 1 – Etapas fundamentales para el sistema de visión artificial

2.2. Procesamiento y Segmentación de la imagen.

Se requiere extraer únicamente la mano del usuario por lo que se ha aplicado segmentación como primera etapa, luego se utiliza procesamiento de la imagen y con esto se extrae la información relevante que sea de utilidad para los clasificadores, las etapas necesarias se presentan en la Figura 2.

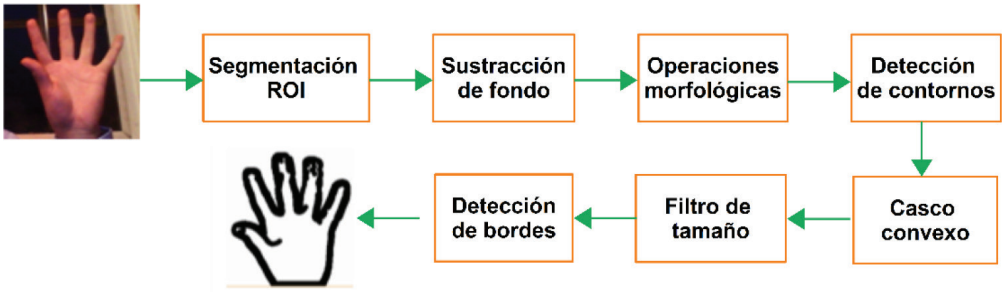


Figura 2 – Procesamiento y segmentación de la imagen

2.2.1. Segmentación área de interés

La segmentación es una etapa que permite excluir objetos que no tienen importancia en una escena y se basa en los principios de semejanza y discontinuidad. En la imagen capturada por la webcam se busca la región de interés (ROI), (Hasan et al., 2017) (Zhang J., & Liu Z., 2007) la cual permitirá separar un cierto número de píxeles de la imagen para convertirlos en un cuerpo entero (Fig. 3).

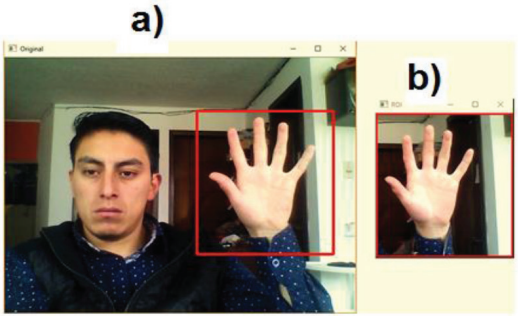


Figura 3 – Extracción de la ROI a) Imagen original b) ROI

2.2.2. Sustracción de fondo

El fin de la sustracción de fondo es obtener una imagen en primer plano la cual goza de un movimiento (Gervasoni et al., 2014), el principio de la sustracción de fondo se basa en restar las tramas actuales con un fondo que posee la parte inmóvil de la escena (Zhang J., & Liu Z., 2007) como se puede observar en la figura 4.

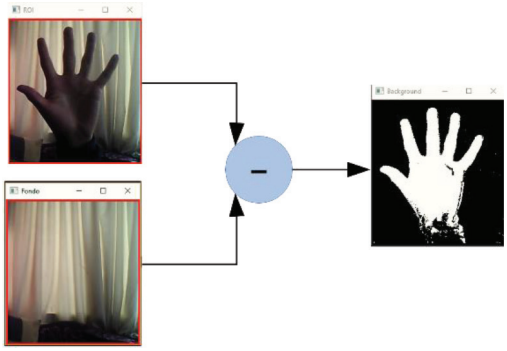


Figura 4 – Sustracción de fondo

2.2.3. Operaciones morfológicas

Las operaciones morfológicas como apertura y cierre (Noor et al., 2017) normalmente son aplicadas en imágenes binarias (Hasan et al., 2017), cuya finalidad es realzar características importantes de una imagen mientras elimina el ruido. Estas operaciones en la mano se aprecian en la figura 5.

Figura 5 – Operaciones morfológicas a) Imagen sustracción de fondo b) Imagen limpia

2.2.4. Detección de contornos, Casco Convexo y Filtro de tamaño.

Se entiende por contorno aquella curva que combina todos los puntos continuos alrededor de una frontera en el caso de estudio la mano del usuario. Lo que busca el casco convexo es encontrar el contorno que encierra una región pero que se adapta a esta por completo. El casco convexo es utilizado para formar la máscara que contendrá la imagen de la mano. Para el etiquetado de la imagen se procede a utilizar un filtro de objeto de acuerdo a su tamaño con el fin de eliminar un porcentaje de ruido que pueda aparecer en la imagen, así como también la reducción del número de objetos presentes en la misma ya que existen cuerpos de determinados tamaños que no interesan en la escena. (Fig. 6)

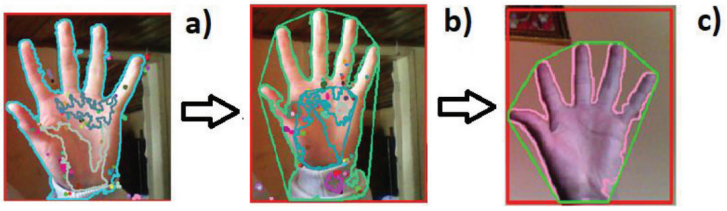


Figura 6 – Mano a) Contorno de la mano b) Casco convexo c) Filtro de tamaño

2.2.5. Detección de Bordes

Tiene como finalidad reconocer los cambios drásticos en el brillo de una imagen o dicho en otras palabras buscar las discontinuidades que esta posea. (Fig. 7)

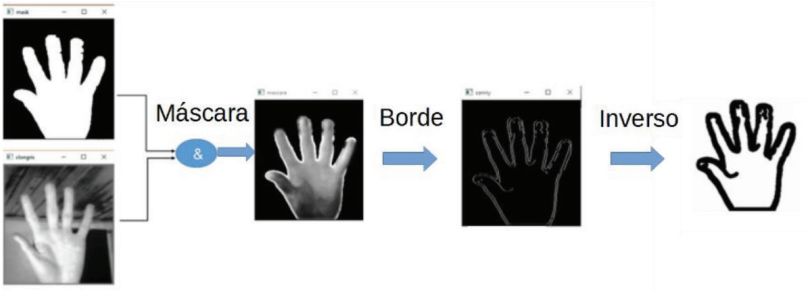


Figura 7 – Detección de bordes

2.3. Entrenamiento

OpenCV ofrece grandes prestaciones referente al uso de Redes Neuronales Artificiales (RNA), usa perceptrones multicapa (MLP) (Fatma Zohra & Amar, 2015) que constan de capa de entrada, capa de salida y una o más capas ocultas.

Se usa neuronas sigmoideas (Constante et al., 2016) (Zuocai, Bin, & Jin, 2018) para el reconocimiento de señas, las muestras tomadas para el entrenamiento pertenecen a personas que son integrantes de la FENASEC y del Instituto Nacional de Audición y Lenguaje - INAL, ya que conocen muy bien la Lengua de Señas Ecuatoriana.

$$f(x)= (\beta*(1-e^{(-\propto x)}))/(1+e^{(-\propto x)}) \quad (1)$$

Se diseña dos clasificadores basados en el modelo de (Raj, 2013) (Fig. 8), un clasificador encargado del reconocimiento de letras y uno para los números.

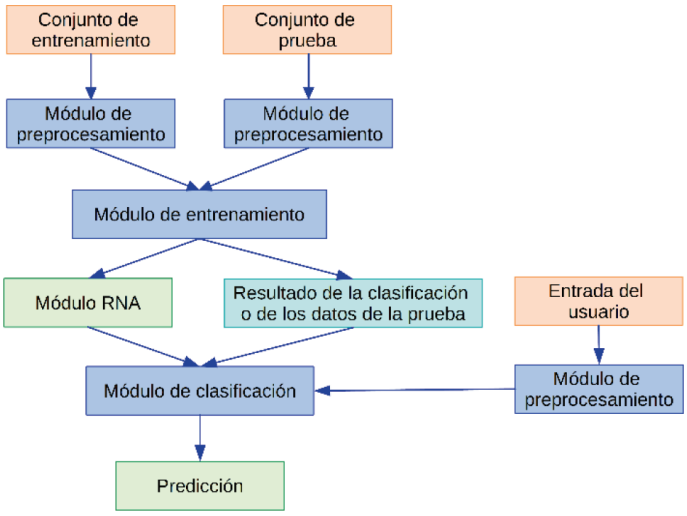


Figura 8 – Modelo para creación de clasificadores

El modelo consta de tres etapas, la etapa de pre-procesamiento, la etapa de entrenamiento y la etapa de clasificación.

La etapa de pre-procesamiento permite transformar cualquier imagen tomada en una de 16x16 píxeles para con esto obtener un conjunto de datos que se constituye en la entrada de la RNA, dichos datos contienen únicamente valores de 0 (color negro) y 1 (color blanco).

Se crean clasificadores con una capa de entrada, una oculta y una de salida (Constante et al., 2016). El entrenamiento termina después de 1000 iteraciones o si los pesos tienen valores por debajo de 0.00001, se utiliza retropropagación (Zuocai, Bin, & Jin, 2018) con coeficientes estándares que brinda el software OpenCV.

La RNA para clasificación de los números tiene 256 neuronas en la capa de entrada, 16 neuronas en la capa oculta y 9 en la capa de salida (Fig. 9), con lo que se podrá identificar los números del 1 al 9. Se emplean 1530 muestras de entrenamiento, 170 muestras por dígito y 900 muestras para validación.

La RNA para clasificación de las letras tiene 256 neuronas en la capa de entrada, 34 neuronas en la capa oculta y 25 en la capa de salida (Fig. 10), con lo que se podrá identificar las letras del alfabeto. Se crea un conjunto de entrenamiento de 4250 muestra, 170 muestras por dígito y un conjunto para pruebas de 2500 muestras para validación de la RNA.

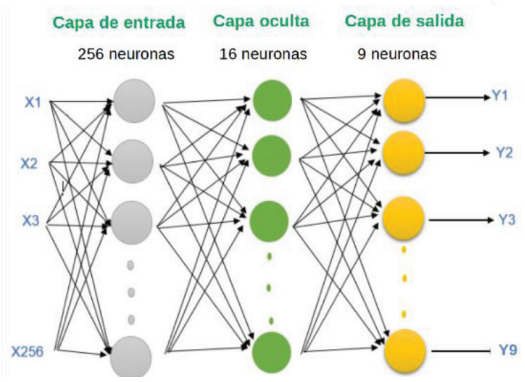


Figura 9 – RNA para clasificación de números

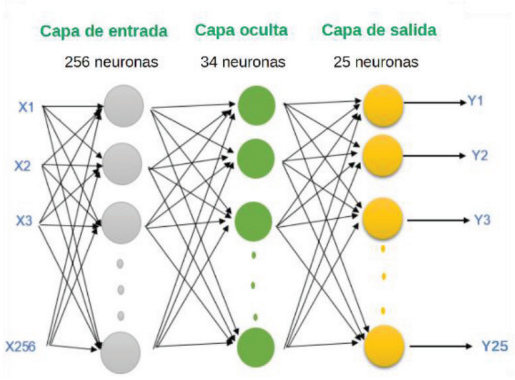


Figura 10 – RNA para clasificación de las letras del alfabeto

La clasificación permite inducir el conocimiento mediante el entrenamiento de una máquina para la interpretación y reconocimiento de clases o categorías a la cual corresponde un objeto (Fig. 11).

Para confirmar el entrenamiento de los agentes inteligentes se realiza una prueba con imágenes que contienen los números/letras en LSEC las mismas que la RNA nunca observó en su etapa de entrenamiento, (Constante et al., 2016) verificando que el reconocimiento sea mayor al 86%. (Tabla 1).

N°- neuronas capa oculta	N.º Muestra	Categoría	Imágenes reconocidas	% Error
16	100	números	92	8 %
34	200	letras	197	1,5 %

Tabla 1 – Porcentaje de error en la clasificación de cada número y letra en LSEC

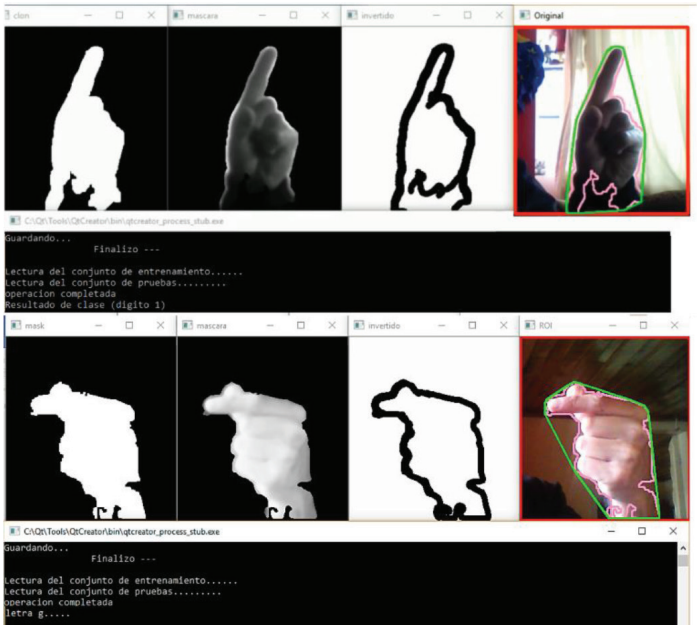


Figura 11 – Reconocimiento del lenguaje de señas a) número 1 b) letra g

2.4. Interfaz Gráfica de Usuario (GUI)

Se propone el diseño de una interfaz gráfica que permita al usuario interactuar de manera intuitiva y sencilla con el programa (Fig. 12).

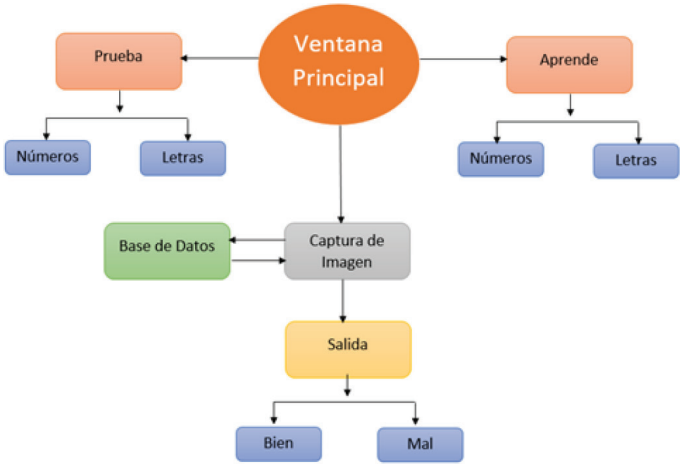


Figura 12 – Funcionamiento de la GUI

Se toma en cuenta que la GUI es orientada a la enseñanza por tanto se realiza una interfaz que capte la atención del usuario y el aprendizaje sea de manera interactiva y emocionante (Fig. 13).

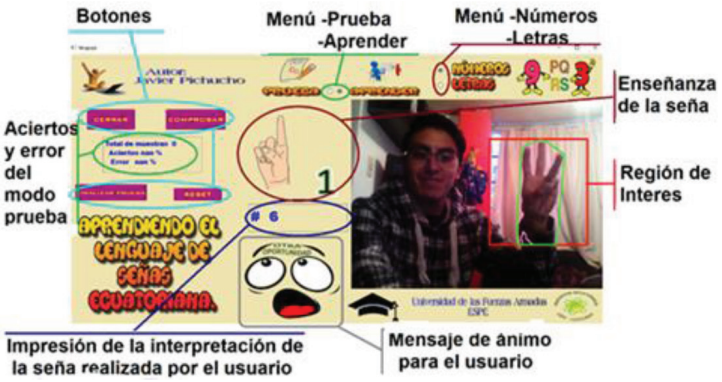


Figura 13 – Bosquejo de la interfaz gráfica

3. Experimentación

La luminosidad en varios proyectos de visión artificial es muy importante ya que para un correcto funcionamiento se requiere de entornos controlados donde no cambie constantemente la luz es decir permanezca estática, para la presente aplicación el entorno varía, no permanece constante por tal motivo el algoritmo se implementa buscando compensar cambios moderados en el ambiente, ante los cuales el sistema responde satisfactoriamente (fig 14.a-14.c), sin embargo es suceptible a cambios bruscos de iluminación o un fondo inadecuado como una ventana.

El funcionamiento de la aplicación denominada “Aprendiendo el lenguaje de señas ecuatoriano” se ha probado en diferentes usuarios, puede ser este niño o niña, adolescente, adulto o adulto mayor, sin interesar su tonalidad de piel y el tamaño de sus manos. (Fig 14.d- 14.f)

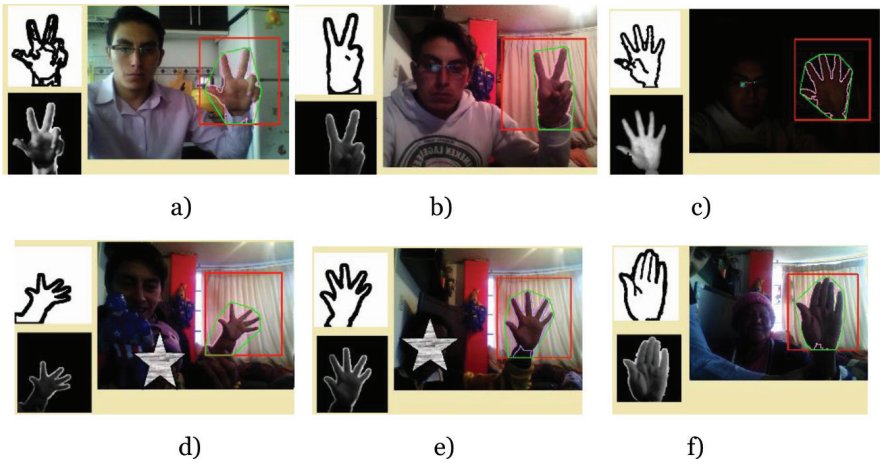


Figura 14 – Detección de la mano del usuario

a) Cocina (toma diurna) b) Dormitorio (toma nocturna) c) Dormitorio (muy poca iluminación) d) niño de 2 años e) niño de 11 años f) adulto de difernte tonalidad

Las figuras a, b y c evidencian la detección de la mano del usuario en condiciones cambiantes de iluminación, d, e y f evidencian la detección de la mano del usuario pese al tamaño y la tonalidad de piel que este tenga.

Para verificar la capacidad de reconocimiento del sistema se evalúa con 30 muestras para cada número/letra obteniendo, 270 muestras para los números y 750 para las letras (Tabla 2)

S	**A**	**S**	**A**	**S**	**A**	**S**	**A**
1	26	A	28	I	30	S	29
2	30	B	30	K	29	T	30
3	27	C	29	L	30	U	28
4	30	Ch	30	M	27	V	30
5	30	D	28	N	26	W	30
6	26	E	28	O	27	X	25
7	25	F	28	P	30	Y	29
8	30	G	28	Q	28		
9	25	H	25	R	30		

*S: Símbolo, A: Acierto

Tabla 2 – Número de aciertos en la interpretación de símbolos.

En la tabla 3 se puede observar el porcentaje promedio de todas las pruebas realizadas, obteniéndose un porcentaje de confiabilidad de reconocimiento y clasificación de las letras del abecedario en LSEC del 94,93% y 92,22% para los números. Un video demostrativo se presenta en https://youtu.be/MEeNB3BseFU

Categoría	**Nº muestras**	**Coincidencias**	**Fallo**	**Reconocimiento (%)**
Números	270	249	21	92,22 %
Letras	750	712	38	94,93 %

Tabla 3 – Porcentaje de confiabilidad en la clasificación de números y alfabeto en LSEC.

4. Conclusiones

Se ha desarrollado un sistema de visión artificial que permite el reconocimiento de números del 1 al 9 y de 25 letras del lenguaje de señas ecuatorianas.

Se ha aplicado técnicas de visión por computador que permiten al sistema el reconocimiento de la mano en condiciones variables de luz, variación de tonalidad de la pie, orientación y tamaño de la misma, lo cual permite usarlo como una herramienta que contribuya al aprendizaje de LSEC en niñas, niños y adultos.

Dos RNA son utilizadas en el aprendizaje de las características representativas de cada letra del abecedario y de número comprendido entre el uno y el nueve, y las letras del alfabeto sin considerar las que presentan movimiento para su reconocimiento es decir el sistema no permite el reconocimiento del 10, "j", "ll", "m", "ñ", "rr" y "z"

Se establece un porcentaje de confiabilidad superior al 92% a los clasificadores creados a partir del entrenamiento de las RNA con la ayuda de imágenes tomadas de las manos formando los diferentes números o letras de personas que conocen muy bien el LSEC como son los integrantes de la FENASEC y del INAL, con lo que el sistema es capaz de predecir en forma efectiva cuando el usuario realiza la seña de manera adecuada.

Referencias

Badi, H., Hamza, A., & Hasan, S. (2017). New method for optimization of static hand gesture recognition. *Intelligent Systems Conference (IntelliSys)*. doi:10.1109/intellisys.2017.8324347

Bueno, A. B., & Limiñana, A. R. (2002). Guía didáctica de psicología de la intervención social. Club Universitario.

Caliwag, A., Angsanto, S., & Lim, W. (2018). Korean Sign Language Translation Using Machine Learning. 2018 Tenth International Conference on Ubiquitous and Future Networks (ICUFN). doi:10.1109/icufn.2018.8436747

Constante P., Gordon A., Chang O., Pruna E., & Escobar I. (2016). Neural networks for optic nerve detection in digital optic fundus images. 2016 IEEE International Conference on Automatica (ICA-ACCA). doi:10.1109/ica-acca.2016.7778415

Constante P., Gordon A., Chang O., Pruna E., Acuña F. & Escobar I. (2016). Artificial Vision Techniques to Optimize Strawberry's Industrial Classification. *IEEE Latin America Transactions IEEE Latin Am. Trans., 14, 6*, 2576–2581.

Díez, R. P., & Gómez, A. G. (2001). Introducción a la inteligencia artificial: sistemas expertos, redes neuronales artificiales y computación evolutiva.

Escudeiro, P., Escudeiro, N., Reis, R., Lopes, J., Norberto, M., Baltasar, A. B., & Bidarra, J. (2015). Virtual Sign – A Real Time Bidirectional Translator of Portuguese Sign Language. *Procedia Computer Science*, 253–262. Obtenido de https://doi.org/10.1016/j.procs.2015.09.269

Fatma Zohra, C., & Amar, D. (2015). Face Recognition Using MLP and RBF Neural Network with Gabor and Discrete Wavelet Transform Characterization: A Comparative Study. *Mathematical Problems in Engineering*. doi:https://doi.org/10.1155/2015/523603

Gervasoni L., D'Amato J., Barbuzza R., Vénere M., Bertolino G., Cantero M., Storti M., & Teruel F. (2014). Un Método Eficiente para la Sustracción de Fondo en Videos Usando GPU. doi:10.13140/2.1.1478.1123.

Hasan M., Khaliluzzaman M., Himel S. & Chowdhury R. (2017). Hand sign language recognition for Bangla alphabet based on Freeman Chain Code and ANN. 2017 4th International Conference on Advances in Electrical Engineering (ICAEE). doi:10.1109/icaee.2017.8255454

Ji, Y., Kim, S., & Lee, K.-B. (2017). Sign Language Learning System with Image Sampling and Convolutional Neural Network. 2017 First IEEE International Conference on Robotic Computing (IRC). doi:10.1109/irc.2017.40

Noor, S., Shohan, J. A., Waresi, M. S., Rhythm, S. A., & Shahnaz, C. (2017). Real time hand movement controlled robotic arm for risk prevention. *IEEE Region 10 Humanitarian Technology Conference (R10-HTC)*. doi:IEEE Region 10 Humanitarian Technology Conference (R10-HTC).

Piedra, J. A. (2008). Aplicación de los sistemas neurodifusos a la interpretación automática de imágenes de satélite. Universidad Almería.

Raj, N. (2013). OCR using Artificial Neural Network (OpenCV) – Part 1. Obtenido de https://archive.is/0J4Vm

Riotordo, A. (2012). ¿Cultura Sorda? - A.P.A.S.U. Recuperado el 23 de Enero de 2018, de http://www.apasu.org.uy/b-sordos/d-cultura-sorda

Rivas y otros. (2017). Learning Sign Language of Ecuador. *In: Huang TC., Lau R., Huang YM., Spaniol M., Yuen CH. (eds) Emerging Technologies for Education. SETE, 10676*. Springer, Cham.

Salian, S., Dokare, I., Serai, D., Suresh, A., & Ganorkar, P. (2017). Proposed system for sign language recognition. *International Conference on Computation of Power, Energy Information and Commuincation (ICCPEIC)*. doi:10.1109/iccpeic.2017.8290339

Wang, G.-W., Zhang, C., & Zhuang, J. (2012). An Application of Classifier Combination Methods in Hand Gesture Recognition. *Mathematical Problems in Engineering*, 17. Obtenido de https://doi.org/10.1155/2012/346951.

Zhang J., & Liu Z. (2007). A Vision-Based Road Surveillance System Using Improved Background Subtraction and Region Growing Approach. Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing (SNPD 2007). doi:10.1109/snpd.2007.495

Zuocai, W., Bin, C., & Jin, W. (2018). Effective Inertial Hand Gesture Recognition Using Particle Filtering Based Trajectory Matching. *Journal of Electrical and Computer Engineering*, 9. Obtenido de https://doi.org/10.1155/2018/6296013.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 25/10/2018
Aceitação/Acceptance: 07/12/2018

Realidad Aumentada como recurso de apoyo en el proceso enseñanza-aprendizaje en Medicina Legal

Jennifer Betzabe Samaniego-Franco[1], Martha Vanessa Agila-Palacios[1], Dunia Inés Jara-Roa[1], Celia Paola Sarango-Lapo[1]

jbsamaniego@utpl.edu.ec, mvagila@utpl.edu.ec, dijara@utpl.edu.ec, cpsarango@utpl.edu.ec.

[1] Universidad Técnica Particular de Loja, Dirección de Tecnologías para la Educación, 1101608, Loja Ecuador.

Pages: 972–984

Resumen: El presente artículo se describe la metodología para el desarrollo de Objetos de Aprendizaje (OA) en 3D para la aplicación (app) de realidad aumentada (RA) "Virtopsia" en Medicina Legal, así como, los resultados del nivel de aceptación de los estudiantes de la carrera Derecho. La implementación de la app en el aula, se enmarcó en el ciclo de aprendizaje experiencial de Kolb, tomando en cuenta que el uso de la tecnología por sí sola, no genera aprendizajes significativos. Para poder analizar el criterio y percepción en cuanto a usabilidad del aplicativo se aplicó un instrumento de evaluación al estudiante de modalidad a distancia de la Universidad Técnica Particular de Loja de la asignatura de "Medicina Legal" en los siguientes ámbitos: 1) contenido, 2) usabilidad, mediante el uso de "Virtopsia" con RA. Los resultados demuestran una alta aceptación por parte de los estudiantes y manifiestan que el aplicativo representó un apoyo en el proceso de enseñanza-aprendizaje. A futuro se complementará esta investigación con un estudio del impacto del ciclo de aprendizaje experiencial en estas prácticas.

Palabras-clave: Realidad aumentada; aprendizaje experiencial; medicina legal; aprendizaje activo, autopsia.

Augmented Reality as a support resource in the teaching-learning process in Legal Medicine

Abstract: This article describes the methodology for the development of Learning Object (OA) in 3D for the application (app) of Augmented Reality (AR) "Virtopsia" in Legal Medicine, and shows the results of the level of acceptance of Law students. The implementation of the app in the classroom was part of Kolb's experiential learning cycle, taking into account that the use of technology alone does not generate significant learning. In order to analyze the criterion and perception, an instrument of evaluation was applied to the student of distance modality of the Universidad Técnica Particular de Loja of the subject of "Legal Medicine" in the following areas: 1) content, 2) usability, through the use of of "Virtopsia" with RA. The results show a high acceptance by the students and show that the application represented a support in the teaching-learning process. In the future, this research will be complemented by a study of the impact of the experiential learning cycle on these practices.

Keywords*:* Augmented reality; experiential learning; legal Medicine; active learning, autopsy.

1. Introducción

La educación en la modalidad a distancia cada vez presenta diferencias insignificantes con la modalidad presencial, esto gracias a los avances de la tecnología que permite a las instituciones educativas brindar un acompañamiento más cercano al estudiante. Es así que la educación a distancia, según sus avances con la tecnología, se categoriza en etapas o generaciones de innovación tecnológica, actualmente nos encontramos con la cuarta generación como lo menciona (García Aretio, 2003). Esta generación está basada en un conjunto de sistemas, algunos como soporte del funcionamiento y otros para entrega de actividades síncronas y asíncronas en diferentes formatos; la implementación de esta generación es con el propósito de brindar una interacción directa entre el docente-estudiante y estudiante-estudiante, es así que la tecnología proporciona la capacidad de transformar de forma radical el sistema educativo; por ello se debe elegir la tecnología que se va a implementar tomando en cuenta la relación con los objetivos planteados y las circunstancias de la institución para que su implementación tenga resultados efectivos en el proceso enseñanza-aprendizaje (Bates, 1999).

La Realidad Aumentada (RA) es una tecnología innovadora que brinda un apoyo potencial en el proceso de enseñanza-aprendizaje, se la utiliza como una nueva estrategia en los diferentes enfoques pedagógicos (Billinghurst & Dünser, 2012; Dunleavy, Dede, & Mitchell, 2009; Johnson, Adams, Cummins, Estrada, Freeman, y Hall, 2016), debido a los 3 subsistemas fundamentales (Agudelo, 2005):

- Visualización. – corresponde a la tecnología y dispositivos que se utilizan para acceder, pueden ser: computador, teléfonos móviles, gafas o cascos.
- Registro de objetos virtuales. – se refiere a la construcción de los objetos virtuales en 3D que serán visualizados mediante los dispositivos anteriormente mencionados en el mundo real.
- Interacción. - son las opciones que se le ofrece al usuario para poder manipular o modificar el objeto 3D.

La integración de estos tres subsistemas, permite al estudiante tener una mejor compresión e interacción con el contenido de la temática a tratar. Tomando en cuenta que numerosas investigaciones indican que al utilizar RA se refuerza el aprendizaje e incrementa la motivación por aprender (Reinoso, 2012).

La implementación de la RA se presenta en diferentes ámbitos o áreas de conocimiento, una de ellas es medicina en la que han existido varios estudios e implementaciones de nuevas tecnologías educativas (Ruiz-Parra, Angel-Müller, & Guevara, 2009) no obstante, se observa ausencia de estudios relacionados con la medicina legal. Varios autores O'Sullivan, Martin, & Murray (2000); Roberts, Bell, & Duffy (2006); Ypinazar & Margolis (2006) mencionan algunos aspectos claves por los que cada vez es necesario nuevas implementaciones tecnológicas como apoyo al proceso de enseñanza aprendizaje en el área de medicina general, entre ellos:

- Cuando un estudiante requiere acceder a prácticas hospitalarias o forenses depende mucho de la disponibilidad o existencia de casos de estudio, por ello no es posible una planificación de un calendario de prácticas, pues los casos de estudio son impredecibles(Ruiz-Parra et al., 2009).

- Actualmente las implicaciones éticas y legales son más exigentes por lo cual es complicado mantener las prácticas "en vivo". Además, es inevitable la perspectiva de las personas, frente a un estudiante de medicina y su notable inseguridad de ser atendido por un "inexperto"(Ypinazar & Margolis, 2006).
- Aumento de estudiantes y por ende la falta de sitios de prácticas y difícil acceso a los mismos (Ruiz-Parra et al., 2009).

Entre las diferentes implementaciones tecnológicas en el área de medicina han surgido las simulaciones y el aprendizaje virtual que permiten recrear escenarios del mundo real y facilitan la adquisición de habilidades y destrezas sin ser expuestos a riesgos ni limitaciones de espacio y tiempo. Las principales ventajas que presta esta tecnología son:

- Se puede utilizar el mismo simulador para varios escenarios con diferentes grupos de estudiantes(Ypinazar & Margolis, 2006).
- Mejora las destrezas y habilidades antes de enfrentarse a un caso real.
- Propicia la práctica continua y repetitiva (O'Sullivan et al., 2000).
- El aprendizaje es a partir de errores; es decir, cuenta con una retroalimentación que permite al estudiante conocer sus falencias y ensayar nuevas respuestas(Roberts et al., 2006).
- Ofrece la oportunidad de conocer equipos e instrumentos reales(Ruiz-Parra et al., 2009).

Se debe tomar en cuenta que la tecnología por sí sola no genera un aprendizaje significativo, es por ello que siempre debe ir acompañada de una metodología de enseñanza-aprendizaje (Cabero, 2015). Las metodologías que favorecen la adquisición de aprendizajes significativos, son aquellas que inducen al estudiante a hacer y en lo que hace (Bonwell, & Eison, 1991) a través de técnicas y estrategias que promuevan un aprendizaje experiencial. Según Kolb (1981) el aprendizaje experiencial se da en un ciclo de cuatro fases: experimentación concreta, observación reflexiva, conceptualización abstracta y testeo de implicaciones de conceptos en nuevas situaciones; las que pueden ser desarrolladas en entornos presenciales o simulados (Ruiz-Parra et al., 2009).

El proyecto de RA "Virtopsia", nace para responder a la necesidad de realizar prácticas en la titulación de Derecho de la Universidad Técnica Particular de Loja (UTPL), en la asignatura de Medicina Legal, tomando en cuenta que el acceso al centro forense es restringido y los estudiantes son, en su mayoría, de Modalidad Abierta y a Distancia. Antes de la implementación del proyecto los docentes trabajaban con el contenido que encontraban en la prensa y los recursos de libre de acceso disponibles en la web, lo que dificultaba poder explicar estos procesos biológicos a los estudiantes; por ello se encontró en la RA la posibilidad de simular autopsias de casos reales, de una forma didáctica desde su dispositivo móvil y de manera offline, sin necesidad de ir al centro forense. Se trabajó los 4 tipos de muertes más importantes, acoplando el contenido a la parte médico-legal, este aplicativo es de libre acceso para cualquier estudiante o institución educativa.

2. Metodología

En esta fase se utilizó la metodología para el desarrollo de Objetos de Aprendizaje (OA) en 3D para aplicaciones de RA propuesta por Samaniego-Franco, Jara-Roa, Sarango-

Lapo, Agila-Palacios, Guamán-Jaramillo, Contreras-Mendieta (2018); la que se planteó en base a un estudio comparativo de metodologías para el desarrollo de aplicaciones 3D. A continuación, se detallan cada una de las fases de la metodología:

2.1. Selección del tema

Para la selección del tema del aplicativo de RA se tomó en cuenta varios criterios: a) la necesidad de realizar prácticas y material de acceso libre, b) nivel de complejidad de la temática, c) la posibilidad de incrementar la motivación en el estudiante y d) el resultado de aprendizaje a lograr (Cueva Ludeña, 2018) "Capacidad para establecer los criterios médico-legales sobre la muerte, conocer las formas de la comprobación de la muerte, identificar el procedimiento y determinar los objetivos de la necropsia médico-legal."(p.17). El tema del aplicativo es autopsia o necropsia virtual, el que permitirá al estudiante conocer el procedimiento y determinar las causas de la muerte de la persona y la estructura de un informe médico-legal completo.

2.2. Contenido

En el plan de acción para la convergencia europea, de la Universidad Politécnica de Valencia, se señala que para poder construir un OA es necesario determinar el objetivo que se pretende alcanzar; y en base a ello se define el contenido a estudiar. El objetivo de aprendizaje del presente aplicativo es de carácter procedimental ya que los estudiantes realizarán prácticas en los diferentes tipos de muerte donde deberán

DIRECCIÓN DE TECNOLOGÍAS PARA LA EDUCACIÓN			
DOCENTE:	Rene cueva	**TITULAR:**	X
PERIODO:	Abril/Agosto 2018	**AUXILIAR:**	
TITULACION:	Derecho		
FECHA:	Diciembre 2017		
ASIGNATURA:	Medicina Legal	**TIPO:**	

RESULTADO DE APRENDIZAJE:

- Identificar el procedimiento de una autopsia.
- Determinar las características de diferentes muertes.
- Conocer la estructura del informe forense.

OBJETIVO DE RA:

Procedimental

PUBLICO META:

Estudiantes de derecho de la asignatura de Medicina Legal

OBJETO VIRTUAL DE APRENDIZAJE REALIDAD AUMENTADA (OA-RA)			
Tema:			
N°	**Nombre del objeto**	**Descripción**	**Formato**
Principal	Cadáver general	Cadáver que aparecerá en el marcador para iniciar el aplicativo	FBX
2	Muerte por accidente de tránsito	Tendrá las opciones de examen externo, interno, evaluación e incisiones	FBX, DOC
3	Muerte por arma de fuego	Tendrá las opciones de examen externo, interno, evaluación e incisiones	FBX, DOC
4	Muerte por arma blanca	Tendrá las opciones de examen externo, interno, evaluación e incisiones	FBX, DOC
5	Muerte por ahorcadura	Tendrá las opciones de examen externo, interno, evaluación e incisiones	FBX, DOC

Figura 1 – Plantilla de OA de granularidad gruesa (Plantilla contenido OA 3D UTPL, (Samaniego-Franco et al., 2018)

manejar y utilizar los cadáver e incisiones, verificar las características de los tipos de muerte y aplicar los conocimientos obtenidos en la navegación de las autopsias dentro de la evaluación del informe forense.

Como lo menciona (Wiley, 2003) el tamaño de un OA se puede categorizar según su granularidad y existe una relación inversa entre el tamaño del OA y su reutilización, es decir, mientras el tamaño del OA es menor o tiene granularidad fina la posibilidad de reutilización aumenta y, si la granularidad es gruesa o el tamaño del OA es mayor la posibilidad de reutilización disminuye. La figura 1 expone la plantilla del contenido con granularidad gruesa (diseño general del aplicativo) en donde se plantea resultados de aprendizaje, objetivo de RA, público meta y los temas generales a tratar.

2.3. Creación y diseño de la historia de interacción

Para la creación y diseño de la historia de la interacción es necesario el detalle del objeto con un nivel de granularidad fina, el mismo que es especificado en dos plantillas (figura 2 y figura 3), en las que se detalla las imágenes referenciales y las características de cada uno de los objetos que se indican en la plantilla de OA de granularidad gruesa (figura 1), que le permiten al diseñador tener una base sólida para el diseño del objeto.

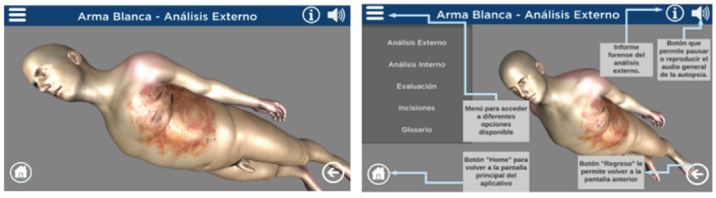
OBJETO VIRTUAL DE APRENDIZAJE REALIDAD AUMENTADA (OA-RA)					
Tema:	Hemotórax debido a trauma penetrante por arma blanca (Análisis externo)				
Nombre del modelado:	Muerte por arma blanca	**Tipo de modelado:**		**Estático (x)**	**Animado ( )**
Imagen referencial					
					
Nota: La imagen que se coloca es resultado de la virtualización; pero en la plantilla original deben ir las fotos fuente pero en este caso por el tipo de imágenes que son confidenciales se colocó el resultado de las imágenes ya finalizadas.					
Descripción general del diseño					
En el presente cadáver irá las características más importante que son: a) Lesiones en la cara, b) Lesiones en el tórax					
Desarrollo del contenido					
Icono identificativo	**Texto**	**Audio**	**Imagen/objeto3D**	**Video**	**URL**	**Contenido**
Menú de opciones	x					Irán las opciones que tiene el aplicativo: 1) análisis externo 2) análisis interno 3) evaluación 4) incisiones 5) glosario
Botón de informe forense	x					Informe forense con la estructura correspondiente
Botón de audio		x				Cadáver de sexo de 25 años de edad

Figura 2 – Plantilla de OA de granularidad fina (Plantilla adaptada de historia de interacción, (Samaniego-Franco et al., 2018)

1. **Plantilla historia de interacción**, en la figura 2 se detalla la historia de la interacción y contiene lo siguiente: 1) nombre del modelado, 2) tipo de modelado (en donde se especifica el grado de interacción de cada objeto), 3) imágenes referenciales que tendrán el bosquejo final del objeto (imagen base de las cuales se realizará el objeto 3D), 4) iconos identificativos, texto, audio y contenido.
2. **Plantilla contenido del objeto 3D,** en esta plantilla (figura 3) se detalla con una granularidad fina el contenido de cada objeto, es decir se describe el nombre de cada una de las muertes con las características de los daños encontrados (generales y específicos) de cada una, considerando la estructura del informe forense en el examen interno y externo.

PLANTILLA CONTENIDO DEL OBJETO 3D	
<TÍTULO DE LA MUERTE>	
Causa de la muerte:	
Tipo de muerte:	
Examen Externo	
•	
•	
Partes del cuerpo	**Descripción**
<Parte 1>	
<Parte 2>	
<subparte>	
<subparte>	
Examen Interno	
•	
•	
Partes del cuerpo	**Descripción**
<Parte 1>	
<Parte 2>	
<subparte>	
<subparte>	

Figura 3 – Plantilla de OA de granularidad fina del contenido del objeto 3D

Esta plantilla (figura 3) se utilizó para los 4 tipos de muerte, así mismo, contiene imágenes y una explicación detallada de las lesiones como apoyo para el diseñador de objetos 3D. en la sección de examen interno, se añade los órganos más comprometidos según el tipo de muerte para que el diseñador tenga en cuenta los órganos independientes a diseñar.

2.4. Desarrollo de objetos 3D

Para el desarrollo de objetos 3D el equipo de diseñadores tomó en cuenta los criterios más relevante mencionados por (Hippolyte, 2011; Saavedra, Arteaga, & Álvarez Rodríguez, 2007; Samaniego-Franco et al., 2018):

- Fácil manipulación.
- Minimización del número de vértices para poder maximizar el rendimiento en dispositivos de poca capacidad.
- Trabajar texturas como imagen para un adecuado reconocimiento y visualización.

- La ubicación del OA 3D en escena (plano (x, y, z)) debe ser tomada en cuenta desde la creación del OA 3D, para evitar la superposición de OA 3D entre sí. Esto es de vital importancia cuando se trabaja con múltiples objetos en un mismo marcador.

En base a lo antes mencionado se elaboró todos los objetos 3D por cada una de las muertes, en la figura 4 se muestran los objetos 3D de muerte por arma blanca:

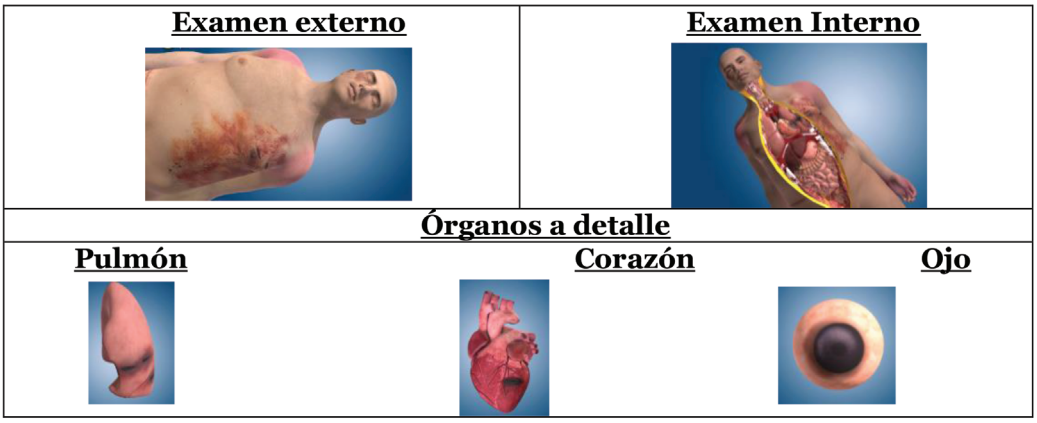


Figura 4 – Objetos 3D muerte por arma blanca

2.5. Desarrollo del aplicativo

En esta fase se desarrolla el aplicativo final, es decir se unifica el contenido de todas las plantillas; adicional a esto se debe plantear la plataforma tecnológica y el proceso interno que tendrá el aplicativo.

Plataforma tecnológica. - la plataforma tecnológica que se utilizó para el aplicativo "Virtopsia" es la siguiente:

- Virtopsia AR Android: Android API Level 25 (Android 4.1 a 7.1.1) y Unity 5.6

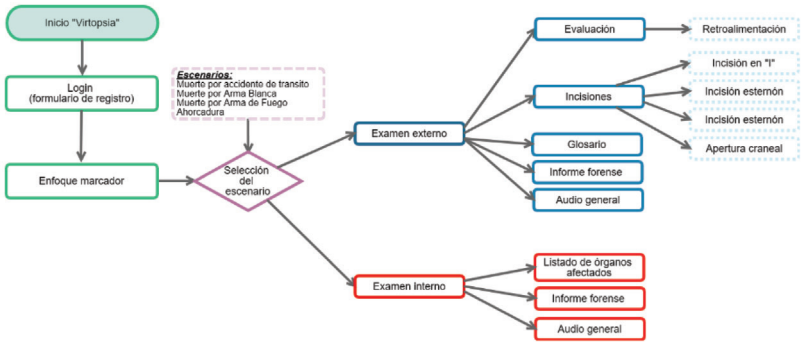


Figura 5 – Vista del proceso

- Virtopsia AR IOs: C Sharp (C#) y Unity 5.6
- Base de Datos: MySQL
- Lenguaje de programación: PHP

Proceso utilizado. – la figura 5 detalla el proceso que llevará internamente el aplicativo para poder conocer exactamente las actividades que puede o no realizar el usuario.

El proceso planteado dentro del aplicativo le permite al estudiante analizar varías autopsias y dentro de cada una navegar en el examen externo e interno, acceder a una evaluación con retroalimentación, glosario, tipos de incisiones y exploración de cada órgano afectado, según el tipo de muerte; el proceso del aplicativo es completo, permite al estudiante alcanzar los objetivos de aprendizaje planteados.

La figura 6 muestra la versión final del aplicativo a nivel general, contiene una ventana donde al solicitar enfocar el marcador se observará el cadáver con cada uno de los tipos de muerte.

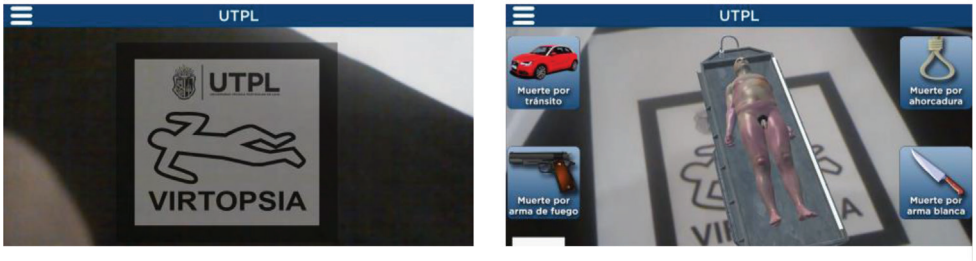


Figura 6 - Imágenes aplicativo final

2.6. Publicación en repositorios

Todos los objetos son desarrollados de manera individual, por ende, cada objeto es publicado en la cuenta de la UTPL en el repositorio de objetos 3D **"Sketchfab"**; con el objetivo de poder compartir recursos educativos abiertos a la comunidad. Para la publicación de los objetos se tomó en cuenta lo siguiente:

1. Título identificativo del objeto.
2. Archivo fbx del objeto.
3. Licencia institucional Creative Commons "Atribución, no comercial y compartir igual".
4. Permiso de descarga.
5. Restricción de edad para personas mayores a 18 años (considerando las políticas para la publicación de imágenes con contenido fuertes).

La figura 7 muestra un objeto publicado en el repositorio Sketchfab:

2.7. Plan de difusión y capacitación

El plan de difusión y capacitación debe contribuir a aumentar el uso de los aplicativos de RA y brindar ayuda para el manejo de los mismos a través de un curso de soporte en el Entorno Virtual de Aprendizaje (EVA) para el uso de aplicativos con RA.

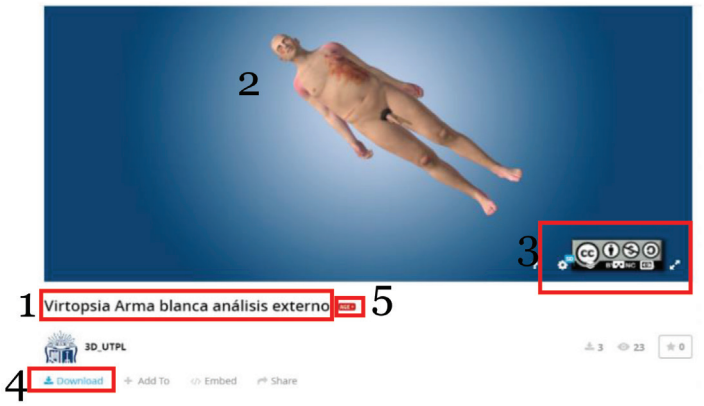


Figura 7 - Objeto 3D en repositorio Sketchfab

3. Implementación en el aula

Para la implementación en el aula, se adaptó la metodología del ciclo de aprendizaje experiencial de Kolb, cuyas fases se describen a continuación:

4. Nivel de aceptación del aplicativo Virtopsia

Para medir el nivel de aceptación del aplicativo Virtopsia se elaboró un cuestionario de 26 items, categorizado en 4 áreas: contenido, usabilidad, soporte y recomendaciones. Para el presente estudio se consideró las dos primeras áreas que están en relación directa con el nivel de aceptación. El instrumento se diseñó con escala de Likert, donde 1 significa muy en desacuerdo y 4 muy de acuerdo. El nivel de confiabilidad de Alpha de Cronbach = 0.87.

Los estudiantes que utilizaron el aplicativo fueron 450, de los cuales 110 (24,4%) respondieron la encuesta de forma completa al finalizar el segundo bimestre.

De acuerdo a los resultados obtenidos, se puede observar que el 82% de estudiantes en promedio, están entre muy de acuerdo y de acuerdo con las primeras 10 preguntas, esto indica que el aplicativo ha brindado el apoyo esperado en la asignatura, el 18% que está en desacuerdo, según las recomendaciones indicadas, se debe a la dificultad en el uso del aplicativo, por la ausencia de manuales o videotutoriales.

El 50% de estudiantes indican haber utilizado el aplicativo de 2 a 3 veces a la semana, el 31% una vez a la semana y el 19 % una vez al mes, estos resultados afirman el cumplimiento del trabajo planteado con el aplicativo en la mayoría de estudiantes.

5. Conclusiones

- Con la presente investigación se corrobora lo indicado por (Ruiz-Parra et al., 2009) cuando menciona que las nuevas tecnologías, en especial la RA, son un apoyo valioso en el proceso enseñanza-aprendizaje.

Fase	Descripción	Herramienta / Actividades
Experimentación concreta	Esta fase se experimenta desde el momento en el que el participante luego de haber descargado el aplicativo, ingresa al mismo relacionándose con cada uno de los casos de muerte. La experimentación se facilita a través del texto, audio e interacción con el cuerpo.	App Virtopsia
Observaciones y reflexiones	En el mismo proceso de experimentación el estudiante observa y reflexiona sobre los elementos que visualiza, la relación entre ellos, dependencia o independencia, detalles sobre sus ideas iniciales y construye las primeras aproximaciones a los conceptos. Todo esto en el primer tipo de muerte: por tránsito	Registro de las observaciones en un documento, y respuesta a preguntas planteadas por el profesor para promover la reflexión; planteamiento de nuevas preguntas temas o problemas que necesitan nuevas aportaciones para su respuesta o resolución. y, la socialización a través del foro académico.
Formación de conceptos abstractos y generalizaciones	Basada en la experiencia de los objetos 3D y la interacción con las opciones dentro del aplicativo, como informe y descripciones de los objetos, los estudiantes van formando sus conceptos abstractos planteando una hipótesis en torno a una autopsia, en un caso específico de una muerte (muerte de tránsito, ahorcadura, arma de fuego y arma blanca). Para dar respuesta a esa hipótesis y a las interrogantes resultantes de la fase anterior, es necesario la revisión de los contenidos de la asignatura.	Envío de la matriz de relación entre observaciones, interrogantes y conceptos o contenidos teóricos. Aquí se realiza la actividad de tarea.
Testeo de implicaciones de conceptos en nuevas situaciones.	En esta fase a través de su experiencias realizando la autopsia virtuales del primer tipo de muerte, debe estar en la capacidad de elaborar en forma adecuada un informe de la autopsia y de poder identificar diferentes características en otro tipo de muerte	Evaluación incluida en la aplicación de RA Cuestionario con preguntas que promuevan la reflexión en otro tipo de muerte de la app. De esta manera se inicia nuevamente el ciclo de aprendizaje experiencial, además la discusión de nuevos casos planteados por el docente mediante una videocolaboración.

Tabla 1 – Implementación de fases de Kolb

Tema	Pregunta	Opciones de Respuesta	F	%
Contenido	1. Las incisiones, glosario y evaluaciones le han ayudado con su aprendizaje.	Muy de acuerdo/de acuerdo	97	**88**
		Muy en desacuerdo/en desacuerdo	13	12
	2. Los objetos 3D y contenidos, le ayudaron a reconocer las características de muerte en diferentes circunstancias.	Muy de acuerdo/de acuerdo	**94**	**86**
		Muy en desacuerdo/en desacuerdo	16	14
	3. Considera que los contenidos desarrollados en el aplicativo son más llamativos que los señalados en la guía didáctica y el libro base.	Muy de acuerdo/de acuerdo	**94**	**86**
		Muy en desacuerdo/en desacuerdo	16	14
	4. Las explicaciones de audios lo orientan a explorar de mejor manera los diferentes casos de muerte.	Muy de acuerdo/de acuerdo	**84**	**85**
		Muy en desacuerdo/en desacuerdo	16	15
	5. La evaluación de cada caso le ayudó a conocer las partes fundamentales de un informe forense.	Muy de acuerdo/de acuerdo	**91**	**83**
		Muy en desacuerdo/en desacuerdo	19	17
	6. ¿El producto de aprendizaje (Matriz observaciones), solicitado le ayudó a relacionar la teoría(concepto) con la práctica (aplicativo "Virtopsia UTPL")?	Muy de acuerdo/de acuerdo	**88**	**80**
		Muy en desacuerdo/en desacuerdo	22	20
	7. Las retroalimentaciones de las evaluaciones le ayudaron a solventar sus dudas e inquietudes	Muy de acuerdo/de acuerdo	**87**	**79**
		Muy en desacuerdo/en desacuerdo	23	21
Usabilidad	8. ¿Al hacer uso por primera vez del aplicativo lo incentivó a revisar todos los temas tratados dentro del mismo?	Muy de acuerdo/de acuerdo	88	**80**
		Muy en desacuerdo/en desacuerdo	22	20
	9. Las orientaciones que el docente realizó para el uso del aplicativo fue adecuada	Muy de acuerdo/de acuerdo	**84**	**85**
		Muy en desacuerdo/en desacuerdo	16	15
	10. Cree usted que el uso de la aplicación "Virtopsia UTPL" le permite realizar autopsias como si estuviera en un entorno real.	Muy de acuerdo/de acuerdo	**75**	**68**
		Muy en desacuerdo/en desacuerdo	35	32
	11. ¿Cuántas veces considera que usó el aplicativo en el transcurso del II bimestre?	Más de 3 veces a la semana	19	17
		1 a 3 veces a la semana	36	33
		1 vez a la semana	35	31
		1 vez al mes	21	19

Tabla 2 - Tabla de resultados

- Como menciona (García Aretio, 2003) las nuevas tecnologías, en este caso RA, ayuda a eliminar las diferencias entre la Modalidad Presencial y la Modalidad Abierta y a Distancia permitiendo una experiencia inmersiva con casos de la vida real.
- La implementación de la metodología de Kolb más el aplicativo ayudó al estudiante a relacionar la parte teórica con la práctica en un ciclo de aprendizaje experiencial, de acuerdo a lo señalado por Cabero (2015).
- De acuerdo a los resultados, podemos afirmar lo mencionado por (Reinoso, 2012), ya que el aplicativo resultó llamativo e intuitivo motivando al estudiante a revisar todo el contenido.
- La presente investigación puede tomarse como referencia para futuros trabajos en cuanto a implementación de RA en un sistema de estudios a distancia.
- A futuro es necesario ampliar las investigaciones en cuanto a los resultados del ciclo de aprendizaje experiencial de Kolb con aplicaciones de RA.

Referencias

Agudelo, A. (2005). *Modelo de contexto para Realidad Aumentada. Revista Universidad EAFIT*. Retrieved from http://publicaciones.eafit.edu.co/index.php/revista-universidad-eafit/article/view/856/762

Bates, A. W. (1999). La tecnología en la enseñanza abierta y la educación a distancia, 37–51.

Billinghurst, M., & Dünser, A. (2012). Augmented Reality in the Classroom.

Bonwell, C., y Eison, J. (1991). Active Learning: Creating Excitement in the Classroom. ASHE-ERIC Higher Education Report.

Cabero, A. J. (2015). Reflexiones educativas sobre las tecnologías de la información y la comunicación (TIC), 20–22. Retrieved from https://idus.us.es/xmlui/bitstream/handle/11441/32285/Reflexiones_educativas_sobre_las_Tecnolo.pdf?sequence=1

Cueva Ludeña, R. (2018). No Title. Retrieved from https://utpl.instructure.com/courses/602/files/140426/download?download_frd=1

Dunleavy, M., Dede, C., & Mitchell, R. (2009). Affordances and limitations of immersive participatory augmented reality simulations for teaching and learning. *Journal of Science Education and Technology*, *18*(1), 7–22. https://doi.org/10.1007/s10956-008-9119-1

García Aretio, L. (2003). La Educación a Distancia, (May), 18. Retrieved from http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1024-94352003000100002

Hippolyte, P. (2011). *Técnicas de modelado 3D en Sketchup, para el manejo de modelos urbanos complejos dentro de Google Earth*. Retrieved from https://www.fau.ucv.ve/trienal2011/cd/documentos/irg/IRG-1.pdf

Johnson, L., Adams Becker, S., Cummins, M., Estrada, V., Freeman, A., & Hall, C. (2016). *The NMC Horizon Report: Edición Educación Superior 2016. 50 NMC Horizon Report: Edición Educación Superior 2016*. Retrieved from http://cdn.nmc.org/media/2016-nmc-horizon-report-HE-ES.pdf

Kolb, David a. (1981): "Learning styles and disciplinary differences", en chickering, Arthur W. (ed.), the modern American college, jossey- bass publishers, San Francisco, pp. 232-255.

O'Sullivan, M., Martin, J., & Murray, E. (2000). Students' perceptions of the relative advantages and disadvantages of community-based and hospital-based teaching: A qualitative study. *Medical Education*, *34*(8), 648–655. https://doi.org/10.1046/j.1365-2923.2000.00623.x

Reinoso, R. (2012). *Tendencias emergentes en Educación con TIC*. (E. y T. Asociación Espiral, Ed.) (1era ed.). Barcelona. https://doi.org/10.7203/attic.12.3544

Roberts, K. E., Bell, R. L., & Duffy, A. J. (2006). Evolution of surgical skills training. *World J Gastroenterol*, *12*(20), 3219–3224. Retrieved from www.wjgnet.comhttp://www.wjgnet.com/1007-9327/12/3219.asp

Ruiz-Parra, I., Angel-Müller, E., & Guevara, O. (2009). La simulación clínica y el aprendizaje virtual. Tecnologías complementarias para la educación médica. *Revista Facultad de Medicina …*, *57*, 67–79. Retrieved from http://www.scielo.co/scielo.php?script=sci_arttext&pid=S0120-00112009000100009&lng=es&nrm=iso

Saavedra, A. B., Arteaga, J. M., & Álvarez Rodríguez, F. J. (2007). Modelo Instruccional para el Diseño de Objetos de Aprendizaje: Modelo MIDOA. Retrieved from http://e-spacio.uned.es/fez/eserv/bibliuned:19245/n03barajas07.pdf

Samaniego Franco, J. B., Jara Roa, D. I., Sarango Lapo, C. P., Agila, M. V., Guamán Jaramillo, J. E., & Contreras Mendieta, J. A. (2018). "Case study: Methodology for the development of learning objects (OA) in 3D for applications of augmented reality (AR)," 2018 13th Iberian Conference on Information Systems and Technologies (CISTI).

Wiley, D. A. (2003). Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy. Retrieved from http://www.reusability.org/read/chapters/wiley.doc

Ypinazar, V. A., & Margolis, S. A. (2006). Clinical simulators: applications and implications for rural medical education. *Rural and Remote Health [Electronic Resource].*, *6*(2), 527.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 22/09/2018
Aceitação/Acceptance: 06/11/2018

Métricas que podrían usarse en el Desarrollo de Aplicaciones Móviles para Personas con deficiencias Visuales: Una Revisión de Literatura Preliminar

Erick Zambrano[1], Rolando P. Reyes Ch.[1], John W. Castro[2], Efraín R. Fonseca C.[1]

ejzambrano, rpreyes1, erfonseca@espe.edu.ec, john.castro@uda.cl

[1] Universidad de las Fuerzas Armadas ESPE, Departamento de Ciencias de la Computación, Av. General Rumiñahui s/n y Ambato, EC171103, Sangolquí, Ecuador.

[2] Universidad de Atacama, Departamento de Ingeniería Informática y Ciencias de la Computación, Av. Copayapu 485, 1530000, Copiapó, Chile.

Pages: 985–999

Resumen: Antecedentes: La masificación del uso de dispositivos móviles ha propiciado el desarrollo paralelo de un sinnúmero de aplicaciones móviles (apps), cuyas características son explotadas por la mayoría de las personas. No obstante, existen segmentos de personas con deficiencias visuales (PDV) que han sido marginadas por temas de usabilidad y/o accesibilidad. **Objetivo:** Identificar estudios primarios que reporten métricas de usabilidad y/o accesibilidad aplicable al desarrollo y evaluación de sistemas móviles, incluyendo las necesidades de las PDV. **Método:** Se llevó a cabo una Revisión de Literatura Preliminar (PLR) basada en las directrices de una Revisión Sistemática de Literatura (SLR). **Resultados:** Se obtienen 6 estudios primarios que reportan 73 métricas para apps, las cuales fueron agrupadas en 8 contextos de usabilidad: Contenido visual, eficacia, navegabilidad, eficiencia, aprendizaje, satisfacción, comprensión y accesibilidad. **Conclusión:** Los contextos de mayor relevancia son: Contenido visual, eficacia, navegabilidad, comprensión y accesibilidad, siendo los 3 primeros los que agrupan la mayor cantidad de métricas de usabilidad.

Palabras-clave: Usabilidad; Accesibilidad; Métricas; Ingeniería de Software; Revisión de Literatura

Metrics That Could Be Used In Mobile Applications Development For People With Visual Impairments: A Preliminary Literature Review

Abstract: **Background:** The widespread use of mobile devices has motivated the parallel development of countless mobile applications (apps), whose characteristics benefit most people. However, there are segments of people with visual impairments (PVI) who have been marginalized due to usability and accessibility issues. **Aim:** To identify primary studies which report metrics of usability, accessibility or both, applicable to the development and assessing of mobile systems, including requirements of PVI. **Method:** A Preliminary Literature Review (PLR) was done

based on the guidelines of a Systematic Literature Review (SLR). **Results:** 6 primary studies reported 73 metrics for apps, which were grouped into 8 usability contexts: visual content, effectiveness, navigability, efficiency, learnability, satisfaction, understandability and accessibility. **Conclusions:** The most relevant contexts are Visual content, effectiveness, navigability, understandability, and accessibility; of which, the first three group the highest number of usability metrics.

***Keywords**:* Usability; Accessibility; Metrics; Software Engineering; Literature Review.

1. Introducción

La masificación del uso de dispositivos móviles para realizar tareas cotidianas desde cualquier lugar a través de internet ha crecido exponencialmente en los últimos años (Dias Silva, Carvalho Nepomuceno, Silva, & Freire Pimenta, 2018). De hecho, se espera que en 2019 la cantidad de usuarios de teléfonos móviles a nivel mundial supere los cinco mil millones, lo que representa aproximadamente el 67 % de la población mundial (Baron, 2018). Con este fin, se han desarrollado paralelamente un sinnúmero de aplicaciones móviles (apps), las cuales son ampliamente utilizadas por la mayoría de los usuarios. No obstante, segmentos de usuarios con deficiencias físicas han sido marginados, dado que enfrentan dificultades al utilizar los sistemas móviles, especialmente las personas con deficiencias visuales (PDV) (Dias Silva, Carvalho Nepomuceno, Silva, & Freire Pimenta, 2018).

Las PDV enfrentan una serie de retos al momento de interactuar con un sistema móvil, principalmente relacionados con la usabilidad y accesibilidad referidas al software y hardware respectivamente (Carvalho & Freire, 2017). En lo que respecta al software, está comprobado que, durante el desarrollo de la mayoría de las apps, no son considerados criterios ni parámetros referentes a la usabilidad para las PDV; mientras que, en lo relacionado con el hardware, los dispositivos móviles presentan una gran variedad de características (tamaños variados, colores diversos, etc.), sien-do la más común el uso de pantallas táctiles como medio de interacción (TIC, 2013). Esta situación dificulta e incluso imposibilita el uso de ciertas apps; así como también, reduce significativamente la interacción PDV - sistema móvil (Kane, 2009).

La usabilidad es un atributo de calidad muy importante dentro del desarrollo de software; de hecho, existe una innumerable cantidad de estudios al respecto (P. Di , y otros, 2012). No obstante, los estudios de usabilidad enfocados en las apps han llamado la atención recientemente, al parecer por el incremento en su uso. Para el caso de las PDV, la importancia de la usabilidad se acrecienta aún más, por obvias razones. Sin embargo, el número y la madurez de los estudios al respecto nos indican que, al parecer, no se está considerando a las PDV al momento de construir y evaluar la usabilidad de las apps, limitando así la inclusión de este segmento de la población en la sociedad (Facanha Rocha, Cavalcante, & Campos, 2014).

Esta situación ha puesto de manifiesto la necesidad de identificar trabajos en los que se estudien métricas de usabilidad y accesibilidad que podrían ser aplicadas en el desarrollo y evaluación de los sistemas móviles, incluyendo las necesidades de las PDV. Como estrategia metodológica para obtener los estudios requeridos, consideramos

la realización de una Revisión de Literatura Preliminar (PLR de las siglas del inglés Preliminary Literature Review). La propuesta de un estudio preliminar responde a dos situaciones: por una parte, a optimizar el tiempo que normalmente es utilizado en una Revisión Sistemática de Literatura (SLR de las siglas del inglés Systematic Literature Review); y por otra parte, la necesidad de obtener de manera rigurosa estudios fiables. Como resultado, se obtuvieron 6 estudios primarios en los que se hacen referencia a contextos de métricas de: contenido visual (por ejemplo, visual coherence), eficacia (por ejemplo, ease of use), navegabilidad (por ejemplo, display self adjustment), satisfacción (por ejemplo, reach goal accurately), comprensión (por ejemplo, display load) y accesibilidad (por ejemplo, information access easily). Estos contextos de métricas fueron considerados idóneos para evaluar la usabilidad de apps que podrían ser utilizadas por las PDV, dada su afinidad con las características de los diferentes tipos de deficiencias visuales.

2. Antecedentes

La usabilidad es considerada como una característica de calidad muy importante dentro de la Ingeniería de Software (IS). Esta característica normalmente suele ser considerada durante el desarrollo de las aplicaciones de escritorio y web (Carvajal, Moreno, & M. Segura, 2013). No obstante, en los últimos años el desarrollo se ha orientado hacia las apps, debido al incremento exponencial en el uso de los dispositivos móviles (Joorabchi, Mesbah, & Kruchten, 2013). Ahora bien, evaluar la usabilidad de las apps, no es una tarea fácil. Los ingenieros de software necesitan conocer exactamente qué es lo que hay que medir, a partir de un sinnúmero de dominios que han sido propuestos en la literatura (Baharuddin , Singh , & Razali , 2013). Razón por la cual, es frecuente la selección errónea de métricas de usabilidad para las apps, dado que inicialmente las métricas fueron definidas para aplicaciones de escritorio o web. El problema radica en que los sistemas móviles presentan características particulares como: dimensiones reducidas, ancho de banda limitado, tamaño de la pantalla, entre otros (Hanlu , Malu, & Findla, 2014).

La literatura que trata la usabilidad es muy extensa. Hemos clasificado la literatura relacionada con métricas de usabilidad en dos grupos. En el primero, se encuentran los trabajos que identifican métricas para aplicaciones de escritorio o web, como el estándar ISO/IEC 9126 que define métricas de usabilidad. Sagar and Saha (Anju & Sagar, 2017) realizan una SLR para identificar características de usabilidad generales (por ejemplo: estándares, métodos, métricas, dominios de usabilidad). Los autores identificaron 150 estudios primarios, a partir de los cuales concluyen que no existe un modelo consensuado para evaluar la usabilidad en diferentes dominios. El trabajo de Sauro y Lewis (Sauro & Lewis R., 2009) reporta factores que pueden incidir en la correlación entre distintas métricas de usabilidad, a partir de una base de información constituida de 90 estudios de usabilidad.

El segundo grupo está relacionado con la evaluación de la usabilidad de las apps. Ahmad & otros (Ahmad, Rextin, & Kulsoom, 2018) logran sintetizar una lista de guías relacionadas con la usabilidad para el desarrollo de aplicaciones multiplataforma en dispositivos móviles. Los autores obtienen 148 estudios primarios, que proponen un total

de 359 guías que fueron organizadas en 7 categorías. Además, aplican un experimento controlado con estudiantes de grado y posgrado con el objetivo de evaluar algunas apps y detectar problemas de usabilidad. Kumar y Mohite (Kumar & Mohite, 2018) realizan una SLR en la que identifican 23 estudios primarios relacionados con la evaluación de la usabilidad de apps para el aprendizaje. Lettner y Holzmann (Lettner & Holzmann, 2012) proponen un nuevo enfoque, de la mano de un conjunto de herramientas para la evaluación automatizada y no supervisada de apps. Estas herramientas permiten identificar y visualizar fallos en el diseño (por ejemplo, errores en la navegación o eficiencia). Kronbauer & otros (Kronbauer H & Santos S., 2012) proponen un modelo de evaluación de usabilidad híbrido para apps, el cual describe cómo supervisar y recopilar automáticamente datos de contexto y métricas de usabilidad.

En resumen, podemos indicar que existen trabajos relacionados con las métricas de usabilidad para aplicaciones de escritorio o web y otros relacionados con las métricas de usabilidad para apps. No obstante, no encontramos estudios sobre métricas de usabilidad para apps que consideren las necesidades de las PDV, lo cual nos pareció extraño, dado que según la Organización Mundial de la Salud (OMS) (World Health Organization, 2017), actualmente existen 314 millones de personas en el mundo que tienen discapacidades visuales y que por lo antes mencionado son excluidos del uso de la tecnología (Kulpa Costa & Amaral Gon). Por tanto, se hace necesario investigar que métricas de accesibilidad o usabilidad podrían ser aplicadas durante el desarrollo de apps para las PDV. Para tal efecto, nos hemos planteado la siguiente pregunta de investigación (RQ):

RQ: ¿Qué métricas de accesibilidad y/o usabilidad, consideradas comúnmente en la construcción del software, podrían ser utilizadas para evaluar las apps desde la perspectiva de las PDV?

3. Método de Investigación

Para dar respuesta a nuestra RQ realizamos una PLR como estrategia metodológica para la cual consideramos utilizar algunas fases, etapas y actividades propuestas por Kitchenham & otros (Kitchenham, Brereton, O , & Linkman, 2009) para realizar una SLR, dada su representatividad en este tipo de estudios en la IS. Más específicamente, consideramos: planificar la revisión y ejecutar la revisión. De la fase planificar la revisión consideramos las etapas: (1) identificar la necesidad de una revisión, (2) especificar la pregunta de investigación y (3) desarrollar el protocolo de investigación. Estas etapas permitieron orientar nuestra PLR. Para la fase ejecución de la revisión seleccionamos las etapas: (1) identificación de la investigación, (2) selección de estudios, (3) evaluación de calidad de los estudios y (4) extracción de datos y síntesis. Esta última etapa es una fusión de las etapas extracción de datos y síntesis de datos. Esta fusión responde a que posiblemente los datos resultantes no son extensos y por tanto, permitan una síntesis mayormente descriptiva (Velásquez, 2015).

3.1. Desarrollo del Protocolo e Identificación de la Investigación

Las actividades realizadas en esta etapa son: (1) Conformación del grupo de control, (2) identificación y selección de las palabras claves, (3) especificación de los criterios

de inclusión y exclusión y (4) conformación de la cadena de búsqueda. A continuación, describimos cada una de estas etapas.

***Conformación del Grupo de Control*:** Según Zhang & otros (Zhanga, Ali Babarb, & Tellb, 2010), un grupo de control es un conjunto de estudios relevantes que permiten establecer otros estudios similares bajo criterios explícitos establecidos en las RQ. Desde esta perspectiva, nuestro grupo de control de literatura nos permitió identificar y seleccionar las palabras comúnmente utilizadas en el contexto de la temática de investigación, las cuales constituyeron la base para conformar de la cadena de búsqueda. Así mismo, el grupo de control nos permitió afinar los criterios de inclusión y exclusión que nos habíamos planteado inicialmente para definir las características de los estudios que buscamos. Para la conformación del grupo de control son necesarios 4 pasos (Zhanga, Ali Babarb, & Tellb, 2010).

- Identificación de Investigadores. Contamos con un investigador novato y un investigador experto con una amplia experiencia en la ejecución de revisiones sistemáticas de literatura. Estos investigadores se encargaron de conformar el grupo de control.
- Búsqueda de Artículos Candidatos para el Grupo de Control. De forma individual, los investigadores buscan artículos aparentemente idóneos para el grupo de control. Esta búsqueda fue realizada en bases de datos digitales tales como Scopus, ACM, IEEE, entre otras (Kitchenham, Brereton, O , & Linkman, 2009). Posteriormente, los investigadores examinaron el título, resumen y palabras claves, guiados por los criterios de inclusión y exclusión preliminares. Como resultado, un total de 27 artículos fueron propuestos para la conformación del grupo de control (17 por el investigador novato y 10 por el investigador experto).
- Validación y Conformación del Grupo de Control. Una vez obtenidos los artículos candidatos para conformar el Grupo de Control, se realiza una validación cruzada entre los investigadores. Esta validación consiste en la revisión del texto completo de cada uno de los 27 artículos candidatos, con el propósito de verificar que los artículos tengan una relación directa con la RQ. Como resultado de la validación cruzada se conformó el Grupo de Control con los siguientes artículos: (Campanella, 2012), (Ba Matraf1 & Hussain, 2018), (Molano I., Yara S. , & García K. , 2015), (FortesPontin , Salgado, de Souza Santos, Agostini do Amaral, & Nogueira da Silva , 2017), (Cruz Reis, Gomes, Dias-Neto, & de L. Fontão, 2015), (Baumel, Faber, Mathur, Kane, & Muench, 2017), (Sánchez Sierra & Togores, 2012), (Guerrero & Vega, 2018), (Zali & Ahmad Fadzlah , 2016), (Kumanan , y otros, 2016), (Olmedo-Moreno & López-Delgado , 2015). En https://goo.gl/nmYxZw puede consultarse el detalle de la validación cruzada y la selección del grupo de control.

Identificación y Selección de las Palabras Claves: En los artículos del grupo de control identificamos las palabras que aparecen con mayor frecuencia (Bolderston, 2008) con la ayuda del software ATLAS.ti[1], lo cual puede ser revisado en detalle en[2]. La **Tabla 1** presenta, para cada una de las 10 palabras con mayor frecuencia, el número de veces que es usada en el grupo de control y la cobertura que consiste en el porcentaje

[1] https://atlasti.com/
[2] https://goo.gl/nmYxZw

de artículos del grupo de control donde aparece la palabra. Las palabras seleccionadas fueron consideradas para conformar y afinar la cadena de búsqueda.

Palabras	Conteo	Cobertura	Palabras	Conteo	Cobertura
usability	451	90 %	quality	149	64 %
mobile	378	90 %	applications	130	72 %
evaluation	227	90 %	application	118	55 %
accessibility	191	72 %	model	103	55 %
app	154	36 %	metrics	74	45 %

Tabla 1 – Palabras con Mayor Frecuencia en el Grupo de Control

Conformación de la Cadena de Búsqueda: Con base en las palabras seleccionadas se construyó la cadena de búsqueda luego de varios ciclos de afinamiento. Para la construcción de la cadena de búsqueda se utilizó el operador lógico AND para incluir palabras que pertenecen a contextos diferentes y el operador OR para incluir sinónimos de palabras que pertenecen al mismo contexto. El afinamiento consiste en probar distintas cadenas variando su contenido de palabras a modo prueba-error, con el propósito de obtener:

- Un número razonable de estudios (cientos no miles)
- Un alto porcentaje de estudios que se relacionen directamente con la temática de investigación
- La mayoría de los estudios del Grupo de Control dentro de la respuesta de la cadena
- En el proceso de afinamiento se generaron un total de 6 cadenas de búsqueda hasta encontrar la cadena ideal. Este proceso se realizó en la base digital Scopus. Para validar la cadena de búsqueda más adecuada, se utilizaron diferentes ratios:
- Ratio-X (ERGC/TER): Orientado a establecer el ratio de los artículos positivos - positivos recuperados del total de artículos encontrados con la cadena de búsqueda. Este ratio se calcula dividiendo la cantidad de estudios recuperados del grupo de control (ERGC) con la cadena X sobre el número total de estudios recuperados con la cadena X (TER).
- Ratio-Y (ERGC/GC): Establece el ratio de los artículos del grupo de control recuperados. Este ratio se calcula dividiendo la cantidad de estudios recuperados del grupo de control (ERGC) con la cadena X sobre el número total de estudios del Grupo de Control (GC).

Con base en el cálculo de estos ratios, se logró determinar que la cadena C6 (Ratio-X = 0.09, Ratio-Y = 0.64) es la más adecuada para encontrar los estudios primarios. Las cadenas afinadas (C1-C5) y la cadena C6 se muestran en la **Tabla 2**.

***Especificación de los Criterios de Inclusión y Exclusión*:** A partir de las características y particularidades de los artículos del grupo de control, los investigadores afinaron los criterios de inclusión y exclusión. Los criterios de inclusión corresponden a artículos que:

	Search Strings	Recov. studies	Ratio-X	Cont. Group	Ratio-Y
C1	usability AND metrics AND software AND quality AND model	88	0.02	[25], [30]	0.18
C2	usability AND metrics AND software AND quality AND mobile AND apps	3	0.67	[25], [32]	0.18
C3	usability AND metrics AND mobile AND (applications OR app OR apps) AND design AND development	17	0.12	[27], [28]	0.18
C4	(accessibility OR usability) AND metrics AND mobile AND (apps OR app OR applications) AND design AND development	19	0.10	[27], [28]	0.18
C5	accessibility AND metrics AND mobile	53	0.05	[23], [24], [26]	0.27
C6	(accessibility OR usability) AND metrics AND evaluation AND mobile AND (apps OR app OR applications)	78	0.09	[23], [24], [25], [27], [28], [31], [32]	0.64

Tabla 2 – Cadenas de Búsqueda Piloteadas y sus Ratios Asociados

- Describen métricas de accesibilidad o usabilidad utilizadas en el desarrollo de apps.
- Proponen métricas de usabilidad para el desarrollo de apps.
- Reportan métricas de usabilidad para personas con discapacidad.
- Han sido publicados en revistas, congresos, o en capítulos de libros (indexa-dos por lo menos en Scopus).
- Los criterios de exclusión son los siguientes:
- Artículos que no describen o no presenten información sobre métricas de accesibilidad o usabilidad utilizadas durante el desarrollo de apps.
- Artículos que reporten casos de estudio o metodologías para el diseño de apps.
- Artículos escritos en idiomas diferentes al inglés.

3.2. Selección de Estudios

La búsqueda de estudios fue realizada en las bases digitales IEEE Xplore[3], Scopus[4] y ACM Digital Library[5] usando la cadena **C6**. En una primera instancia, logramos encontrar un total de 108 estudios candidatos distribuidos de la siguiente manera: 78 estudios candidatos en Scopus, 3 estudios en IEEE Xplore y 27 estudios en ACM Digital Library. Posteriormente, eliminamos estudios duplicados entre las bases de datos. De esta depuración se descarta 1 estudio candidato de IEEE Xplore y 10 estudios candidatos de ACM Digital Library. Como resultado de esta primera depuración se obtienen 97 estudios candidatos.

[3] http://ieeexplore.ieee.org/
[4] https://www.scopus.com/
[5] https://dl.acm.org/

A paso seguido realizamos (investigadores novato y experto) un primer filtro aplicando los criterios de inclusión y exclusión sobre el título, resumen y palabras claves de cada uno de los 97 estudios candidatos. Como resultado de este primer filtro se obtuvieron 14 estudios (83 estudios candidatos fueron descartados). Los estudios obtenidos a partir del primer filtro fueron evaluados nuevamente en un segundo filtro, lo cual se describe a continuación.

3.3. Evaluación de Calidad de los Estudios

Los 14 estudios obtenidos a partir del primer filtro fueron sometidos a un segundo filtro. En este filtro interviene un tercer investigador, el cual, participa con fines de rigurosidad en la validación cruzada. Previo a la validación, cada investigador aplica los criterios de inclusión y exclusión al texto completo de cada uno de los 14 estudios.

Como resultado del segundo filtro se obtienen 6 estudios primarios (ver **Tabla 3**), que hemos codificado como "SP#". En el enlace https://goo.gl/nmYxZw se encuentra el detalle del proceso de selección de estudios primarios.

Cod.	Año	Título	Ref.
SP1	2001	QUIM: A Framework for Quantifying Usability Metrics in Soft-ware Quality Models	[35]
SP2	2012	Mobile Learning New forms of education	[36]
SP3	2016	An initial theoretical usability evaluation model for assessing defence mobile e-based application system	[31]
SP4	2018	Modeling Measurement Metrics for E-Book App on Mobile Devices	[24]
SP5	2017	Enlight: A Comprehensive Quality and Therapeutic Potential Evaluation Tool for Mobile and Web-Based eHealth Interventions	[28]
SP6	2018	Usability analysis: Is our software inclusive?	[30]

Tabla 3 – Estudios Primarios

4. Resultados e Interpretación

Esta sección presenta los resultados obtenidos en la PLR, los cuales están divididos en dos partes: Generalidades de los estudios primarios y la respuesta a la pregunta de investigación.

La primera parte se encuentra resumida en la Figura 1, la cual presenta una visión sintética de los seis estudios primarios identificados. En esta figura, los resultados han sido segmentados en dos áreas. La primera (lado izquierdo), consiste básicamente en dos gráficos (parte superior e inferior) de dispersión XY con burbujas en las intersecciones de las categorías tipo-año de publicación (lado izquierdo - parte superior) y tipo de publicación-contexto de usabilidad (lado izquierdo - parte inferior). Los tipos de publicación son conferencias y revistas. El tamaño de cada burbuja está determinado por el número de estudios primarios que se han clasificado como pertenecientes a cada categoría. La segunda área (lado derecho) de la Figura 1, presenta el número de estudios primarios por año de publicación (lado derecho - parte superior) y el número de métricas que han sido clasificadas en cada uno de los ocho contextos de usabilidad identificados

(lado derecho - parte inferior). Como se observa en la parte superior del lado derecho de la Figura 1, el interés en las métricas de usabilidad que pueden ser utilizadas en el desarrollo de apps ha aumentado a partir del 2016.

La segunda parte de esta sección presenta la síntesis del análisis de los estudios primarios, a partir del cual se respondió la pregunta de investigación.

RQ: ¿Qué métricas de accesibilidad y/o usabilidad, consideradas comúnmente en la construcción del software, podrían ser utilizadas para evaluar las apps desde la perspectiva de las PDV?

Como respuesta a la RQ fueron identificadas 73 métricas dentro de los estudios primarios. Las métricas se refieren o están relacionadas con la usabilidad en el desarrollo de aplicaciones móviles. Durante la extracción de las métricas, nos dimos cuenta que cada autor nombra de forma particular las métricas que estudia. Sin embargo, indistintamente de su denominación, las métricas se refieren a un contexto específico de usabilidad para dispositivos móviles. Por ello, para nuestra síntesis, hemos decidido agrupar las métricas identificadas en 8 contextos de usabilidad: Contenido visual, eficacia, navegabilidad, eficiencia, aprendizaje, satisfacción, comprensión y accesibilidad. Hemos codificado estos contextos como "CMA#" y a las métricas las hemos codificado como "CM#".

No obstante, la agrupación permitió albergar 40 métricas en los diferentes contextos, como se muestra en la parte inferior del lado derecho de la Figura 1 (ver rectángulo con fondo gris y esquinas redondeadas). Por ejemplo, los contextos de Eficacia (CMA2) y Navegabilidad (CMA3) están compuestos por ocho métricas cada uno. Un total de 33 métricas no pudieron ser agrupadas en los contextos definidos, dado que su descripción no era clara. Las métricas agrupadas se encuentran en la **Tabla 4**. Los detalles de las métricas no agrupadas se pueden consultar en la sección "Metrics" de: https://goo.gl/nmYxZw.

En resumen, los resultados son interesantes. Por un lado, se verifica que todos los estudios primarios consideran por lo menos una de las métricas relacionadas con la eficacia (CMA2): Task Concordance (CM2), Usage charac-terization (CM5), Task used completely (CM25), No errors detected (CM26), Goal achieve correctly (CMA27), Effectiveness (CM62), Ease of Use (CM68), Effectiveness (CM71). Consideramos que estas métricas son importantes al tra-tarse de aplicaciones para dispositivos móviles. Por otro lado, las métricas más nombradas son: Visual coherence (CM1), graph properties (CM3), icons located correctly (CM13), image appeared correctly (CM14), text highlighted correctly (CM15), readability (CM61). Estas métricas podrían ser consideradas relevantes para apps que serán utilizadas por PDV, dada su relación muy cercana con temas de visualización.

5. Amenazas a la validez

El método de investigación utilizado en este trabajo corresponde a una versión reducida de una SLR, en la cual fueron consideradas algunas fases y etapas propuestas por Kitchenham & otros (Kitchenham, Brereton, O , & Linkman, 2009). A pesar de ser una revisión inicial de literatura, no fueron descuidados los criterios de rigurosidad para obtener estudios relevantes relacionados con la RQ. La **Tabla 5** muestra una evaluación

de la PLR, la cual fue realizada de acuerdo a los criterios sugeridos por Thompson otros (Thompson , Tiwari , Fu, Moe, & Buckley, 2012) , dado que son simples e independientes del dominio. Los resultados de la evaluación muestran que las decisiones tomadas en la PLR (basadas en su mayoría en validaciones cruzadas) fueron las adecuadas, lo que da muestras de la fiabilidad del método de investigación, y por ende de su validez.

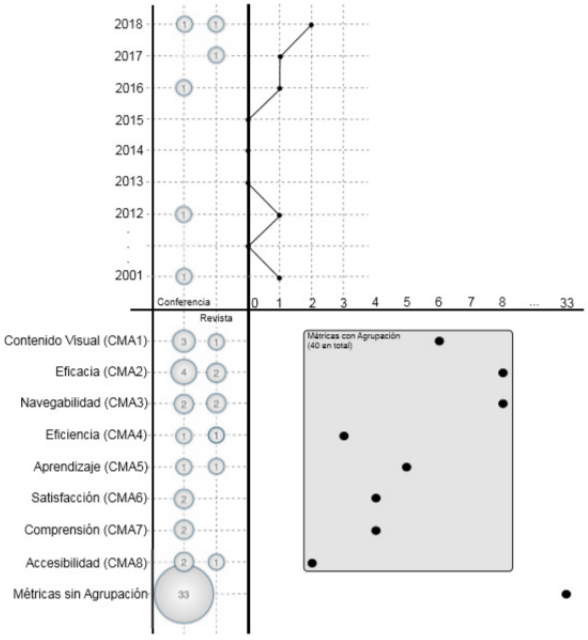


Figura 1 – Mapeo con la distribución de estudios primarios según los contextos de las métricas, incluyendo tipo, año de publicación y número de métricas por agrupación

Cod.	Agrupación	PS1	PS2	PS3	PS4	PS5	PS6
CMA1	Contenido visual	**CM1:** Visual coherence	**CM3:** Graph properties	**CM13:** Icons located correctly,		**CM61:** readability	
				CM14: Image appeared correctly,			
				CM15: Text highlighted correctly			
CMA2	Eficacia	**CM2:** Task Concordance	**CM5:** Usage characterization	**CM25:** Task used completely,	**CM68:** Ease of Use	**CM62:** Effectiveness	**CM71:** Effectiveness
				CM26: No errors detected,			
				CM27: Goal achieve correctly			
CMA3	Navegabilidad		**CM6:** Similarity	**CM43:** Input entering easily,	**CM66:** Navigation	**CM65:** Navigation	
				CM44: Display self adjustment,			

		CM45: Task based on location accurately,			
		CM46: Less input error,			
		CM47: Surrounding interaction			
CMA4	Eficiencia	**CM28:** Task complete quickly,		**CM64:** Efficiency	
		CM29: Action done accurately			
CMA5	Aprendizaje	**CM37:** Skilladaptively,	**CM67:** Learnability		**CM69:** Learnability
		CM38: Learning to operate easily,			
		CM39: Task gain proficiency,			
CMA6	Satisfacción	**CM48:** Reach goal accurately,			**CM73:** Satisfaction
		CM49: Communication easily,			
		CM50: Icon button connect properly			
CMA7	Comprensión	**CM58:** Display load,			**CM70:** Understand-ability
		CM59: Operation clearly,			
		CM60: Operation completely,			
CMA8	Accesibilidad	**CM30:** Information Access easily		**CM63:** Accessibility	

Tabla 4 – Resumen de Métricas encontradas en los Estudios Primarios (PS)

Criterios de evaluación	**Resultado Evaluación RQ**
¿La búsqueda bibliográfica fue exhaustiva?	Si: Se realizaron varias iteraciones de selección aplicando filtros basados en los criterios de inclusión y exclusión.
¿Se utilizaron criterios apropiados en la selección de artículos para su inclusión?	Parcialmente: Los artículos seleccionados fueron apropiados, pero no especializados. Pudieron omitirse artículos que proponían métricas de usabilidad orientadas a las PDV.
¿Se incluyeron estudios que fueron suficientemente válidos para el tipo de pregunta formulada?	Si: Los artículos, tanto del grupo de control como los estudios primarios, abordaron específicamente el tema de métricas de usabilidad para el desarrollo de apps.
¿Los resultados fueron similares entre los estudios?	Si: fueron muy consistentes. Varias métricas se identificaron simultáneamente en dos o tres estudios primarios

Tabla 5 – Criterios para evaluar la PLR

6. Discusión, Conclusiones y Trabajos Futuros

Este trabajo describe una PLR realizada para responder a la pregunta de investigación: ¿Qué métricas de accesibilidad y/o usabilidad, consideradas comúnmente en la construcción del software, podrían ser utilizadas para evaluar las apps des-de la perspectiva de las PDV? Encontramos un total de 6 estudios primarios publicados en conferencias y revistas. Ninguno de estos estudios propone o plan-tea explícitamente métricas de usabilidad para el desarrollo de apps destinadas a las PDV.

Identificamos un total de 73 métricas de usabilidad reportadas en la literatura, de las cuales pudimos agrupar 40 en 8 contextos de usabilidad diferentes: contenido visual, eficacia, navegabilidad, eficiencia, aprendizaje, satisfacción, comprensión y accesibilidad. Los 3 primeros contextos agrupan la mayor cantidad de métricas de usabilidad.

Para las PDV es importante que las aplicaciones consideren principalmente 5 aspectos (Carvalho & Freire, 2017). En primer lugar, las apps deben proporcionar una retroalimentación adecuada para informar que una acción ha tenido un efecto. La retroalimentación está relacionada con los contextos de Comprensión y Contenido Visual. En segundo lugar, a partir de los elementos utilizados en la interfaz debe ser posible inferir correctamente la existencia de una funcionalidad. Es decir, los elementos de la interfaz de usuario utilizados (tales como, campos de entrada de texto, botones de opción) deben tener una asociación adecuada para que las PDV puedan inferir, por ejemplo, que están en una pantalla para el ingreso de datos (si este fuera el caso). Por tanto, este aspecto tiene relación con el contexto de Comprensión. En tercer lugar, la secuencia de interacción debe ser clara para que los usuarios comprendan los pasos que deben seguir para llevar a cabo una funcionalidad. Por ende, este aspecto tiene relación con los contextos de Comprensión y Navegabilidad. En cuarto lugar, debe ser posible realizar la identificación de los botones, de tal manera que cuando un botón específico sea tocado, el lector de pantalla no solamente lea “botón”, sino que también informe de la acción que realiza. Es decir, debe existir una asociación entre los botones y los elementos usados para etiquetarlos expresada en el código, de manera que las tecnologías de asistencia brinden una orientación adecuada. La identificación de los botones está relacionada con los contextos de Accesibilidad y Navegabilidad. Finalmente, las imágenes deben tener un texto alternativo cuando se utilice un lector de pantalla. De esta manera, es posible identificar si la imagen tiene asociada una acción y el usuario pueda, por ejemplo, iniciar una tarea. Este aspecto tiene relación con el contexto Eficacia. De lo anterior, podemos concluir que los contextos de usabilidad de mayor relevancia son: Contenido Visual, Eficacia, Navegabilidad, Comprensión y Accesibilidad. Los 3 primeros contextos coinciden con los que agrupan la mayor cantidad de métricas de usabilidad.

Nuestro estudio permitió establecer métricas de usabilidad que podrían ser adecuadas para el desarrollo de aplicaciones móviles dirigidas a las PDV. Sin embargo, es necesario realizar más esfuerzos de investigación en este sentido. Por tal razón, uno de los trabajos futuros consiste en evaluar algunas métricas (por ejemplo, navigation (CM65), information access easily (CM30)) en el desarrollo de una app para PDV. Otro de nuestros trabajos futuros es realizar una SLR completa y ampliar la búsqueda a otras bases digitales.

Agradecimientos: Esta investigación ha sido parcialmente financiada por el proyecto TIN 2014-60490-P (Ministerio de Economía y Competitividad, Gobierno de España), SENESCYT, ESPE y Armada del Ecuador.

Referencias

Ahmad, N., Rextin, A., & Kulsoom, U. (2018). Perspectives on usability guidelines for smartphone applications: An empirical investigation and systematic literature review. *Information and Software Technology*, 130–149.

Anju, S., & Sagar, K. (2017). A systematic review of software usability studies. *International Journal of Information Technology*, 1–24.

Ba Matraf1, M. S., & Hussain, A. (2018). Modeling Measurement Metrics for E-Book App on Mobile Devices. *Journal of Telecommunication, Electronic and Computer Engineering*, 63–67.

Baharuddin , R., Singh , D., & Razali , R. (2013). Usability dimensions for mobile applications. *Research Journal of Applied Sciences, Engineering and Technology*, 2225–2231.

Baron, C. (2018). *The Statistics Portal*. Obtenido de The Statistics Portal: https://www.statista.com/statistics/274774/forecast-of-mobile-phone-users-worldwide/

Baumel, A., Faber, K., Mathur, N., Kane, J. M., & Muench, F. (2017). Enlight: A Comprehensive Quality and Therapeutic Potential Evaluation Tool for Mobile and Web-Based eHealth Interventions. *Journal of medical internet research*.

Bolderston, A. (2008). Writing an Effective Literature Review. *Journal Medical Imaging and Radiation Scien*, 86–92.

Campanella, P. (2012). Mobile Learning: New forms of education. *ICETA 2012 • 10th IEEE International Conference on Emerging eLeaming Technologies and Applications*, 51–56.

Carvajal, L., Moreno, A., & M. Segura, S. (2013). Usability through Software Design. *IEEE Transactions on Software Engineering*, 1582–1596.

Carvalho, L. P., & Freire, A. P. (2017). Native or Web-Hybrid Apps?: An Analysis of the Adequacy for Accessibility of Android Interface Components Used with Screen Readers. *Proceedings of the XVI Brazilian Symposium on Human Factors in Computing Systems*, 38:1–38:10.

Cruz Reis, R., Gomes, L., Dias-Neto, A., & de L. Fontão, A. (2015). Usability Evaluation Approaches for (Ubiquitous) Mobile Applications: A Systematic Mapping Study. *The Ninth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies (UBICOMM)*.

Dias Silva, F., Carvalho Nepomuceno, M. C., Silva, A. G., & Freire Pimenta, A. (2018). Accessibility and Usability Problems Encountered on Websites and Applications in Mobile Devices by Blind and Normal-vision Users. *Proceedings of the 33rd Annual ACM Symposium on Applied Computing (SAC)*, págs. 2022–2029.

Facanha Rocha, A., Cavalcante, M., & Campos, B. (2014). Touchscreen Mobile Phones Virtual Keyboarding for People with Visual Disabilities. *Human-Computer Interaction. Applications and Services: 16th International Conference, HCI International 2014, Heraklion*, 134–145.

FortesPontin , R. M., Salgado, A. d., de Souza Santos, F., Agostini do Amaral, L., & Nogueira da Silva , E. A. (2017). Game Accessibility Evaluation Methods: A Literature Survey. *Springer International Publishing*, 182–192.

Guerrero, H., & Vega, V. (2018). Usability analysis: Is our software inclusive. *Springer International Publishing AG*.

Hanlu , Y., Malu, M., & Findla, U. (2014). Current and Future Mobile and Wearable Device Use by People with Visual Impairments. *Proceedings of the Annual ACM Conference on Human Factors in Computing Systems (CHI)*, 3123–3132.

Joorabchi, M. E., Mesbah, A., & Kruchten, P. (2013). Real Challenges in Mobile App Development. *2013 ACM / IEEE International Symposium on Empirical Software Engineering and Measurement*, 15–24.

Kane, S. K. (2009). Freedom to Roam: A Study of Mobile Device Adoption and Accessibility for People with Visual and Motor Disabilities. *Proceedings of the 11th International ACM SIGACCESS Conference on Computers and Accessibility*, 115–122.

Kitchenham, B., Brereton, O , P., & Linkman, S. (2009). Systematic literature reviews in software engineering--a systematic literature review. *Information and software technology*, 7–15.

Kronbauer H, A., & Santos S., C. A. (2012). Smartphone Applications Usability Evaluation: A Hybrid Model and Its Implementation. *Human-Centered Software Engineering*, 146–163.

Kulpa Costa, C., & Amaral Gon, F. (s.f.). Avaliação da interação entre usuários de baixa visão e as interfaces gráficas digitais de um tablet: foco na usabilidade. *Blucher Design Proceedings. Blucher*, 2015.

Kumanan , W., Atkinson , K., Westeinde , J., Bell, C., Marty , K., Dean , F., . . . Bett, J. (2016). An evaluation of the feasibility and usability of a proof of concept mobile app for adverse event reporting post influenza vaccination. *Human Vaccines \& Immunotherapeutics*, 1738–1748.

Kumar, B. A., & Mohite, P. (2018). Usability of mobile learning applications: a systematic literature review. *Journal of Computers in Education*, 1–17.

Lettner, F., & Holzmann, C. (2012). Automated and Unsupervised User Interaction Logging As Basis for Usability Evaluation of Mobile Applications. *Proceedings of the 10th International Conference on Advances in Mobile Computing \& Multimedia*, 118–127.

Molano I., J. R., Yara S. , E., & García K. , L. J. (2015). Model for measuring usability of survey mobile apps, by analysis of usability evaluation methods and attributes. *2015 10th Iberian Conference on Information Systems and Technologies (CISTI)*, 1–6.

Olmedo-Moreno , E. M., & López-Delgado , A. (2015). Apps. accessibility and usability by people with visual disabilities. *12th International Conference on Cognition and Exploratory Learning in Digital Age (CELDA)*, 267–270.

P. Di , G., M., R., M. , S., G. , T., G. , V., T. , G., . . . G. , W. (2012). User centered scenario based approach for developing mobile interfaces for Social Life Networks. *First International Workshop UsARE*, 18–24.

Sánchez Sierra, J., & Togores, J. (2012). Designing mobile apps for visually impaired and blind users. *The Fifth International Conference on Advances in Computer-Human Interactions*, 47–52.

Sauro, J., & Lewis R., J. (2009). Correlations Among Prototypical Usability Metrics: Evidence for the Construct of Usability. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1609–1618.

Seffah , A., Kececi, N., & Donyaee, M. (2001). QUIM: A Framework for Quantifying Usability Metrics in Software Quality Models. *Proceedings Second Asia-Pacific Conference on Quality Software*, 311–318.

Thompson , M., Tiwari , A., Fu, R., Moe, E., & Buckley, D. I. (2012). A Framework to Facilitate the Use of Systematic Reviews and Meta-analyses in the Design of Primary Research Studies. *Agency for Healthcare Research and Quality - Department of Health and Human Services*.

TIC, O. A. (27 de Agosto de 2013). discapnet. *Fundación Once*. Obtenido de disc@pnet: http://www.discapnet.es/sites/default/files/areas-tematicas/tecnologia/informe_sintetico_observatorio_aplicaciones_moviles_27-08-2013.pdf

Velásquez, J. (2015). Una Guía Corta para Escribir Revisiones Sistemáticas de Literatura. *Research Gate*.

World Health Organization. (2017). Obtenido de Blindness and visual impairment: http://www.who.int/en/news-room/fact-sheets/detail/blindness-and-visual-impair ment

Zali , Z., & Ahmad Fadzlah , A. F. (2016). An initial theoretical usability evaluation model for assessing defence mobile e-based application system . *International Conference on Information and Communication Technology (ICICTM)*.

Zhanga, H., Ali Babarb, M., & Tellb, P. (2010). Identifying Relevant Studies in Software Engineering. *National ICT Australia, University of New South Wales, Australia*.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 11/10/2018
Aceitação/Acceptance: 23/12/2018

MOMMIV: Modelo para descomposición de una arquitectura monolítica hacia una arquitectura de microservicios bajo el Principio de Ocultación de Información

Víctor Velepucha[1], Pamela Flores[1], Jenny Torres[1]

victor.velepucha@epn.edu.ec, pamela.flores@epn.edu.ec, jenny.torres@epn.edu.ec

[1] Facultad de Ingeniería de Sistemas, Escuela Politécnica Nacional, Quito, Ecuador

Pages: 1000–1009

Resumen: Los cambios tecnológicos actuales y la necesidad de las empresas de migrar sus aplicaciones legadas hacia plataformas modernas han permitido plantear nuevas formas de migrar aplicaciones hacia otras arquitecturas como son los microservicios. Sin embargo, hay desafíos que deben considerarse en el proceso de migración, como son la complejidad de extraer la lógica de negocio embebida en las aplicaciones, las dificultades por cambios en las bases de datos relacionales de la aplicación que afectan la transaccionalidad de la aplicación, la creación de microservicios duplicados o que realicen más de una funcionalidad de negocio. Hasta la actualidad, no se ha encontrado en la literatura un modelo que permita tener un fundamento teórico que solvente dichos problemas. Este trabajo de investigación recoge aspectos importantes de la descomposición y la migración a través de una revisión de literatura, de donde nace y se expone un modelo que permita descomponer arquitecturas monolíticas a microservicios basado en el Principio de Ocultación de Información.

Palabras-clave: migración; modernización; monolítica; microservicio; Principio de Ocultación de Información.

MOMMIV: Model for the decomposition of a monolithic architecture towards a microservices architecture under the Principle of Information Hiding

Abstract: The current technological changes and the need for companies to migrate their legacy applications to modern platforms have allowed to find new ways to migrate applications to other architectures such as microservices. However, there are challenges that must be considered in the migration process, such as the complexity of extracting the business logic embedded in the applications, the difficulties due to changes in the relational databases of the application that affect the transactionality of the application, the creation of duplicate microservices or that perform more than one business functionality. Up to the present time, a model has not been found in the literature that allows having a theoretical foundation that solves these problems. This research work collects important aspects of decomposition and migration through a literature review, from which a model that allows the decomposition of monolithic architectures towards microservices is born and exposed based on the Principle of Information Hiding.

***Keywords**:* migration; modernization; monolithic; microservice; Information Hiding Principle

1. Introducción

Las empresas para su continuo funcionamiento dependen de aplicaciones, las cuales se acumulan con el paso del tiempo. A estas aplicaciones antiguas se las conoce como legadas. Muchas de estas aplicaciones legadas se construyen internamente por las empresas o se compran a proveedores externos y se encuentran desplegadas generalmente en Centros de Datos propios de la organización conocido como Datacenter *On Premise*. Entre los problemas que actualmente se evidencian en las empresas, es que la mayoría de estas aplicaciones legadas son monolíticas y tienen obsolescencia, lo que aumenta la complejidad para su mantenimiento y evolución. El software legado tiene dificultades para ejecutarse sobre nueva infraestructura y sistemas operativos modernos de forma nativa. Existe escasez de profesionales calificados para realizar cambios en un aplicativo legado por lo que la calidad del código fuente nuevo desarrollado se reduce. El abstraer la lógica de negocio de las aplicaciones para crear los nuevos microservicios puede ser un proceso complejo por lo que puede haber muchos intentos fallidos (Furda, Fidge, Zimmermann, Kelly, & Barros, 2018; Gholami, Daneshgar, Beydoun, & Rabhi, 2017; Reza Bazi, Hassanzadeh, & Moeini, 2017).

Todos estos problemas conllevan a que las empresas busquen alternativas para eliminar la obsolescencia de sus aplicaciones y a la vez modernizarlas utilizando tecnología de punta, siendo una solución viable la migración de aplicaciones monolíticas hacia una arquitectura de microservicios.

En nuestra investigación hacemos referencia a la siguiente pregunta de investigación: ¿Cómo descomponer arquitecturas monolíticas para obtener una arquitectura de microservicios basándose en el Principio de Ocultación de Información?

Conducidos con esta pregunta y apegados a una revisión de literatura se plantea un modelo que permita llevar a cabo este proceso de migración.

En la Sección 2, se presenta el fundamento teórico de lo que es un microservicio y el Principio de Ocultación de Información, en la Sección 3 se resume el estado del arte, en la Sección 4 se explica la propuesta de investigación y finalmente en la Sección 5 se exponen las conclusiones y trabajo futuro.

2. Antecedentes

2.1. Microservicio

Un microservicio es un conjunto de pequeños servicios que pueden ser desarrollados independientemente, desplegados ya sea en Centros de datos On-Premise o en la Computación en la Nube (Clarke, Elger, & O'Connor, 2016). Al ser desplegados en la computación en la nube existe la ventaja que pueden escalar de forma independiente, más rápida que en On-Premise. Un microservicio consiste en partir un sistema en pequeños componentes cooperantes, donde estos componentes interactúan unos con otros mediante interfaces (Thönes, 2015). Cada microservicio al ser desarrollado y desplegado

de forma independiente, puede ser creado en cualquier lenguaje de programación y cualquier plataforma (Gutiérrez–Fernández, Resinas, & Ruiz–Cortés, 2017). Desde el año 2014 la arquitectura de microservicios ha emergido como una aproximación ágil para desarrollo de aplicaciones web escalables a entornos en la nube (Villamizar et al., 2017). El desarrollo de pequeños componentes en microservicios hace que el desarrollo sea más ágil e independiente de la tecnología y lenguaje de programación (Lewis J Fowler M, 2015).

Se puede resumir que la arquitectura de un microservicio propone descomponer aplicaciones en pequeños componentes, donde cada componente realiza una funcionalidad puntual de negocio. Esto difiere al desarrollo de aplicaciones basados en arquitecturas monolíticas o arquitectura basadas en n-capas.

La descomposición de aplicaciones en microservicios incrementa la complejidad de integración, por lo que se opta por usar comunicación asíncrona basada en eventos para desacoplar interfaces de integración entre los servicios.

Un desafío al que se debe enfrentar los que desarrollan microservicios es identificar correctamente cuales son los límites que debe tener cada microservicio, más que todo cuando estos microservicios exponen sus interfaces a otros microservicios, ya que cualquier cambio en la interfaz puede afectar a otros microservicios.

Otro desafío es ver como se autoorganizan los equipos de trabajo, pues un equipo de trabajo con poco conocimiento técnico puede crear microservicios en desorden y en lugar de simplificar se hace más complejo.

2.2. Principio de Ocultación de Información

David Parnas (Parnas, 1972) en el año 1972, introdujo el Principio de Ocultación de Información que en inglés se denomina “Information Hiding Principle”, y en su artículo *On the Criteria to Be Used in Decomposing Systems into Modules*, recomienda descomponer sistemas en módulos cuasi- descomponibles, que dicho de otra manera es dividir sistemas en módulos, que tengan la menor dependencia unos con otros logrando sistemas cuasi- descomponibles y obtener estos beneficios: a) Tiempo de desarrollo cortos, b) Flexibilidad en el desarrollo de un sistema, c) Comprensibilidad de un sistema.

Los **tiempos de desarrollos cortos,** se logra cuando es posible separar una aplicación en módulos que permitan realizar el trabajo por grupos con la menor comunicación posible. La **flexibilidad** consiste en que se pueden realizar cambios drásticos o de gran impacto en módulos sin que esto afecte a otros módulos. La **comprensibilidad** se refiere a que se puede entender un módulo a la vez, sin tener que entender todo el sistema. Esto puede beneficiar a que un equipo desarrollo pueda estar especializado para dar mantenimiento y añadir nueva funcionalidad en un módulo específico.

Adicionalmente Parnas (Parnas, 1972) propone los siguientes lineamientos para descomponer módulos. a) Cada módulo debe caracterizado por su conocimiento de una decisión de diseño, b) Ocultar las decisiones de diseño de otros módulos.

El que cada módulo esté caracterizado por su conocimiento de una decisión de diseño empata con la definición de lo que es un microservicio y sirve para el propósito de descomponer una aplicación monolítica hacia microservicios.

El ocultar las decisiones de diseño de otros módulos, permite que se puedan realizar cambios internos en cada módulo sin que el que consume el módulo deba preocuparse de cómo está realizada la implementación interna.

3. Trabajos relacionados

Basados en la revisión del estado del arte, se identifican diversas metodologías, modelos, procesos, frameworks (marcos de trabajo) que abordan desde varias perspectivas el cómo migrar aplicaciones monolíticas. En las siguientes subsecciones mostramos cómo llevan a cabo este proceso.

3.1. Metodologías para migración

Mohagheghi et al. (Mohagheghi & Sæther, 2011) proponen una metodología y herramientas para la migración de aplicaciones basados en el modelo "Model-Driven" para llevar aplicaciones legadas que se encuentren en una arquitectura SOA hacia la computación en la nube, donde plantean que la arquitectura de la aplicación sea descompuesta, desacoplable y escalable, además hacen énfasis en que se debe tomar en cuenta los requerimientos relacionados con servidores, almacenamiento, seguridad de datos, red y tiempo de respuesta, precios y se apoyan con un estudio de caso con el proyecto REMICS que tiene como objetivo evaluar la metodología propuesta. Este modelo está enfocado en una arquitectura SOA, y no se adapta hacia una arquitectura de microservicios. Strauch et al. (Strauch et al., 2014), presenta una metodología para migrar la capa de base de datos de la aplicación hacia la computación en la nube. Esta metodología consiste en realizar una serie de pasos tales como la creación del escenario a ser migrado, indicar donde se realizará el alojamiento con un proveedor en la nube, la identificación de patrones de migración y la ejecución del proceso de migración. Una debilidad de esta propuesta es que se puede notar la falta de soporte en la fase de pruebas, y que la metodología considera solo la migración de la capa de base de datos de una aplicación y no se adapta hacia una arquitectura de microservicios.

3.2. Modelos para migración

Jamshidi et al. (Jamshidi, Ahmad, & Pahl, 2013) presenta un modelo llamado Cloud-RMM que contiene veinte tareas de migración de aplicaciones hacia la computación en la nube y clasifica estas tareas en tres procesos principales: proceso de planificación de la migración, proceso de migración y proceso de evaluación de la migración. Frey et al. (Frey & Hasselbring, 2010) propone un modelo de migración de aplicaciones hacia la computación en la nube llamado CloudMIG que consiste en seis actividades para migrar una aplicación hacia un entorno de nube, teniendo como punto positivo el revelar la arquitectura de la aplicación en sus primeras actividades e indicar cómo se realiza la transformación cuando se despliega en la computación en la nube. Sin embargo, una debilidad de este modelo es que todavía es una propuesta temprana y se enfoca únicamente en la perspectiva de las aplicaciones en un modelo de Software como Servicio (SaaS) para llevarlas hacia Infraestructura como Servicio (IaaS) o Plataforma como Servicio (PaaS). Mehfuz et. al. (Rashmi & Sahoo, 2012) proponen un modelo en cascada iterativo donde indican cómo realizar el proceso de migración tomando las aplicaciones desde un On Premise y llevarlo hacia la computación en la nube de tal

forma que este modelo pueda ser utilizado por cualquier tipo de organización. Un punto importante es que propone realizar pruebas funcionales en las primeras iteraciones de la migración para determinar el retorno de la inversión (ROI) y el costo total de propiedad (TCO) de modo que desde las primeras iteraciones se puedan mostrar los beneficios en costos de esta migración.

3.3. Procesos para migración

Bucchiarone et al. (Bucchiarone, Dragoni, Dustdar, Larsen, & Mazzara, 2018) indican que el proceso de migración se lo debe realizar desde dentro hacia afuera, implementando una única funcionalidad a la vez creando equipos de desarrollo dedicados a trabajar exclusivamente con microservicios, revelando la arquitectura actual de la aplicación y diseñar correctamente la nueva arquitectura futura que incluya un mecanismo de comunicación mediante un bus de mensajería y automatizando la creación de los nuevos microservicios mediante el uso de DevOps (Roche, 2013) y los despliegues se realicen en contenedores. Furda et al. (Furda et al., 2018) indica el proceso de comprender el estado actual de una aplicación legada y cómo realizar proceso de transformación hacia microservicios siguiendo el Principio de Responsabilidad Simple (SRP) donde cada microservicio realiza únicamente una única funcionalidad de negocio y proponen un modelo centrado en procesos para analizar la estructura y las dependencias entre el la capa de negocio y la capa de acceso a los datos de las aplicaciones. Este estudio lo revisaremos para comparar contra el Principio de Ocultación de Información de nuestro estudio. Aggarwal et al. (Aggarwal, Mathur, & Saraswat, 2013) propone un enfoque que consta de cinco fases y sugiere la migración de datos y aplicaciones en pequeños incrementos, sin embargo, los autores no especificaron como solventar inconvenientes en caso de tener problemas durante la migración. Houssaini et al. (Houssaini, Hafiddi, Nassar, & Kriouile, 2015) proponen un proceso de migración en la nube que se centra en cuáles son las aplicaciones prioritarias para una empresa y su proceso se centra en la selección del mejor proveedor de la nube hacia donde se migrara la aplicación, sin embargo, en este estudio se centraron solo se centra en la fase de implementación de la estrategia de migración y no detallan lo suficientemente de cómo es el proceso de migración de las otras fases. Sun et al. (Sun & Li, 2013) proponen un proceso de migración de aplicaciones hacia la computación en la nube compuesta de tres fases: planificación, migración y validación, sin embargo, la debilidad de esta propuesta es que se centra únicamente en la migración a nivel de infraestructura. Beserra et al. (Beserra, Camara, Ximenes, Albuquerque, & Mendonca, 2012) y Malouche et al. (Malouche, Ben Halima, & Ben Ghezala, 2017) proponen un proceso paso a paso para realizar la migración de aplicaciones a la nube, sin embargo, sus propuestas se centra únicamente en cómo resolver las restricciones al realizar la migración de una aplicación legada y los proveedores de la nube involucrados, proponen un proceso de migración de aplicaciones hacia la computación en la nube y se centran únicamente en la fase de evaluación utilizando un ejemplo específico en un estudio de caso.

3.4. Frameworks para migración

Razavian et al. (Razavian & Lago, 2010) proponen un framework para migrar aplicaciones legadas hacia una arquitectura SOA, que pueda servir como referencia para guiarnos de manera ordenada en los paso a paso que se debe seguir, sin embargo, no está adaptado

a una arquitectura de microservicios. Finalmente, Klems et al. (Klems, Nimis, & Tai, 2008) definen la necesidad del valor o costos de la computación en la nube, identificar componentes clave y estructurar estos componentes en un framework, para ayudar a los responsables de la toma de decisiones a estimar los costos de la migración hacia la computación en la nube, comparar estos costos con soluciones de TI convencionales, sin embargo, su propuesta se enfoca únicamente en los costos directos e indirectos de la infraestructura.

A partir de la revisión de la literatura, encontramos que la migración de aplicaciones monolíticas hacia microservicios aún se encuentra en las primeras etapas de madurez, con una carencia de modelos, métodos, procesos y herramientas para realizar las tareas de migración. Muchas de las propuestas se centran en fases específicas del ciclo de vida de desarrollo de software y otras propuestas aún no están maduras o suficientemente probadas con pocos estudios de caso. Este estudio plantea una propuesta de migración que considera una visión holística de todo el ciclo de vida del desarrollo de software para la creación de microservicios a partir de la descomposición de arquitecturas monolíticas tomando como fundamento teórico una forma de descomposición basados en el Principio de Ocultación de Información.

4. Propuesta de investigación

En esta sección se presenta la pregunta de investigación, la metodología de investigación y el modelo propuesto de migración de arquitecturas monolíticas hacia microservicios bajo la visión del Principio de Ocultación de Información.

4.1. Pregunta de investigación

Luego de realizar la revisión de literatura presentada en la Sección 3, donde emergen modelos, procesos y frameworks existentes para la migración de aplicaciones monolíticas hacia microservicios, nace la siguiente pregunta de investigación:

¿Cómo descomponer arquitecturas monolíticas para obtener una arquitectura de microservicios basándose en el Principio de Ocultación de Información?

4.2. Metodología de Investigación

La propuesta metodológica se basa en los principios de la investigación empírica donde se escoge como método de investigación un estudio de caso. Un estudio de caso se lleva a cabo para investigar una sola entidad o fenómeno en su contexto de la vida real y en un espacio de tiempo específico. Típicamente, el fenómeno puede ser difícil de distinguir claramente de su entorno y el investigador tiene que recopilar información detallada sobre el fenómeno a investigar durante un período de tiempo. Durante la realización de un estudio de caso, se debe aplicar una variedad de diferentes procedimientos de recolección de datos y perspectivas de análisis (Wohlin et al., 2012).

4.3. MOMMIV: Modelo de migración

De lo expuesto en el estado del arte, existen varias propuestas para la migración de aplicaciones, sin embargo, ningún modelo realiza una propuesta de migración de

aplicaciones monolíticas hacia microservicios basándose en el Principio de Ocultación de Información. A continuación, presentamos nuestro modelo llamado MOMMIV que significa Modelo de Migración a Microservicios Versátil, el cual se presenta en la Figura 1.

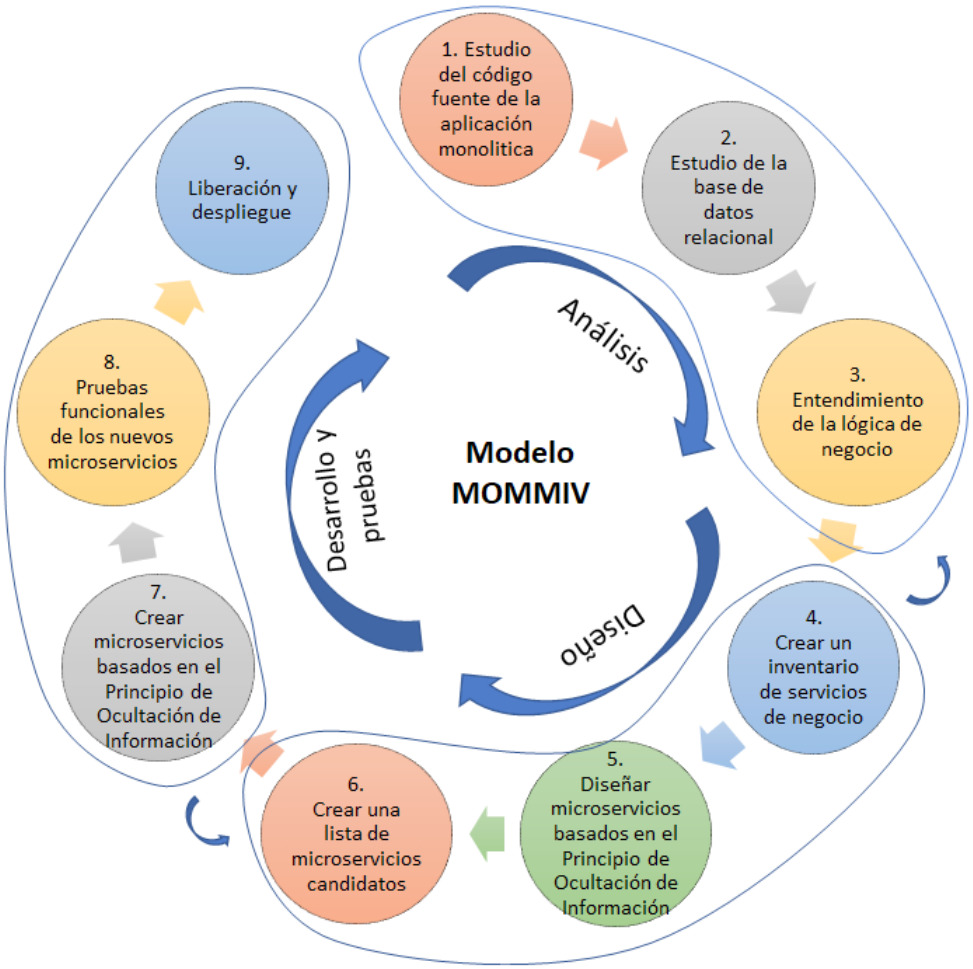


Figura 1 – Modelo MOMMIV para migrar aplicaciones monolíticas hacia microservicios

Este modelo se divide en 3 fases que son: a) Análisis, b) Diseño y c) Desarrollo y pruebas. En la fase de **Análisis** se realiza un entendimiento de una aplicación objetivo, revisando su código fuente y entendiendo su arquitectura, componentes y dependencias y una revisión de su base de datos relacional con el fin de adquirir un entendimiento de la funcionalidad de negocio de la aplicación monolítica. En la fase de **Diseño** se crea un inventario de la funcionalidad de negocio identificada de la fase de análisis. Con este inventario se procede a diseñar los nuevos microservicios basándonos en los lineamientos del Principio de Ocultación de Información. Con los resultados de los diseños de microservicios se levanta un listado de microservicios candidatos

tomando en cuenta que exista trazabilidad con la funcionalidad de negocio. En la fase de **Desarrollo y Pruebas** se crean los microservicios basándose en el Principio de Ocultación de Información, se desarrollan los mecanismos de comunicación como son API's de mensajería, y posibles llamadas entre microservicios, pruebas unitarias. Luego se realizan pruebas de la funcionalidad de negocio de tal forma que se tengan los mismos resultados que en la aplicación monolítica. Finalmente, si el resultado de pruebas es exitoso se realiza una liberación de los microservicios en un ambiente productivo con un despliegue controlado con herramientas DevOps. Cabe indicar que en el modelo propuesto se pueden regresar a pasos anteriores para realizar ajustes, mejoras o cambios de ser necesario.

5. Conclusiones y trabajo futuro

Este trabajo de investigación presenta un modelo que respalda la descomposición de arquitecturas monolíticas hacia microservicios bajo una visión del Principio de Ocultación de Información propuesto por David Parnas. El modelo llamado MOMMIV consta de tres fases: a) Análisis, b) Diseño y c) Desarrollo y pruebas. Cada fase tiene embebida un número de pasos, los cuales permiten dar lineamientos para la migración de arquitecturas monolíticas hacia microservicios. Un paso importante del modelo se basa principalmente en diseñar los microservicios bajo los dos lineamientos del Principio de Ocultación de Información que son: a) cada módulo debe hacer una tarea y b) ocultar las decisiones de diseño de otros módulos.

Como trabajo futuro se plantea realizar un estudio de caso identificando un aplicativo con arquitectura monolítica que se esté usando en la vida real en un entorno productivo, para aplicar la validez de nuestro modelo.

Referencias

Aggarwal, V., Mathur, M., & Saraswat, N. (2013). Comprehensive Cloud Incremental Data-Application Migration-A Proposed Model for Cloud Migration. *International Journal of Computer Applications in Engineering Sciences*, *3*(1), 20.

Beserra, P. V., Camara, A., Ximenes, R., Albuquerque, A. B., & Mendonca, N. C. (2012). Cloudstep: A step-by-step decision process to support legacy application migration to the cloud. In *2012 IEEE 6th International Workshop on the Maintenance and Evolution of Service-Oriented and Cloud-Based Systems (MESOCA)* (pp. 7–16). IEEE. https://doi.org/10.1109/MESOCA.2012.6392602

Bucchiarone, A., Dragoni, N., Dustdar, S., Larsen, S. T., & Mazzara, M. (2018). From Monolithic to Microservices: An Experience Report from the Banking Domain. *IEEE Software*, *35*(3), 50–55.

Clarke, P. M., Elger, P., & O'Connor, R. V. (2016). Technology enabled continuous software development. In *Proceedings of the International Workshop on Continuous Software Evolution and Delivery - CSED '16* (pp. 48–48). New York, New York, USA: ACM Press. https://doi.org/10.1145/2896941.2896943

Frey, S., & Hasselbring, W. (2010). Model-based migration of legacy software systems to scalable and resource-efficient cloud-based applications: The cloudmig approach. *Proceedings of the First International Conference on Cloud Computing, GRIDs and Virtualization*, (c), 155–158. Retrieved from http://www.thinkmind.org/index.php?view=article&articleid=cloud_computing_2010_6_40_50065

Furda, A., Fidge, C., Zimmermann, O., Kelly, W., & Barros, A. (2018). Migrating Enterprise Legacy Source Code to Microservices: On Multitenancy, Statefulness, and Data Consistency. *IEEE Software*, *35*(3), 63–72.

Gholami, M. F., Daneshgar, F., Beydoun, G., & Rabhi, F. (2017). Key challenges during legacy software system migration to cloud computing platforms — an empirical study. *Information Systems*, *67*, 100–113. https://doi.org/10.1016/j.is.2017.03.008

Gutiérrez–Fernández, A. M., Resinas, M., & Ruiz–Cortés, A. (2017). Redefining a process engine as a microservice platform (Vol. 281, pp. 252–263). Springer Verlag. https://doi.org/10.1007/978-3-319-58457-7_19

Houssaini, C. El, Hafiddi, H., Nassar, M., & Kriouile, A. (2015). CM-factory for enabling enterprise migration to cloud. *International Journal of Cloud Computing*, *4*(3), 211. https://doi.org/10.1504/IJCC.2015.071718

Jamshidi, P., Ahmad, A., & Pahl, C. (2013). Cloud Migration Research: A Systematic Review. *IEEE Transactions on Cloud Computing*, *1*(2), 142–157. https://doi.org/10.1109/TCC.2013.10

Klems, M., Nimis, J., & Tai, S. (2008). Do clouds compute? a framework for estimating the value of cloud computing. In *Workshop on E-Business* (pp. 110–123).

Lewis J Fowler M. (2015). Microservice. July. Retrieved from https://martinfowler.com/articles/microservices.html

Malouche, H., Ben Halima, Y., & Ben Ghezala, H. (2017). Enterprise information system migration to the cloud: Assessment phase. *Proceedings of IEEE/ACS International Conference on Computer Systems and Applications, AICCSA*. https://doi.org/10.1109/AICCSA.2016.7945761

Mohagheghi, P., & Sæther, T. (2011). Software engineering challenges for migration to the service cloud paradigm: Ongoing work in the remics project. In *Services (SERVICES), 2011 IEEE World Congress on* (pp. 507–514).

Parnas, D. L. (1972). On the criteria to be used in decomposing systems into modules. *Communications of the ACM*, *15*(12), 1053–1058.

Rashmi, M. S., & Sahoo, G. (2012). A five-phased approach for the cloud migration. *Int J Emerg Technol Adv Eng*, *2*(4), 286–291.

Razavian, M., & Lago, P. (2010). Towards a conceptual framework for legacy to soa migration. In *Service-Oriented Computing. ICSOC/ServiceWave 2009 Workshops* (pp. 445–455).

Reza Bazi, H., Hassanzadeh, A., & Moeini, A. (2017). A comprehensive framework for cloud computing migration using Meta-synthesis approach. *Journal of Systems and Software*, *128*, 87–105.

Roche, J. (2013). Adopting DevOps practices in quality assurance. *Communications of the ACM*, *56*(11), 38–43.

Strauch, S., Andrikopoulos, V., Karastoyanova, D., Leymann, F., Nachev, N., Stäbler, A., … Sheridan, C. (2014). Migrating enterprise applications to the cloud: methodology and evaluation. *International Journal of Big Data Intelligence*, *1*(3), 127. https://doi.org/10.1504/IJBDI.2014.066319

Sun, K., & Li, Y. (2013). Effort estimation in cloud migration process. *Proceedings - 2013 IEEE 7th International Symposium on Service-Oriented System Engineering, SOSE 2013*, 84–91. https://doi.org/10.1109/SOSE.2013.29

Thönes, J. (2015, January). Microservices. *IEEE Software*. https://doi.org/10.1109/MS.2015.11

Villamizar, M., Garcés, O., Ochoa, L., Castro, H., Salamanca, L., Verano, M., … Lang, M. (2017). Cost comparison of running web applications in the cloud using monolithic, microservice, and AWS Lambda architectures. *Service Oriented Computing and Applications*, *11*(2), 233–247. https://doi.org/10.1007/s11761-017-0208-y

Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., & Wesslén, A. (2012). *Experimentation in Software Engineering*. Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-29044-2

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 07/09/2018
Aceitação/Acceptance: 08/11/2018

Aplicación del modelo COBIT para la administración de datos (DS11): el caso de un centro de análisis clínico

Jorge Buele[1,2], John Espinoza[2], Daniel Fiallos[3], Mónica Ruiz R.[1], Ana-María Pilco-Salazar[1], John Reyes[1], Franklin Salazar L.[1]

jl.buele@uta.edu.ec, jjespinoza6@espe.edu.ec, dfiallos.mti@uisek.edu.ec, mp.ruiz@uta.edu.ec, am.pilco@uta.edu.ec, johnpreyes@uta.edu.ec, fw.salazar@uta.edu.ec

[1] Facultad de Ingeniería en Sistemas, Electrónica e Industrial, Universidad Técnica de Ambato, 180103, Ambato, Ecuador

[2] Departamento de Eléctrica y Electrónica, Universidad de las Fuerzas Armadas ESPE, 050104, Latacunga, Ecuador

[3] Facultad de Arquitectura e Ingenierías, Universidad Internacional SEK, 170302, Quito, Ecuador

Pages: 1010–1021

Resumen: Un eje prioritario en la administración organizacional de los centros de análisis clínico es la correcta utilización de las tecnologías de la información (TI) y la gestión segura de los datos y registros. En este trabajo se implementa el modelo COBIT 4.1 y como estudio de caso se ilustra el proceso de administración de datos (DS11). Se implementa en una institución que ha operado sin ninguna herramienta tecnológica, en cuanto al manejo de su información interna. Los niveles de seguridad en la administración de los datos se los determina aplicando una encuesta a una muestra del personal y en base a ello, se calcula el nivel de madurez y se presentan los indicadores estadísticos respectivos. De esta forma se plantean y ejecutan acciones correctivas que permiten elevar estas estadísticas, evitando vulneraciones y fugas de información. Esta propuesta es validada y se brinda un precedente para futuras investigaciones en América Latina.

Palabras-clave: Gestión de datos; servicios y asistencia (DS); gobierno de TI; gestión de seguridad.

Application of COBIT model for data management (DS11): the case of a clinical analysis center

Abstract: A priority axis in the organizational management of clinical analysis centers is the correct use of information technology (IT) and the secure management of data and records. In this work, the COBIT 4.1 model is implemented and the data management process (DS11) is illustrated as a case study. It is implemented in an institution that has been operating without any technological tool in terms of the management of its internal information. Security levels in data administration are determined by applying a survey to a sample of the personnel and, based on it, the

level of maturity is calculated and the respective statistical indicators are presented. In this way, corrective actions are proposed and executed to allow to raise these statistics, avoiding information breaches and leaks. This proposal is validated and provides a precedent for future research in Latin America.

***Keywords**:* Data management; services and support (DS); IT governance; security management.

1. Introducción

La importancia que las tecnologías de la información (TI) poseen en el desarrollo de la sociedad moderna aumenta exponencialmente (Harris, 2017) (Wang, Zhang & Chen, 2008). Es así que se puede influenciar y modificar los procedimientos para la gestión de las empresas tanto públicas como privadas (Johnson, Lukaszewski & Stone, 2016). La gestión de las TI requiere organización y una utilización óptima, eficaz y eficiente de los recursos de la empresa para alcanzar los objetivos planteados por la misma (Aharonson & Schilling, 2016). El éxito de la organización se ve relacionado con la utilización de las TI para responder a las presiones comerciales y económicas que se puedan suscitar en las actividades desarrolladas (Jha, Pinsonneault & Dubé, 2016). En la búsqueda del mejoramiento de los procesos se han implementado procesos de control como las auditorías, que buscan corregir errores en niveles tanto operativos como gerenciales. De esta manera se podrán tomar acciones correctivas que prioricen el capital humano y se entregue mejores productos y servicios (Espejo Chavarría, Bayona Oré & Pastor, 2016).

Para determinar si las acciones que se realizan periódicamente están contribuyendo en el proceso de mejora continua, se realiza el análisis de los datos que se recolectan en un determinado intervalo de tiempo (Roa Mantilla & Giraldo, 2013). Esta información representa la esencia de la compañía, sin importar su tamaño, refleja cuantitativa y cualitativamente su crecimiento o deterioro (Jasperson, Carter & Zmud, 2005). Es por ello que la correcta administración de dichos datos se convierte en una actividad de prioridad elevada. Se busca constantemente prevenir vulnerabilidades y fugas de información, que podrían ser aprovechadas por sus competidores, que buscan sobresalir en un entorno comercial tan disputado (Ibrahim & Nurpulaela, 2016). Por ello es necesario administrar la librería de medios y su correcta eliminación, el respaldo y la recuperación de los datos. El uso de herramientas tecnológicas que promuevan la administración de los datos ha tomado mayor relevancia en este siglo. Una de ellas es COBIT (Control Objectives Control Objectives for Information and related Technology), un modelo para el gobierno de las TI presentado por la ISACA (Information Systems Audit and Control Association) y el ITGI (IT Governance Institute). Esta herramienta busca optimizar las TI, equilibrando beneficios, uso de recursos y niveles de riesgos. Este modelo abarca 34 objetivos de nivel alto que cubren 215 objetivos de control agrupados en cuatro dominios (Peter & Vijayadhasan, 2018).

En tal contexto, en este trabajo se describe el proceso de aplicación del proceso DS11 del modelo COBIT 4.1 y sus objetivos de control en los procesos de una organización. La administración de los datos es evaluada siguiendo los objetivos de control del proceso antes mencionado, con la intención de identificar falencias en el sistema usado durante años. A través de la presentación de resultados preliminares, se determinaron actividades para la corrección de errores y una optimización del manejo interno de la información.

Por tanto, este artículo se organiza de la siguiente manera: la introducción en la sección 1, en la sección 2 se muestra el estado del arte de investigaciones relacionadas. En la sección 3 se presenta la metodología utilizada. Los resultados se ilustran en la sección 4 y las conclusiones y trabajos futuros se describen en la sección 5.

2. Estado del arte

La fundamentación de esta investigación se basó en trabajos relacionados que muestran la influencia de este modelo en la administración empresarial y como se puede realizar una correcta aplicación del mismo. En (Mantilla & Giraldo, 2013) se presentan conceptos relacionados con la arquitectura empresarial (AE) y su alineación con la tecnología de la información (TI) en los negocios. De esta forma se presenta un modelo de mantenimiento para estructurar los diversos elementos que componen el proceso. Con ello se define la trazabilidad de los diferentes cambios que surgen durante la implementación de COBIT y se controla dicho proceso de mantenimiento. En (Hussain & Siddiqui, 2005) se describe un modelo realizado mediante el desarrollo de matrices para los cuatro dominios de COBIT. Con ello se proporciona un enfoque cuantitativo para una mejor implementación del gobierno corporativo de TI a través de una auditoría y controles efectivos. En (Ozkan, 2015) se busca identificar los riesgos, desafíos y problemas de la implementación de Scrum dentro de un entorno COBIT. Este estudio funciona como un mapa de riesgos para las organizaciones que tienen la oportunidad de adaptar esta herramienta. Se exhibe una lista de desafíos para las instituciones y se las motiva a cumplirlos. La presentación de un enfoque para crear un modelo formal de COBIT utilizando una ontología se describe en (Textor & Geihs, 2015). Se establece una mejora con una formalización de las métricas, la forma de utilizar las reglas de selección y las asignaciones para calcular los valores de las mediciones en tiempo de ejecución. El diseño del modelado se muestra detalladamente y su validación se la realiza mediante pruebas de ejecución.

Una forma de mejorar los sistemas es sometiéndolos a procesos de auditoría como se muestra en (Suryani, Sasmita & Purnawan, 2005), donde se realiza una auditoría al sistema de información contable de una universidad de Indonesia. El objetivo de esta investigación es determinar el nivel de madurez de los sistemas de TI de esta institución como soporte en la gestión de datos financieros. En (Ibrahim & Nurpulaela, 2016) se evalúa el nivel de madurez de la gobernanza de TI en una organización. Se detalla la utilización de los diversos dominios de COBIT 4.1 y los resultados obtenidos luego de aplicar esta herramienta. De manera similar en (Wijayanti, Setiawan & Sukamto, 2017) se utiliza COBIT 4.1 combinado con BSC (Balanced Score Card) en un entorno educativo de nivel superior. Se los utiliza para establecer indicadores de evaluación en forma de cuestionario. Al analizar los 15 objetivos de TI de la Universitas Pendidikan Indonesia (UPI) se han desplegado 17 planes estratégicos, que serán ejecutados con pleno respaldo y acuerdo de los propietarios de los procesos que requieren una mejora. Como ya se habló previamente la seguridad de la información es una prioridad en las empresas tanto privadas, como públicas y es por ello necesaria la implementación de herramientas como se muestra en (Morimoto, 2009). Donde se busca implementar una guía de gestión de seguridad, que permita mantener resguardada la información de relevancia.

3. Metodología

El proceso de administración de datos (DS11) del modelo COBIT en su versión 4.1 consta de cinco actividades que se mencionan a continuación: Traducir el almacenamiento de datos y la disposición de retención en procedimientos. Definir, mantener e implementar procedimientos para administrar la biblioteca de medios. Definir, mantener e implementar procedimientos para la eliminación segura de medios y equipos. Disponer de una copia de seguridad de los datos según el esquema. Definir, mantener e implementar procedimientos para la restauración de datos (Ariyadi & Dirgahayu, 2015). Este proceso está regido por un conjunto de objetivos de control que pueden ser políticas, procedimientos, prácticas o estructuras organizativas, como se muestra en la Tabla 1. Así, se optimiza el uso de la información y se garantiza la disponibilidad de la misma cuando se la requiera. Enfocándose en mantener la integridad, exactitud, disponibilidad y protección de los datos.

Objetivo de control	Descripción
DS11.1	Requerimientos del negocio para la administración de datos.
DS11.2	Acuerdos de almacenamiento y conservación.
DS11.3	Sistema de administración de librerías de medios.
DS11.4	Eliminación.
DS11.5	Respaldo y restauración.
DS11.6	Requerimientos de seguridad para la administración de datos.

Tabla 1 – Objetivos del control del proceso DS11

Un proceso es la unidad más pequeña del gobierno de TI y la aplicación de un modelo de madurez evidencia el desempeño del proceso. El modelo de madurez genérico de COBIT 4.1 se presenta en la Tabla 2, mencionando que se adapta dependiendo de las necesidades de esta institución. En este contexto, el análisis y estudio realizados están orientados a contribuir con el proceso de mejora continua de un centro de toma de muestra clínicas, cuyo diagrama de arquitectura de procesos se presenta en la Figura 1. Para la aplicación de una auditoría de TI, la recopilación de información se realiza mediante observación directa de los procesos y entrevistas al personal de los diferentes departamentos y laboratorios.

Nivel	Estado	Descripción	Índice de madurez
0	No existente	No se aplican procesos administrativos en lo absoluto.	0 – 0,50
1	Inicial/Ad hoc	Los procesos son ad-hoc y desorganizados.	0,51 – 1,50
2	Repetible pero intuitivo	Los procesos siguen un patrón regular.	1,51 – 2,50
3	Proceso definido	Los procesos se documentan y se comunican.	2,51 – 3,50
4	Administrado y medible	Los procesos se monitorean y se miden.	3,51 – 4,50
5	Optimizado	Las buenas prácticas se siguen y se automatizan.	4,51 – 5,00

Tabla 2 –Modelo de madurez genérico

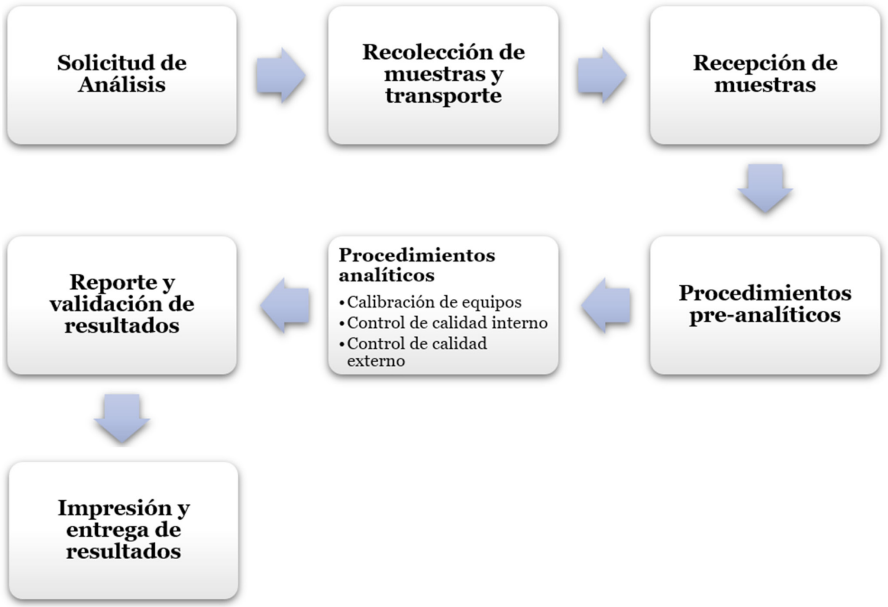


Figura 1 – Arquitectura de procesos de la organización

3.3. Participantes

Para la aplicación de este modelo, se debe establecer un análisis de la situación actual del establecimiento mediante una toma de muestra que sea significativa. Para fines aplicativos se procederá a escoger 10 personas, de las 58 personas que laboran aquí (5 hombres y 5 mujeres). Estas poseen un promedio de edad de 30 años, y poseen conocimientos aceptables sobre las TI. El objetivo de escoger una muestra de población mixta, es poder constatar la información entregada y poder revisarla de una manera enfocada, sin ningún tipo de discriminación. Las áreas seleccionadas para este proceso son: dirección técnica, responsables de los distintos laboratorios, control de calidad, despacho y departamento de TI.

3.2. Instrumentos y técnicas de recolección de datos

Esta investigación utiliza instrumentación en forma de encuesta aplicada a la muestra de la población seleccionada previamente. Ésta se ha formulado con la intención de cumplir cada uno de los objetivos de control del nivel de madurez en el dominio de Entrega y Soporte (DS) (específicamente el proceso DS11). Estos datos obtenidos son del tipo primario, ya que se toman directamente en el lugar de investigación, en base a la observación y preguntas al personal. Con esto, se pretende realizar un proceso general, teniendo en cuenta que, en cada empresa pública o privada varían los enfoques de seguridad de los registros. También se recopilan datos secundarios como la historia de la institución, misión, visión y estructura (aunque no se presentan en el estudio).

3.4. Plan de análisis de datos

Se han utilizado dos técnicas de análisis de datos: descriptiva y cuantitativa. Un análisis descriptivo permite descubrir cómo se gobiernan las TI en la institución actualmente, con la finalidad de implementar las soluciones respectivas. El método cuantitativo brinda un resultado concreto, que permite determinar si los objetivos de control están cumpliendo con la finalidad de mitigar los riesgos y vulnerabilidades en todos los procesos referentes al manejo de datos. Así, mediante el procesamiento y tabulación de las respuestas obtenidas calcular el nivel de madurez. El análisis de los datos obtenidos, debe adecuarse a la realidad, verificando que no fueron alterados y cumplen con los estándares de calidad en cuanto a seguridad.

3.5. Procedimiento

La información solicitada a la institución es relevante para diagnosticar su situación actual y con la cual se puede establecer una metodología. Luego de realizar el cuestionario que proporcione una noción de la eficacia que poseen los diferentes procesos, se implementan gráficas estadísticas que permiten presentar formalmente los datos adquiridos. Con ello, al definir el nivel madurez que posee la organización, se presenta un plan metodológico que contenga lo siguiente:

- Descripción actual de la situación de la empresa.
- Análisis de las vulnerabilidades con respecto al manejo de datos.
- Instauración de medidas que permitan solventar estos problemas
- Análisis de los controles establecidos.
- Verificación de la mejora con los controles.
- Planificación de mejoras a corto y mediano plazo.
- Evaluación luego de un periodo de seis meses.
- Entrega formal de resultados y socialización del proceso implementado.

4. Resultados

Se realizó una encuesta de 12 preguntas enfocadas a la administración óptima de los datos, para verificar su fiabilidad y disponibilidad en cualquier momento. A continuación, en la Tabla 3 se presentan las preguntas realizadas a la muestra de 10 personas.

El análisis de la información derivada luego de aplicar la encuesta de la tabla anterior, permite obtener los siguientes valores porcentuales. El 31,6% de los encuestados asegura que recibe documentos originales con los datos necesarios para procesar la información de forma rápida y precisa, para una entrega oportuna y adecuadamente soportada con su respectiva documentación. El 43,4% indica que tiene un procedimiento definido para el archivo y almacenamiento de datos. Además, afirman que su procedimiento cumple con las exigencias de recuperación, rentabilidad, integridad continua y seguridad en los datos. Un 37,6% posee un procedimiento para mantener un inventario en sitio, el cual le permite garantizar su integridad y uso. El 35,2% posee un manejo adecuado para el acceso a datos sensitivos y al software desde equipos o medios una vez que son eliminados o transferidos para otro uso. Un 48% tiene un plan de continuidad de negocio que respalde los sistemas, aplicaciones y documentación en línea. También posee una verificación para los procedimientos de respaldo y tiempos requeridos para

una restauración completa y exitosa. Finalmente, el 38,2% posee políticas de seguridad aplicables a la recepción, procesamiento, almacenamiento y salida de datos. Según los datos obtenidos la empresa presenta un nivel de riesgo moderadamente alto en el dominio de prestación de entrega y soporte, tal como se puede observar en la Figura 2.

Objetivo de control	N°	Pregunta
1	1.1	¿Se reciben los documentos originales con los datos necesarios para procesar la información de una forma precisa y oportuna?
	1.2	¿Se realizan los reportes de salida de una forma clara y completa y su entrega es oportuna?
	1.3	¿Todas las necesidades de reproceso están adecuadamente respaldadas con su respectiva documentación?
2	2.1	¿Existe un procedimiento definido para el archivo y almacenamiento de datos?
	2.2	¿Este procedimiento considera los requerimientos de recuperación, la rentabilidad, la integridad continua y los requerimientos de seguridad de los datos?
	2.3	¿Posee mecanismos de almacenamiento y conservación de información, que contemplen encriptación y autenticación?
3	3.1	¿Posee un procedimiento para mantener un inventario en el sitio?
4	4.1	¿Posee un manejo adecuado del acceso a datos sensitivos?
	4.1	¿Posee un manejo adecuado del software desde los equipos o medios una vez que son eliminados o transferidos para otro uso?
5	5.1	¿Existe un plan de continuidad de negocio que respalde los sistemas, aplicaciones y documentación en línea?
	5.2	¿Existe una verificación adecuada para el cumplimiento de los procedimientos de respaldo y el tiempo requerido para tener una restauración completa y exitosa?
6	6.1	¿Existen políticas de seguridad aplicables a la recepción, procesamiento, almacenamiento y salida de datos?

Tabla 3 – Encuesta para realizar el diagnóstico previo de la institución

4.1. Análisis de la situación inicial de la institución

Realizado un promedio del cumplimiento de los objetivos de control se tiene un valor del 39%, un nivel definido bastante bajo. En base a este valor se puede deducir que hay falencias existentes con respecto a la administración de los datos. Es decir, no se evidencia un correcto manejo y gestión de los registros almacenados utilizando herramientas de TI. Además, se presentó carencia de documentación original que respalden los procesos de control requeridos. En ciertos casos los documentos eran copias y en otros casos presentaban inconsistencias en su llenado. A partir de estos resultados, se determina el nivel de madurez de la organización, cuyos valores se describen en la Tabla 4. Con la intención de mejorar la gobernanza de las TI en esta institución se aplicaron varias medidas correctivas y entre ellas se menciona las principales:

- Organización y complemento de la documentación necesaria.
- Validación de la integridad de la información que se encuentra en la base de datos, eliminación de inconsistencias y restauración de datos perdidos.

- Mejoramiento del protocolo de seguridad e implementación de respaldos para elevar la redundancia de los registros.
- Mejoramiento del almacenamiento y administración de la librería de medios.
- Establecimiento de responsables a cargo del manejo de los registros internos.

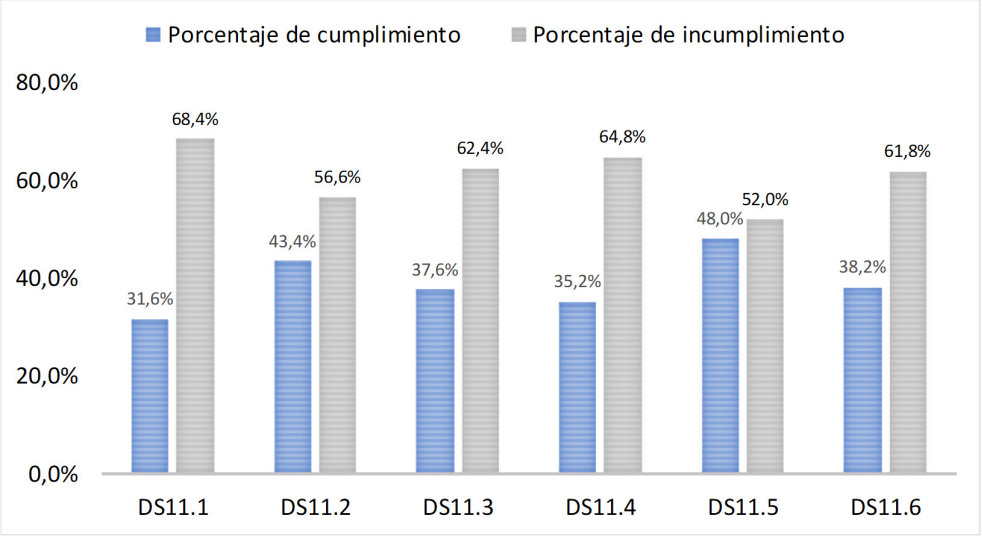


Figura 2 – Resultados de la encuesta inicial

Objetivo de control	Nivel de madurez actual	
	Valor	Nivel
DS11.1	1,58	2 - Repetible pero intuitivo
DS11.2	2,17	2 - Repetible pero intuitivo
DS11.3	1,88	2 - Repetible pero intuitivo
DS11.4	1,76	2 - Repetible pero intuitivo
DS11.5	2,40	2 - Repetible pero intuitivo
DS11.6	1,91	2 - Repetible pero intuitivo
Nivel de madurez promedio	1,95	2 - Repetible pero intuitivo

Tabla 4 – Cálculo del nivel de madurez inicial

3.3. Análisis de la situación de la empresa posterior a la implementación del modelo

Se aplicó los objetivos de control del proceso DS11 por un tiempo de prueba de seis meses a todos los procesos que presentaron ineficiencias. Desde el nivel gerencial hasta los niveles operativos, que incluyen los laboratorios y el despacho de resultados. Con ello, se constató una mejora en el manejo adecuado de la documentación y en la administración digital de los datos. Además, de los requerimientos de seguridad, evitando vulneraciones y pérdidas de información.

Se volvió a aplicar la encuesta establecida en la Tabla 3 y los resultados obtenidos se compararon con los datos previos, como se describe en la Figura 3. El análisis respectivo muestra que hay un incremento del cumplimiento en cada uno de los objetivos de control del proceso DS11. Esta información es la base para un nuevo cálculo de la madurez de la gobernanza de las TI, que se presenta en la Tabla 5.

Como se puede apreciar el nivel de madurez promedio del proceso DS11 incrementó en un 66% luego de la implementación de las medidas correctivas que se mencionaron previamente. A pesar que ahora el nivel se define como un "proceso definido", es evidente que aún se requiere un mayor control en los procesos que involucran TI.

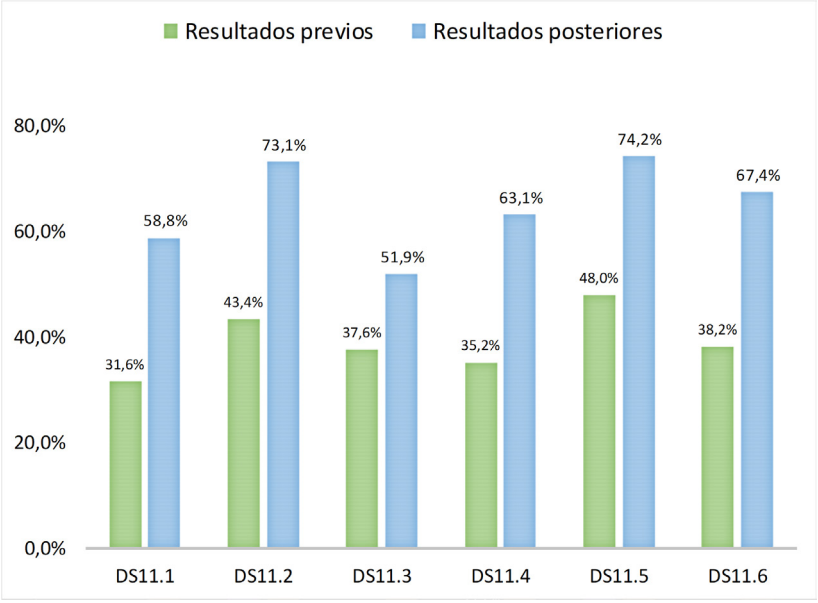


Figura 3 – Comparación del cumplimiento de los objetivos de control antes y después de aplicar la metodología

Objetivo de control	Nivel de madurez actual	
	Valor	Nivel
DS11.1	2,94	3 - Proceso definido
DS11.2	3,66	4 - Administrado y medible
DS11.3	2,60	3 - Proceso definido
DS11.4	3,16	3 - Proceso definido
DS11.5	3,71	4 - Administrado y medible
DS11.6	3,37	3 - Proceso definido
Nivel de madurez promedio	3,24	3 - Proceso definido

Tabla 5 – Cálculo del nivel de madurez posterior

5. Conclusiones y trabajos futuros

La aplicación de una encuesta como parte de una investigación descriptiva a una muestra de 10 personas, permitió evidenciar de una manera clara y concisa la situación actual de un centro de análisis clínico. Previamente la administración de la información, registros y resultados clínicos, poseía la carencia de una metodología de control, lo cual se vio demostrado en inconsistencias de ciertos documentos. Para instituciones de esta índole, la preservación y confidencialidad de la información es una prioridad, que promueve la continuidad de sus funciones. La aplicación de los objetivos de control enfocados al proceso DS11 del modelo COBIT 4.1, logró establecer una mejoría notoria en todos los componentes en un periodo establecido de seis meses. Llegando a aumentar aproximadamente un 66% el nivel de madurez en este proceso en particular. También se ha realizado la socialización de la metodología aplicada, con la intención de incrementar este nivel a futuro y que pueda ser transmitido al nuevo personal que pueda ingresar a laborar. Así se ha logrado robustecer la administración de datos, para garantizar su integridad, confiabilidad, seguridad y disponibilidad.

Cada objetivo de control se ve involucrado con uno o más procesos de TI, por tal razón se debe extender de una manera integral el análisis de los dominios del modelo COBIT. Con esto se logrará entrelazar varios aspectos que podrían detectar a tiempo alguna vulnerabilidad que afecte a la organización. Por tal razón los autores presentan como futura investigación la implementación de los procesos complementarios del modelo COBIT y la utilización de otras herramientas de gestión. Así se obtendrán mejores resultados en donde se establezca de una manera clara la alineación, responsabilidades, procesos y controles de TI.

Referencias

Aharonson, B. S. & Schilling, M. A. (2016). Mapping the technological landscape: Measuring technology distance, technological footprints, and technology evolution. *Research Policy, 45(1)*, 81-96. doi: https://doi.org/10.1016/j.respol.2015.08.001

Ariyadi, D. & Dirgahayu, T. (2015). Assessment to COBIT 4.1 maturity model based on process attributes and control objectives. *2015 International Conference on Science in Information Technology (ICSITech)*, 343–347. IEEE. doi: https://doi.org/10.1109/ICSITech.2015.7407829

Espejo Chavarría, E., Bayona Oré, S. & Pastor, C. (2016). Aseguramiento de la Calidad en el Proceso de Desarrollo de Software utilizando CMMI, TSP y PSP. *RISTI - Revista Iberica de Sistemas e Tecnologias de Informacao, 20*, 62–77. doi: 10.17013/risti.20.62–77

Harris, N. (2017). The Design and Development of Assistive Technology. *IEEE Potentials, 36(1)*, 24–28. doi: https://doi.org/10.1109/MPOT.2016.2615107

Hussain, S. J. & Siddiqui, M. S. (2005). Quantified Model of COBIT for Corporate IT Governance. *2005 International Conference on Information and Communication Technologies*, 158–163. IEEE. doi: https://doi.org/10.1109/ICICT.2005.1598575

Ibrahim & Nurpulaela, L. (2016). Evaluation of IT governance to support IT operation excellent based on COBIT 4.1 at the PT Timah Tbk. *In Information Technology, Computer, and Electrical Engineering (ICITACEE), 2016 3rd International Conference on*, 336–339. IEEE. doi: https://doi.org/10.1109/ICITACEE.2016.7892467

Jasperson, J. Carter, P. E. & Zmud, R. W. (2005). A comprehensive conceptualization of post-adoptive behaviors associated with information technology enabled work systems. *MIS Quarterly: Management Information Systems, 29(3)*, 525–557. doi: https://doi.org/10.25300/MISQ/2016/40.2.08

Jha, S. K., Pinsonneault, A. & Dubé L. (2016). The evolution of an ICT platform-enabled ecosystem for poverty alleviation: the case of eKutir. *MIS Quarterly: Management Information Systems, 40(2)*, 431–445. doi: https://doi.org/10.25300/MISQ/2016/40.2.08

Johnson, R. D., Lukaszewski, K. M. & Stone, D. L. (2016). The Evolution of the Field of Human Resource Information Systems: Co-Evolution of Technology and HR Processes. *Communications of the Association for Information Systems, 38*, 533–553. doi: https://doi.org/10.17705/1CAIS.03828

Mantilla, S. M. R. & Giraldo, O. L. (2013). Model maintenance based in COBIT for the applications architecture in TOGAF enterprise architectures — MOMCAE," *2013 8th Computing Colombian Conference (8CCC)*, 1–6. IEEE. doi: https://doi.org/10.1109/ColombianCC.2013.6637541

Morimoto, S. (2009). Application of COBIT to security management in information systems development. *In 2009 International Conference on Frontier of Computer Science and Technology*, 625–630. IEEE. doi: http://doi.ieeecomputersociety.org/10.1109/FCST.2009.38

Ozkan, N. (2015). Risks, Challenges and Issues in a Possible Scrum and COBIT Marriage. *2015 Asia-Pacific Software Engineering Conference (APSEC)*, 111–118. IEEE. doi: https://doi.org/10.1109/APSEC.2015.29

Peter A., J., & Vijayadhasan G., K. (2018). Design of an energy efficient campus-based on energy audit report. *International Journal of Engineering & Technology, 7(2)*, 370–373. doi: http://dx.doi.org/10.14419/ijet.v7i2.12.11353

Suryani, N. P. S. M., Sasmita, G. M. A., & Purnawan, I. K. A. (2015). Audit of Accounting Information System Using Cobit 4.1 Focus on Deliver and Support Domain. *Journal of Theoretical and Applied Information Technology, 78(3)*, 456–463.

Textor, A. & Geihs, K. (2015). Calculation of COBIT metrics using a formal ontology. *2015 IFIP/IEEE International Symposium on Integrated Network Management (IM)*, 1384–1390. IEEE. doi: https://doi.org/10.1109/INM.2015.7140501

Wang, R., Zhang, Z. & Chen, G. (2008). Energy Function and Energy Evolution on Neuronal Populations. *IEEE Transactions on Neural Networks, 19(3)*, 535–538. doi: https://doi.org/10.1109/TNN.2007.914177

Wijayanti, N. Y., Setiawan, W., & Sukamto, R. A. (2017). Performance Assessment of IT Governance with Balanced Score Card and COBIT 4.1 of Universitas Pendidikan Indonesia. *In Journal of Physics: Conference Series, 812(1)*, 1–7. IOP Publishing. doi: https://doi.org/10.1088/1742-6596/812/1/012072

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 03/10/2018
Aceitação/Acceptance: 23/12/2018

Comparación de comportamiento de redes neuronales recurrentes con otros métodos en pronósticos con series de tiempo – caso series hidrológicas.

Carlos Hernán Fajardo-Toro[1], Luis Cobo[2], Paloma María Teresa Martínez-Sánchez[3]

chfajardo@universidadean.edu.co, lacobo@universidadean.edu.co, martinezpaloma@unbosque.edu.co

[1] Facultad de administración, Finanzas y ciencias económicas, Universidad Ean, 110221, Bogotá, Colombia.

[2] Facultad de Ingeniería, Universidad Ean, 110221, Bogotá, Colombia.

[2] Facultad de Ingeniería, Universidad el Bosque, 110221, Bogotá, Colombia.

Pages: 1022–1034

Resumen: Una de las actividades fundamentales en cualquier organización para poder realizar sus procesos de planificación y presupuestos es la elaboración de pronósticos. Para lograr lo anterior, los responsables de esto deben aplicar técnicas o modelos las cuales pueden ser cualitativos, cuantitativas o una mezcla de ambos. En cuanto a las técnicas cuantitativas, los modelos pueden causales o de series de tiempo, dependiendo de cómo se manejen los datos, manejo que va acorde a como se enfoque el problema. Para ambos casos se utilizan técnicas estadísticas, pero también técnicas de Inteligencia Artificial o adaptativas tales como las redes neuronales o la hibridación de diferentes técnicas. Este artículo se centra en modelos y técnicas para pronósticos de series de tiempo, y para ello presenta una comparación del comportamiento de las redes neuronales recurrentes – RNN tipo Elman contra otras técnicas estadísticas, así como y otras arquitecturas de redes tales como redes perceptrón multicapa o multi-layer perceptron - MLP, FeedForward de una capa, redes de base radial o Radial Basis Function - RBF y una red modular compuesta de redes MLP. Esto se aplica para el cálculo de pronóstico de caudales.

Palabras-clave: Pronósticos, Series de tiempo, Redes neuronales recurrentes, Elman-Jordan, caudales, redes neuronales

Comparison of the behavior of recurrent neural networks against other methods to obtain time series forecasts.

Abstract: One of the fundamental activities of any organization to be able to carry out its planning processes and budgets is the preparation of forecasts. To achieve this, those responsible for the planification must apply techniques or models. These models can be qualitative, quantitative or a mixture of both. As for quantitative

techniques, the models can be causal or for time series analysis; this depending on how the data is handled, according to how the problem is approached. For both cases statistical techniques are used, but also Artificial intelligence techniques such as neural networks or the hybridization of different techniques. This article focuses on models and techniques for time series forecasts, showing a comparison of the behavior of recurrent neural networks, Elman type, against other statistical techniques and different neural network architectures such as Multilayer perceptron - MLP, single layer FeedForward, Radial Basis Function netework - RBF and a modular network composed of MLP networks.

***Keywords**:* Forecasts, Time Series, Recurrent Neural Networks, Elman-Jordan, River Flows, neural networks

1. Introducción

Las redes neuronales artificiales – *Artificial Neural Networks* - ANN son algoritmos implementados en forma de programa informático o modelo electrónico, basados en la estructura del cerebro humano. El cerebro aprende fundamentalmente a base de experimentar situaciones. Es un hecho bien conocido que muchos problemas difíciles de resolver para ordenadores, son fácilmente solucionados por pequeños grupos de energía (neuronas) en los cerebros humanos (Haykin, 1994; Corchado et al., 2000). Los modelos del cerebro humano tienen una forma mucho más práctica para solucionar problemas que los proporcionados por máquinas. Incluso los cerebros de pequeños animales son capaces de realizar funciones que ningún ordenador puede resolver, como por ejemplo el reconocimiento de objetos desde diferentes proyecciones. Los ordenadores realizan operaciones mecánicas sin ningún problema, pero tienen problemas para reconocer patrones y más aún para generalizar eventos futuros a partir de acciones pasadas.

Las ANN actuales, intentan únicamente reproducir de forma simplificada los mecanismos más básicos del cerebro animal. No se trata de reproducir el cerebro humano, tan solo se centran en mecanismos de resolución de problemas individuales. El objetivo de las investigaciones teóricas en el campo de las ANN es el de entender la forma en la que los humanos resuelven los problemas, y utilizarla para mejorar las capacidades de los sistemas de Inteligencia Artificial tradicionales. Mientras que el cerebro humano tiene una estructura tridimensional y sus neuronas son capaces de interconectarse de forma incontrolada, las ANN son completamente diferentes, están compuestas por circuitos integrados y tienen un número limitado de capas e interconexiones entre las neuronas de cada nivel. La realidad física restringe el tipo y el ámbito de las ANN.

Las ANN requieren un proceso de entrenamiento previo a su utilización (Corchado et al., 2000). Este proceso comienza con la asignación de valores a los pesos asociados a cada conexión, y la definición de los parámetros de aprendizaje utilizados. En general, los valores de los pesos se eligen de forma aleatoria. Tras realizarse esta operación se lleva a cabo el entrenamiento de la red, lo cual conlleva la adaptación de estos pesos. Existen dos formas de aprendizaje claramente diferencia-das:

- Aprendizaje supervisado (supervised): requiere la presencia de pares de vectores, cada uno de ellos compuesto por los valores de salida de la red correspondientes a los valores de entrada. En este caso, es necesario que el entrenamiento de la red sea controlado por un supervisor o un conjunto de reglas.

- Aprendizaje no supervisado (unsupervised): en el cual se utilizan sólo los vectores de entrada para entrenar la red. Los valores de las salidas son determinados por la red durante el curso del aprendizaje. Se utiliza para construir modelos internos que capturan similitudes en los valores de los vectores de entrada, sin recibir información adicional.

1.1. Predicciones de series temporales

Las ANN poseen la capacidad de proyectar sucesos futuros a partir de experiencias pasadas (Weigend y Gershenfeld, 1994). Existen numerosas aplicaciones en las que se utilizan predicciones; por ejemplo, para ayudar en la definición de prioridades, anticipar futuros acontecimientos, etc. (Bishop, 1995). Predecir el valor de una variable un intervalo hacia delante, utilizando valores previos procedentes de una serie temporal, requiere la estimación de la función autorregresiva que define el proceso. Por tanto, el objetivo de cualquier sistema flexible es el de disminuir el error esperado de la predicción, de tal forma que la complejidad del modelo creado se ajuste a la complejidad de los datos. Por ejemplo, la red MLP y las regresiones B-splines poseen esta capacidad (Canu et al., 1997).

El análisis de series temporales se basa en:

- Predecir: avanzar acertadamente hechos futuros que tendrán lugar en un espacio corto de tiempo.
- Modelar: definir una descripción, que capture las características que definen la evolución de un sistema a largo plazo de forma adecuada.
- Caracterizar: intento de determinar las propiedades fundamentales que definen un sistema, utilizando una pequeña parte del conocimiento o poca información a priori.

Para el problema planteado en este trabajo se han elegido las redes de base radial o RBF (Radial Basis Function), de arquitectura modular, las redes perceptrón multicapa o Multi-layer Perceptron - MLP, la single-Layer feedforward y la Elman-Jordan, como mecanismos para realizar predicciones en series de tiempo; esto como base para el desarrollo de un modelo híbrido en investigaciones posteriores, así como una comparación con los métodos estadísticos tradicionales. Las capacidades de aprendizaje y adaptación a cambios continuos en las variables de entrada, así como su capacidad para manejar entornos no lineales, las convierten en candidatas ideales para implementar una predicción final revisada.

Este documento pretende mostrar como es el comportamiento de las redes neuronales recurrente tipo Elman en predicción de series de tiempo, caso series hidrológicas, comparándolas con otras topologías de redes neuronales y métodos estadísticos tradicionales, mostrando como el comportamiento de estas es bastante bueno con predictoras en este tipo de problemas.

2. Redes recurrentes

Los problemas de predicción constituyen una clase especial de los problemas de aproximación funcional en los que los valores de las variables de salida (predicciones) se determinan a partir de valores de salida obtenidos en instantes anteriores. Para tratar con

esta clase de problemas se introducen unidades de proceso que evolucionan en tiempo discreto, es decir, en el tiempo t ϵ {0,1,2,....}, de manera que sus salidas son transmitidas como entradas en el tiempo t+1. Estas se comportan como un autómata finito.

Sin embargo, el problema está en cómo entrenar dicho autómata para conseguir las secuencias de valores de salida deseados. Las redes neuronales recurrentes constan de una capa de sensores (recogen la información de las entradas), una capa intermedia de unidades ocultas y una capa de unidades de salida; además contienen conexiones sinápticas que van de los sensores de entrada a las unidades ocultas (también pueden ir directamente a las unidades de salida) y de las unidades ocultas a las unidades de salida, y se permiten conexiones sinápticas entre las unidades de proceso de una misma capa sobre todo entre las unidades de la capa oculta.

Se han propuestos varios modelos de redes recurrentes juntos con sus algoritmos de entrenamiento, como los modelos de Elman (1990), Jordan (1986), cuyo algoritmo de aprendizaje es esencialmente el algoritmo de retropropagación.

2.1. Redes Elman

En 1990, Jeffrey L. Elman en su trabajo Finding Structure in Time (Elman, 1990) desarrolló una red neuronal basada en el trabajo previo de Jordan (Jordan, 1986). Las redes de Jordan que presentan una topología de capas similar a las redes feedforward, con la inclusión de una capa de contexto que realimenta las señales de salida hacia las capas ocultas. En una red de Jordan, existen retardos en todas las realimentaciones. Adicionalmente, se considera que los pesos de las realimentaciones desde la capa de salida hacia la capa de contexto son unitarios.

Elman modifica esta arquitectura, considerando realimentaciones desde capas ocultas hacia la capa de contexto, sin considerar realimentaciones locales. La capacidad principal, consecuencia de su diseño recursivo, es que puede aprender a detectar y generar patrones temporales; de ahí se desprende su aplicación en procedimiento de señales y otros modelos donde el tiempo juega un papel determinante.

Esta red está constituida por una capa de sensores o unidades de entrada, dos capas intermedias de unidades (una capa de unidades ocultas y otra capa de unidades de contexto) y una capa de unidades de salida. Los sensores de entrada y las unidades de salida recogen la información del entorno, es decir, la suministrada por el conjunto de patrones de entrenamiento, mientras que las demás no. Las unidades de salida reciben la señal suministrada por la salida de las unidades de la capa oculta ponderadas por los correspondientes pesos sinápticos y suelen utilizar como función de transferencia la función identidad.

Sin embargo, la novedad de este modelo la introducen las unidades de contexto que se utilizan para memorizar las salidas de las unidades ocultas en la etapa anterior, de manera que cada unidad de contexto tiene como salida la salida de la unidad oculta correspondiente en la etapa anterior. Por lo tanto, esta red es sólo parcialmente recurrente y cada unidad de proceso oculta recibe como entrada las salidas de las unidades de contexto y de los sensores de entrada ponderadas por sus pesos sinápticos. De esta manera la salida de la red depende no sólo del patrón de entrada actual sino también de los patrones anteriores a través de las unidades de contexto.

Normalmente las funciones de activación de la capa recurrente suelen ser de tipo tangente hiperbólica y las de la capa de salida de tipo lineal. Esta configuración, permite a la red aproximar cualquier función (con un número finito de discontinuidades) con la precisión deseada. La única condición es que la capa oculta debe tener suficientes neuronas. Si la función que se necesita aproximar aumenta de complejidad será necesario incrementar el número de neuronas (Al-Hadithi et al., 2006).

2.2. Redes recurrentes y series de tiempo

Las series de tiempo han sido tratadas tradicionalmente como sistemas lineales, pero desde hace algún tiempo, los sistemas caóticos han venido cobrando importancia. Una señal cualquiera se puede representar a través de sistemas no lineales. Las redes neuronales han sido utilizas con bastante éxito para solventar este problema, donde las redes recurrentes han mostrado ser bastante eficientes para trabajar con series de tiempo, esto en sus diferentes topologías y características.

Aussen et al., (1994) implementaron redes recurrentes dinámicas en conjunto con métodos para reconocimiento de patrones en la estimación de temperaturas para realizar predicciones meteorológicas. También aplicaron el mismo tipo de redes para la predicción de niveles de visibilidad, variables que en conjunto permiten una mejor programación del telescopio en el observatorio astronómico en Cerro Paranal en Chile.

Ho et al., (2002) contrastan la eficiencia de las redes recurrentes frente a ARIMA en el pronóstico de la cantidad de fallos por unidad de tiempo en una planta en Noruega. More y Deo (2003) comparan la eficiencia de una red recurrente de Jordan con las MLP y los modelos ARIMA en el pronóstico del viento en zonas costeras.

Binner et al., (2006) comprara las redes recurrentes con dos modelos estadísticos utilizados para el cálculo de la inflación en EEUU. Contrasta las redes con modelos de benchmark y Markov Switching Autoregressive models. Jeong et al., (2008) desarrollan un modelo de pronóstico para predecir la dinámica del fitoplancton. Para lograrlo desarrolla una red neural recurrente temporalmente autorregresiva (TARNN, Temporal Autoregressive Recurrent Neural Network) y la contrasta con los modelos SARIMA (Seasonal Auto-Regressive Integrated Moving Average).

Las redes tipo Elman también han demostrado un buen comportamiento con series de tiempo univariadas, y han demostrado ser tan o más eficientes que los modelos estadísticos. Así como también frente a otras topologías como las MLP, las RBF o las SVM (Gençai y Liu, 1997; Recknagel, 2001).

Baker y Richards (1999) demuestran la eficiencia de la red Elman cuando la aplican para calcular el pronóstico del gasto en educación elemental y secundaria en los EEUU. Cho (2002) también lo demuestra al intentar pronosticar el flujo turístico hacia Hong Kong desde otros países. Brunelli et al., (2008) desarrollan un sistema de pronóstico basado en una red Elman para predecir el nivel de dióxido de sulfuro en la región de Priolo en Grecia. Dicho modelo resultó ser más eficiente que los modelos ARIMA.

Dia (2001) lleva a cabo una comparación de 4 redes neuronales para el pronóstico de tráfico de coches en Australia. Hace pruebas con MLP, redes semirecurrentes, redes dinámicas tipo TLRN (Time-Lag Recurrent Network) y redes híbridas. Las redes híbridas

son una combinación de análisis de componentes principales y TLRN. El autor concluye según sus datos, que las redes dinámicas se adaptan mejor a los procesos no lineales que las de tipo estático como las MLP.

Kalaitzakis et al., (2002) utilizan las redes recurrentes para el diseño e implementación de un sistema de pronósticos a corto plazo acerca de la carga de energía del sistema eléctrico de la Isla de Creta. Compara los métodos autorregresivos utilizados con los resultados obtenidos con una MLP, una RBF, una red neuronal recurrente en tiempo real y una rede neuronal recurrente autorregresiva para pronósticos de con-sumo y carga de 24 horas.

Mandal y Prabaharan (2005) construyen una red recurrente que denominan NARX (Nonlinear AutoRegressive models with eXogenous inputs) para el pronóstico de oleajes. Una red NARX consiste en una red MLP que toma como entrada los valores de entrada y salida proporcionales y genera una nueva salida. Calculan los valores de salida y los comparan con los resultados obtenidos en anteriores modelos desarrollados por otros autores obteniendo mejores índices de correlación.

Goh et al., (2006) Abordan el problema del pronóstico del viento, para lo que desarrollan y comparan dos topologías de redes recurrentes. Comparan el desempeño de las redes neuronales concurrentes totalmente conectadas (FCRNN, Fully Connected Recurrent Neural Network) y las redes neuronales recurrentes “entubadas” valoradas complejamente (CPRNN, Complex-valued Pipelined Recurrent Neural Net-works).

Brunelli et al (2007) desarrollan un modelo de pronóstico de concentraciones de compuestos químicos en la atmosfera para el área de Palermo en Italia. Para lograrlo inicialmente implementan dos redes, una red tipo Elman y una MLP, siendo la primera red la que proporciona los pronósticos más acertados.

Cao y Lin (2008) desarrollan una aplicación que permite pronosticar la irradiación del sol utilizando una red neuronal wavelet recurrente diagonal (DRWNN, Diagonal Recurrent Wavelet Neural Network) asistida con una técnica difusa. Una red neuronal wavelet es una ANN integrada con técnicas wavelet como funciones de activación (Cao y Cao, 2006; Benauda et al., 2006). Una red diagonal recurrente es una estructura relativamente simple basada en el algoritmo de retropropagación. No tiene conexiones cruzadas o transversales entre los nodos de la capa oculta o de la capa de salida, reduciendo el número de conexiones y por tanto el tiempo de proceso. Cuando la red se diseña con bases wavelet como funciones de activación, se convierte en una DRWNN (Cao y Lin, 2008).

En esta misma línea, Becerikli y Oysal (2006) diseñan una red denominada TDDNN (Time Delay Dinamic Neural Network) que definen como una red recurrente de tiempos continuos con retroalimentación con valores retardados. Con ella reali-zan el modelado y pronóstico de series caóticas.

Barbournis y Theocharis (2006) desarrollan su modelo con redes recurrentes para el cálculo de predicciones a largo plazo de la velocidad del viento. Para lograrlo explotan el concepto de las redes localmente recurrente basadas en el algoritmo de predicción recursiva del error. En este tipo de redes la recurrencia es localmente introducida dentro

de cada una de las neuronas de la red, y también se les conoce por el nombre de Local Feedback Multilayer Networks (LF-MLN).

Las redes recurrentes también han presentado una gran eficiencia trabajando con algoritmos evolutivos y genéticos. Pai et al., (2005) se interesa por las máquinas de soporte vectorial recurrente con algoritmos genéticos (RSVMG, Recurrent Support Vector Machine with Genetic Algorithm). Las SVM han sido utilizadas con éxito para resolver problemas no lineales o de series de tiempo, pero no muy frecuentemente en la elaboración de pronósticos (Pai et al., 2005). Los algoritmos genéticos se implementan para determinar los parámetros de la SVM. Los autores realizan una comparación de resultados obtenidos con la RSVMG con los alcanzados por una SVM, una MPL y un modelo autorregresivo, siendo muy superior en su desempeño la RSVMG. Versace et al., (2004), Cai et al., (2007) y Talib et al., (2008) implementan redes recurrentes con algoritmos híbridos evolutivos como algoritmos de aprendizaje, siendo bastantes eficientes a la hora de adaptarse en sistemas caóticos.

3. Diseño experimental y resultados

Los datos brutos utilizados para el presente estudio fueron tomados de la red de monitoreo de la Corporación Autónoma Regional del Valle del Cauca (CVC) que mantiene más de 150 hidrométricos y estaciones de lluvia que realizan la medición y el control de recursos hídricos en el valle y al norte del río Cauca. Bajo un acuerdo con la empresa de energía del pacífico (EPSA), la CVC supervisa el embalse de Salvajina para la producción de energía. En este contexto, uno de los objetivos de CVC es determinar la cantidad de agua que puede ser utilizado para la generación de energía, mientras se restringe el depósito.

Los datos disponibles se reciben y procesan diariamente en la estación hidrométrica de salvajina ubicada en el departamento de Cauca, en Colombia (ver Fig. 1). Esta información representa el registro diario del caudal del río (m3/s), a partir de mediciones de terreno proporcionadas por CVC desde los años 1950 a 2016.

Para generar un conjunto de datos listo para usar para entrenamiento, validación y fines de prueba, era necesario previamente preprocesar y analizar los datos brutos disponibles. El primer paso llevado acabo fue el tratamiento de los valores equivocados y faltantes. Inicialmente se procedió a depurar los datos inválidos o inexistentes. El primer filtro fue respecto a años con demasiados datos perdidos siguiendo una regla en hidrología que establece que, si falta más de un 15% de los datos, la muestra debe ser anulada. Para las series en las cuales había datos faltantes pero que cumplían con la regla de tener menos del 15%, las series fueron completadas a través de un proceso de ponderación en cual se estima el peso ponderado de ese valor en la serie total, en función del peso del mes frente al total de todo el año.

Para validar el enfoque propuesto, se han probado las redes utilizado el 60% de los datos disponibles para entrenar los modelos, 15% para estimar la mejor configuración parámetros y el restante 25% de los datos para fines de prueba. Esta validación el esquema y la distribución son comúnmente aceptados para probar el rendimiento de diferentes enfoques en el pronóstico de series de tiempo problemas (See et al, 2000).

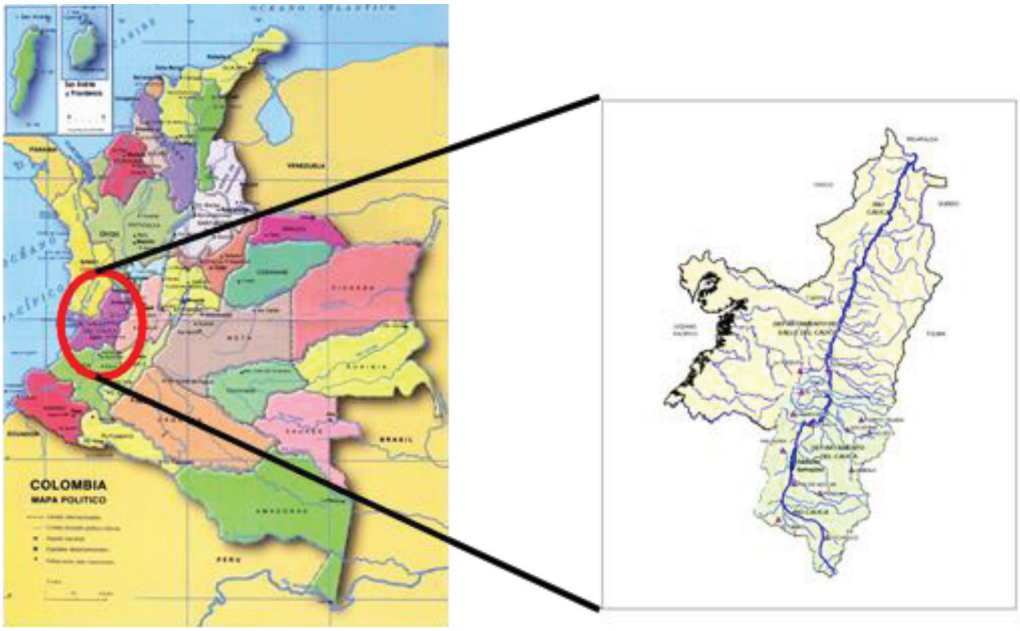


Figura 1 – Localización área de estudio

La topología de la red Elman escogida consta de una capa de entrada con 7 neuronas, que constituyen los días previos al pronóstico a generar, una capa intermedia de 20 neuronas y una capa de salida con una única neurona, correspondiente al caudal a predecir

La red MLP consta también de una capa de entrada de 7 neuronas que constituyen los días previos de pronóstico, y dos capas intermedias de 16 y 8 neuronas respectivamente con una única capa de salida que representa el pronóstico a cal-cular. La red Modular está compuesta de dos redes MLP que tienen la misma configuración de la red neuronal MLP que se evaluó de forma individual. Para la Single-Layer Feedforward también se tomaron 7 neuronas de entrada y una de salida.

El conjunto de datos de entrada son los 7 días previos al día a pronosticar. El entrenamiento se lleva acabo con una secuencia de 5000 iteraciones.

Para la Red RBF se realizaron varias simulaciones dando como mejor resultado el que se tomaran 6 días previos con 100 centros y 500 iteraciones.

Una vez realizadas las simulaciones, así como la aplicación de los modelos estadísticos tradicionales, se obtuvieron los siguientes resultados reflejados en las tablas 1 y 2, las cuales presentan los resultados obtenidos utilizando los modelos estadísticos y las redes neuronales. En cuanto a los modelos estadísticos, se dejaron los que dieron mejor resultado, en especial en lo que a ARIMA se refiere. Estos resultados se pueden ver en la Tabla 1. La tabla 2 muestra los resultados de las simulaciones realizadas con las redes neuronales. Como índice de comparación se decidió utilizar la Media absoluta de los errores - MAE.

Modelo	MAE
Naive	18,849
Media constante = 149,671	59,764
Tendencia lineal = 157,513 + -0,00537375 t	55,124
Suavización exponencial simple con alfa = 0,7486	18,11
Suavización exp. De Brown con alfa = 0,4889	18,807
Suavización exp. De Holt con alfa = 0,6327 y beta = 0,0009	19,143
ARIMA(2,0,2) con constante	17,674
ARIMA(1,0,1) con constante	17,916
ARIMA(2,1,2)	17,978
ARIMA(2,1,2) con constante	17,915
ARIMA(2,1,1)	17,026

Tabla 1 – Resumen de errores estadísticos obtenidos para el caudal.

Para evaluar si realmente hay o no diferencias en las medias de los errores entre los métodos, se requiere que hacer pruebas estadísticas sobre la distribución de los errores que demuestren que hay diferencias significativas. Inicialmente realizar una prueba de diferencias de medias es una alternativa válida, pero para aplicarse debe inicialmente cumplir el supuesto de normalidad.

Modelo	MAE
Pronóstico red feedforward	19,992
Pronóstico red Elman	18,345
Pronóstico red RBF	26,911
Pronóstico red Multilayer	58,210
Pronóstico red Modular	22,232

Tabla 2 – Resumen de los índices de errores por redes neuronales para el Caudal

En concreto, para el estudio de las series de error generadas por cada uno de los modelos analizados, se detectó que no tenían un comportamiento normal al realizar las pruebas de Kolmogorov-Smirnov dado el tamaño de muestra. Lo anterior hace que se invalide cualquier test paramétrico que se realice para comparar las muestras. Por este motivo se ha recurrido al contraste de Kruskall-Wallis. Esta prueba representa la alternativa no paramétrica para situaciones como la anteriormente mencionada, centrándose en el análisis de las medianas en lugar del de las medias. La única exigencia que se establece es la aleatoriedad en la extracción de las muestras, no haciendo referencia a ninguna de las otras condiciones adicionales de homocedasticidad y normalidad necesarias para la aplicación cualquier test paramétrico.

Los resultados generados por el contraste de Kruskall-Wallis sobre los datos analizados, muestran, con un nivel de confianza del 99%, que existen diferencias significativas entre los distintos modelos (valor de $p < 0,01$).

Con el fin de profundizar en el análisis de las diferencias significativas encontradas por el contraste de Kruskall-Wallis, se ha realizado la prueba de Mann-Whitney, que permite comparar los sistemas analizados dos a dos, para detectar entre qué modelos existen las diferencias significativas encontradas por la prueba de Kruskall-Wallis.

La Tabla 3 muestra el resumen de la ejecución de la prueba de Mann-Whitney al 99%, para todas las combinaciones posibles de los modelos analizados

Como se podrá observar, las redes Elman y las Feedforward realmente no presentaron diferencias significativas con varias de las técnicas estadísticas, en especial las más robustas como son los modelos ARIMA. Aunque habría que hacer test con más tipos de series de tiempo, se podría decir que el uso de las redes recurrentes resulta de bastante utilidad, dada la dificultad numérica que representan algunos modelos estadísticos, es decir, el grado de experticia en econometría que requiere el decisor para la configuración de los modelos autorregresivos.

4. Conclusiones y trabajo futuro

En este artículo se presentó un comparativo del comportamiento que tienen las redes neuronales, en especial las de tipo recurrente, frente al pronóstico de series de tiempo las cuales presentan un comportamiento estocástico no lineal.

Para lo anterior se tomaron los métodos estadísticos tradicionales y se compararon con 5 redes diferentes, Feedforward, Elman-Jordan, RBF, MLP y una modular compuesta por 2 MLPs internamente, haciendo una comparación de las Medias Absolutas de las Desviaciones – MAE.

	Arima (1,0,1)	**Arima (2,0,2)**	**Arima (2,1,1)**	**Arima (2,1,2)**	**Arima (2,1,2)**	**Naive**	**M. simple**	**Brown**	**Holt**	**M. Lineal**	**Elman**	**FF**
Arima(1,0,1)												
Arima(2,0,2)	=											
Arima(2,1,1)	*	*										
Arima(2,1,2)	*	*	=									
Arima(2,1,2)	*	*	=	*								
Naive	*	*	*	*	*							
M. simple	*	*	*	*	*	*						
Brown	*	*	*	*	*	*	*					
Holt	*	=	*	*	*	*	*	*				
M. lineal	=	=	*	*	*	=	*	*	*			
Elman	=	=	=	=	=	=	*	=	*	*		
FeedForward	*	*	=	=	=	*	*	=	=	*	=	

Tabla 2 – Resumen de los índices de errores por redes neuronales para el Caudal

Posterior a esta comparación, y en especial que los errores son bastante similares, se busca realizar pruebas estadísticas que indiquen si existen diferencias representativas

y dado que no tiene un comportamiento normal, se decidió aplicar las pruebas no paramétricas de Kruskall-Wallis y Mann-Whitney las cuales, a través del análisis de medianas, nos pueden indicar si existen diferencias significativas de los errores entre los resultados de las redes neuronales y los métodos estadísticos.

Una vez realizadas las pruebas se logra ver cómo, estadísticamente hablando y con base en las series de tiempo de caudal, no se considera que haya diferencias significativas en la distribución de los errores de los métodos de recurrentes con la mayoría de métodos de mejor precisión estadística.

De lo anterior se puede concluir como este tipo de redes resulta especialmente útil debido a su retropropagación para adaptarse bien al comportamiento de las series de tiempo, que como ya se mencionó, son de naturaleza estocástica no lineal.

Por tanto, se puede considerar una ventaja para los tomadores de decisiones el uso de este tipo de redes en pronósticos dado que, para implementarlas, no se requiere un bagaje matemático amplio como si requiere el conocimiento y definición de modelos de series de tiempo, en especial los autorregresivos.

Como trabajo futuro está el construir modelos más complejos, en los cuales se puedan hacer hibridación de diferentes técnicas de clasificación y pronósticos las cuales contribuyan a disminuir el error o hacer que la distribución de los errores sea más reducida, es decir, con una menor variación con trabajos como los realizados por Toro et al (2013) en la cual se hace uso de sistema de razonamiento basado en casos.

Referencias

Al-Hadithi B. M., Rueda M. R., Muro J. S. (2006) Interfaz visual para el prototipado rápido de clasificadores de gajos de mandarina basados en redes neuronales, Revista de Ciencia, Tecnología y Medio Ambiente, VOLUMEN IV.

Aussem, A., Murtagh, F., Sarazin, M. (1994) Dynamical recurrent neural networks and pattern recognition methods for time series prediction: Application to seeing and temperature forecasting in the context of ESO´s VLT astronomical weather station. Vistas in astronomy, 38(3):357–374.

Barbounis, T. G. y Theocharis, J. B. Locally recurrent neural networks for long-term wind speed and power prediction. Neurocumputing, 69:466–496.

Becerikli, Y. y Oysal, Y. (2007). Modeling and prediction with a class of time delay dynamic neural networks. Applied Soft Computing, 7:1164–1169.

Binner, J. M., Elger, C. T., Nilsson, B., Tepper, J. A. (2006). Predictable non-linearities in U.S. inflation. Economics letters, 93:323–328.

Bishop, C. R. (1995). Neural Networks for Pattern Recognition. Carendon Press, Oxford.

Brunelli, U., Piazza, V., Pignato, L., Sorbello, F., Vitabile, S. (2007). Two-days ahead prediction of daily maximum concentrations of SO2, O3, PM10, NO2, CO in the urban area of Palermo, Italy. Atmospheric Environment, 41:2967–2995.

Cai, X., Zhang, N., Venayagamoorthy, G. K., Wunsch II, D. C. (2007). Time series prediction with recurrent neural networks trained by a hybrid PSO-EA algorithm. Neurocomputing, 70:2342–353.

Canu, S., Grandvalet, Y., y Ding, X. (1997). Une application des réseaux de neurones pour la prévision à un pas de temps. Statistiques et méthodes neuronales. (pp. 120-131), Dunod.

Cao, J. C. y Cao, S. H. (2006) Study of forecasting solar irradiance using neural networks with preprocessing sample data by wavelet analysis. Energy, 31:3435–3445.

Cao, J. y Lin, X. (2008). Application of the diagonal recurrent wavelet neural network to solar irradiation forecast assisted with fuzzy technique. Engineering Applications of Artificial Intelligence, article in press.

Cao, J. y Lin, X. (2008). Study of hourly daily solar irradiation forecast using diagonal recurrent wavelet neural networks. Energy Conversion & Management, 49: 1396–1406.

Cho, V. (2003) A comparison of three different approaches to tourist arrival forecasting. Tourism Management, 24:323–330.

Corchado, J. M., y Lees, B. (2000). Adaptation of cases for case-based forecasting with neural network support. En Pal, S. K., Dilon, T. S., y Yeung, D. S. (Eds.). Soft Computing in Case Based Reasoning. Springer-Verlag: London.

Dia, H. (2001). an Object-oriented neural network approach to short-term traffic forecasting. European Journal of operational Research, 131:253–261.

Elman, J. L. (1990): Finding structure in time, Cognitive Science, 14:179–211

Gençay R., Liu, T. (1997) Nonlinear modelling and prediction with feedforward and recurrent networks. Physica D, 108:119–134.

Goh, S. L., Chen, M., Popovic, D. H., Aihara, K., Obradovic, D., Mandic, D. P. (2006). Complex-valued forecasting of wind profile. Renewable Energy. 31:1733–1750.

Haykin S. (1994) Neural Networks, A comprehensive Foundation. Macmillan College Publishing Company, New York, NY.

Ho, S. L., Xie, M., Goh, T. N. (2002). A comparative study of neural network and Box-Jenkins ARIMA modeling in time series prediction. Computers and industrial engineering, 42:371–375.

Jordan, M. I. (1986). Serial order: A parallel distributed processing approach. Institute for Cognitive Science Report. University of California, San Diego.

Kalaitzakis, K., Stavrakakis, G. S., Anagnostakis, E. M. (2002). Short-term forecasting based on artificil neural networks parallel implementation. Electric Power Systems Research, 63:185–196.

Mandal, s. y Prabaharan, N. (2006). Ocean wave forecasting using recurrent neural networks. Ocean Engineering, 33:1401–1410.

More, A. y Deo M. C. (2003). Forecasting wind with neural networks. Marine Structures, 13:35–49.

Pai, P. F. y Hong W. C. (2005). Forecasting regional electricity load based ib recurrent support vector machines with genetic algorithms. Electric Power Systems Research, 74:417–425.

Recknagel, F. (2001). Applications of machine learning to ecological modelling. Ecological Modelling, 146:303-310.

See L, Openshaw, S, (2000) A hybrid multi-model approach to river level forecasting, Hydrological Sciences Journal 45 (4) 523–536.

Talib, A., Recknagel, F., Cao, H., Van der Molen. D. T. (2008). Forecasting and explanation of algal dynamics in two shallow lakes by recurrent artificial neural network and hybrid evolutionary algorithm. Mathematics and Computers in Simulation, 78:424–434.

Toro CHF, Gómez Meire S, Gálvez JF et al (2013) A hybrid artificial intelligence model for river flow forecasting. Appl Soft Computing 13:3449–3458

Versace, M., Bhatt, R., Hinds, O., Shiffer, M. (2004). Predicting the exchange traded find DIA with combination of genetic algorithms and neural networks. Expert systems with Applications, 27:417–425.

Weigend, A. S., y Gershenfeld, N. A. (1994). Time series prediction: Forecasting the future and understanding the past. Santa Fe institute studies in the sciences of complexity, Addison-Wesley Publishing Company.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 04/09/2018
Aceitação/Acceptance: 09/11/2018

Propuesta de aplicación de minería de procesos para evaluar las rutas de aprendizaje de estudiantes con discapacidad visual en cursos en línea

Diego Buenaño-Fernández[1], Sergio Luján-Mora[2]

diego.buenano@udla.edu.ec, sergio.lujan@ua.es

[1]Universidad de Las Américas, José Queri, 170137, Quito, Ecuador.

[2] Universidad de Alicante, San Vicente del Raspeig, 03080, Alicante, España

Pages: 1035–1047

Resumen: La culminación exitosa de un curso en línea por parte de estudiantes con algún tipo de discapacidad está relacionada directamente con la inclusión o no de criterios de accesibilidad. Las aplicaciones web, y en particular los cursos en línea, están conformados por un sinnúmero de componentes cuyo objetivo es hacer más fácil la navegación de los usuarios. La exclusión de criterios de accesibilidad en el diseño de estos componentes puede llegar a convertirse en barreras de aprendizaje para estudiantes con algún tipo de discapacidad. En la literatura revisada, se han identificado varios trabajos relacionados con el análisis de problemas de accesibilidad en entornos educativos virtuales. Sin embargo, hay poca investigación vinculada con el análisis de rutas de aprendizaje seguidas por personas con algún tipo de discapacidad. En este trabajo se propone aplicar herramientas de minería de datos educativos y de minería de procesos para detectar puntos potencialmente peligrosos en las rutas de aprendizaje que siguen este grupo de estudiantes.

Palabras-clave: Analítica de aprendizaje; Discapacidad visual; Minería de procesos educativos; Web Content Accessibility Guidelines (WCAG) 2.1; World Wide Web Consortium (W3C).

Proposal of application of process mining to evaluate the learning routes of students with visual disability in online courses

Abstract: The successful completion of an online course by students with some type of disability is directly related to the inclusion or not of accessibility criteria. Web applications, and in particular online courses, are made up of a number of components whose objective is to make it easier for users to navigate. The exclusion of accessibility criteria in the design of these components can become learning barriers for students with some type of disability. In the literature reviewed, several works related to the analysis of accessibility problems in virtual educational environments have been identified. However, there is little research linked to the analysis of learning routes followed by people with some type of disability. In this project, it is proposed to apply educational data mining and process mining tools

to detect potentially dangerous points in the learning routes followed by this group of students.

***Keywords**:* Learning Analytics; Visual disability; Educational Process Mining; Web Content Accessibility Guidelines (WCAG) 2.0; World Wide Web Consortium (W3C).

1. Introducción

En las plataformas educativas mediadas por las tecnologías de la información y comunicación (TIC) se puede encontrar una amplia gama de recursos interactivos de aprendizaje. Este tipo de recursos va tomando fuerza debido al aparecimiento de nuevos modelos pedagógicos (Amo et al., 2018), que ponen énfasis en las actividades interactivas que realiza el estudiante en la plataforma virtual. El incremento de oportunidades de formación en todas las áreas de conocimiento está relacionado con la aparición de cursos en línea, sean estos masivos (*Massive Open Online Course*, MOOC) o privados (*Small Private Online Course*, SPOC). Sin embargo, es necesario que los administradores de estos cursos no se descuiden en la inclusión de criterios de usabilidad y accesibilidad. Caso contrario, no podrán ofrecer un verdadero aprendizaje a los estudiantes inscritos, incluyendo a las personas con algún tipo de discapacidad y adultos mayores (Iniesto, Mcandrew, Minocha & Coughlan, 2008).

Por otro lado, la investigación en analítica de aprendizaje (AA) ha generado un espacio importante para el desarrollo de herramientas de análisis de datos educativos. En este contexto aparecen los estándares y especificaciones técnicas para tecnologías de aprendizaje tipo SCORM o xAPI. Estas especificaciones permiten la creación de objetos pedagógicos estructurados, que cumple con los requerimientos de: adaptabilidad, interoperabilidad, accesibilidad, reusabilidad. A través de la implementación de las directrices propuestas en xAPI, se podrá entender el significado de los datos almacenados en el repositorio de aprendizaje conocido como repositorio de registros de aprendizaje (*Learning Record Store,* LRS). A través del análisis de los datos almacenados en el LRS de xAPI se pueden identificar las rutas de aprendizaje que siguen los estudiantes al navegar por un curso en línea. Además, que se pueden identificar los posibles puntos críticos de aprendizaje que se generan en dichas rutas.

Varios trabajos previos (Iniesto, McAndrew, Minocha, & Coughlan, 2016), (Iniesto, McAndrew, Minocha, & Coughlan, 2014), (Ngubane-Mokiwa, 2016) se orientan a evaluar y validar la aplicación de criterios de accesibilidad sobre todo en MOOC. Sin embargo, no se han orientado a analizar detalladamente los elementos del curso que pueden presentar barreras de accesibilidad. Por lo tanto, en el presente trabajo, apoyados por la minería de procesos educativos (*Educational Process Mining*, EPM) se plantea identificar y listar los elementos del curso que necesitan una revisión del autor. Estos elementos son los que no permiten que los estudiantes puedan avanzar con el aprendizaje de la manera planificada.

La EPM es una alternativa emergente en el campo de la minería de datos educativos (*Educational Data Mining*, EDM). La EPM analiza los datos de los archivos de registro, archivos de base de datos o archivos de pistas de auditoría almacenados en plataformas educativas. Estos archivos contienen una secuencia de eventos producto de la interacción

de estudiantes con las plataformas virtuales de aprendizaje. El objetivo de la EPM es descubrir, analizar y proporcionar una representación visual del proceso educativo completo (Bogarín, Cerezo, & Romero, 2018).

El documento se estructura de la siguiente manera. La sección 2 constituye una fundamentación teórica, en la que se detallan conceptos relacionados con la propuesta. La sección 3 presenta la metodología utilizada para implementar el trabajo. Finalmente, la sección 4 contiene las conclusiones del trabajo realizado.

2. Fundamentación teórica

2.1. Evaluación de accesibilidad web en MOOC

La continua evolución de los MOOC, desde su nacimiento hace aproximadamente una década, ha potenciado su alcance, brindando un sinnúmero de oportunidades de aprendizaje para personas de todo el mundo. Sin embargo, para alcanzar su máximo potencial, los MOOC deben cumplir con criterios de accesibilidad para poder llegar a diversos estudiantes, con o sin discapacidades. En la revisión de la literatura, hemos encontrado algunas investigaciones publicadas que están relacionadas con la evaluación de accesibilidad de contenido en plataformas MOOC.

Un estudio presentado en el año 2013 (Sanchez-Gordon & Lujan-Mora, 2013) mostró los resultados preliminares de un análisis de accesibilidad realizado sobre un número pequeño de MOOC. Las conclusiones fueron que las WCAG 2.0 no incluyen requisitos de accesibilidad para personas de edad avanzada. Adicionalmente, los autores concluyeron que debería haber un aumento de conciencia por parte de los autores de contenido de cursos web, con el propósito de producir contenido accesible.

En el año 2014, una investigación propuso dos categorías de accesibilidad que deben ser consideradas en los contenidos web y las plataformas MOOC. Las categorías presentadas fueron personales y no personales. Los autores concluyeron que la no inclusión de criterios de accesibilidad web en los MOOC podría discriminar la participación de personas con algún tipo discapacidad. Es decir, los MOOC no cumplían uno de sus principales objetivos: brindar acceso a todas las personas (Sanchez-Gordon & Luján-Mora, 2014).

En el trabajo realizado por Iniesto, McAndrew, Minocha, & Coughlan (2017), se presentó una serie de iniciativas orientadas a apoyar a los proveedores de MOOC para lograr una mayor accesibilidad en sus plataformas. Por otro lado, también se plantearon sugerencias para los estudiantes con discapacidad orientadas a mejorar su aprendizaje permanente. De esta revisión de artículos relacionados con la evaluación de criterios de accesibilidad en las plataformas y recursos de los MOOC, se puede evidenciar una necesidad latente de realizar propuestas orientadas a mitigar la discriminación de acceso para personas con una o varias discapacidades, así como también para adultos mayores.

2.2. Accesibilidad web

La accesibilidad web se conoce como el conjunto de características incluidas en el diseño web, que hacen posible que todas las personas puedan percibir, entender, navegar e

interactuar en la red (Luján-Mora, 2013). El World Wide Web Consortium (W3C) a través de la Iniciativa para la Accesibilidad Web (*Web Accessibility Initiative*, WAI) ha propuesto tres estándares enfocados en la mejora de la accesibilidad web (World Wide Web Consortium, 2018a):

1. Pautas de accesibilidad al contenido web (Web Content Accessibility Guidelines, WCAG).
2. Pautas de accesibilidad de las herramientas de autor (Authoring Tool Accessibility Guidelines, ATAG).
3. Pautas de accesibilidad de los agentes de usuario (User Agent Accessibility Guidelines, UAAG).

Adicionalmente, la WAI ha planteado técnicas que deben aplicarse según las circunstancias y documentación de fallas con el objetivo de desarrollar y evaluar el contenido web, tal como se muestra en la Figura 1.

A continuación, se presenta información de las pautas de accesibilidad WCAG, ATAG y UAAG.

La versión 1.0 de las WCAG fue presentada en mayo de 1999, la versión 2.0 en diciembre del 2008. La versión 2.1 es una recomendación publicada en junio de 2018 (World Wide Web Consortium, 2008), estos lineamientos presentan un conjunto de recomendaciones que explican cómo hacer que el contenido web sea más accesible para las personas con discapacidad.

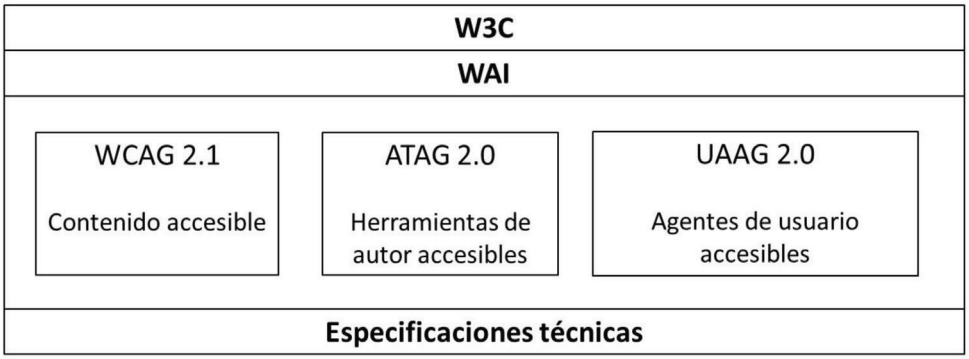


Figura 1 – Pautas de accesibilidad propuestas por W3C

Las WCAG 2.1 consideran esenciales cuatro principios de accesibilidad (World Wide Web Consortium, 2018b):

- Perceptible: la información presentada en una página web debe ser percibida por los usuarios.
- Operable: los usuarios deben poder navegar y operar los componentes de la interfaz.
- Comprensible: los usuarios deben ser capaces de comprender la información y el funcionamiento de la interfaz.

- Robusto: los agentes de usuario, incluidas las tecnologías de apoyo, deben permitir a los usuarios acceder al contenido.

Los lineamientos de la WCAG 2.1 incluyen 13 directrices y 78 criterios de éxito. Los criterios de éxitos de accesibilidad web son clasificados en 3 niveles de conformidad: el nivel de conformidad A (nivel mínimo de accesibilidad), nivel AA (nivel medio de accesibilidad) y nivel AAA (nivel alto de accesibilidad). W3C sugiere utilizar WCAG 2.1 con el fin de atender de mejor manera las necesidades de más personas. Es importante resaltar que la WAI mantiene compatibilidad con versiones anteriores a los lineamientos propuestos en WCAG 2.1 (Kirkpatrick & Michael, 2018).

En relación a las herramientas de edición de contenidos ATAG, son un conjunto de normas que deben cumplir las herramientas de autor para producir contenido web accesible. Las ATAG 2.0 se publicaron en septiembre de 2015 y están constituidas por dos partes (World Wide Web Consortium, 2015a):

- Parte A: incluye principios y pautas para el diseño de herramientas de autor de contenido web que son más accesibles para autores con discapacidad.
- Parte B: incluye principios y pautas para apoyar y promover la producción de contenido web accesible por parte de los autores.

En ATAG 2.0 se incluye 24 pautas (13 parte A, 11 parte B). Las pautas, a la vez, están divididas en 89 criterios de éxito (33 parte A, 56 parte B) y 3 niveles de conformidad (A, AA y AAA).

En relación a las UAAG, son pautas que recomiendan cómo hacer que los agentes de usuario (reproductores multimedia, navegadores, y productos de apoyo) sean accesibles para las personas con discapacidad. Las UAAG 2.0 fueron propuestas en diciembre del 2015 y están organizado en 5 principios (perceptible, operable, comprensible, acceso programático, especificaciones y convenciones) y 25 pautas (World Wide Web Consortium, 2015b).

En la Tabla 1, se describen a modo de ejemplo ocho de las principales barreras de accesibilidad que deben vencer las personas con discapacidad visual al acceder a recursos web.

2.3. La minería de procesos educativos

La minería de procesos tiene como objetivo central extraer conocimiento a partir de los registros de eventos almacenados en un sistema de información. La EPM propone la construcción de modelos de procesos educativos completos y compactos a través de los cuales se pueda reproducir todo el comportamiento de un estudiante. A través de la EPM se puede verificar si el comportamiento modelado coincide con el comportamiento observado (Bogarín et al., 2014).

En el artículo propuesto por (Bogarín et al., 2018) se presentó un marco de referencia específico para ejecutar tareas de minería de procesos en entornos educativos. En la Figura 2 se puede observar un esquema general con el flujo de información y los conceptos involucrados en la EPM. Este flujo está compuesto por los siguientes elementos: actores del proceso enseñanza aprendizaje (estudiantes y docentes), plataformas de aprendizaje virtual, registros de interacción estudiante- plataforma y modelos de procesos.

No.	Barreras de accesibilidad	Criterio de éxito WCAG 2.1	Descripción	Nivel de conformidad WCAG 2.1
1	*Carece del mapa de contenidos o mal estructurado.*	Bloques de derivación.	Un mecanismo está disponible para eludir bloques de contenido que se repiten en múltiples páginas web.	A
2	*Uso de CAPTCHA sonoro con un idioma diferente al del usuario.*	Contenido no textual.	Se proporcionan alternativas de texto que identifican y describen el propósito del contenido no textual, y formas alternativas de CAPTCHA usando modos de salida para diferentes tipos de percepción.	A
3	*Página web sin título.*	Título de la página.	Las páginas web tienen títulos que describen un tema o propósito.	A
4	*Imágenes sin texto alternativo que describa su contenido.*	Contenido no textual.	Todas las imágenes, los botones de imagen y los puntos calientes del mapa de imagen tienen un texto alternativo equivalente apropiado.	A
5	*Los subtítulos no están sincronizados con el video.*	Subtítulos (Pregrabados)	Los subtítulos sincronizados se proporcionan para video no en vivo, basado en la web.	A
6	*Tablas cuyo contenido resulta incompatible cuando se lee de forma secuencial.*	Información y relaciones.	La información, la estructura y las relaciones transmitidas a través de la presentación se pueden determinar mediante programación o están disponibles en texto.	A
7	*Documentos PDF no accesibles.*	Secuencia significativa.	Garantizar la correcta tabulación y orden de lectura en documentos PDF.	A
8	*Al expirar una sesión de trabajo el usuario pierde la información cuando vuelve a ingresar.*	Re/autenticación.	Cuando una sesión autenticada caduca, el usuario puede continuar la actividad sin pérdida de datos después de volver a autenticarse.	AAA

Tabla 1 – Barreras de accesibilidad web para personas con discapacidad considerando las WCAG 2.1

A través de la aplicación de herramientas de AA es posible capturar, en diferentes niveles de granularidad, los eventos, acciones y actividades de los estudiantes. Estos eventos van desde acciones de bajo nivel, como pulsaciones de teclas y clics del ratón, hasta eventos de mayor nivel, como las actividades de aprendizaje de los estudiantes (Trcka, Pechenizkiy, & Van der Aalst, 2010).

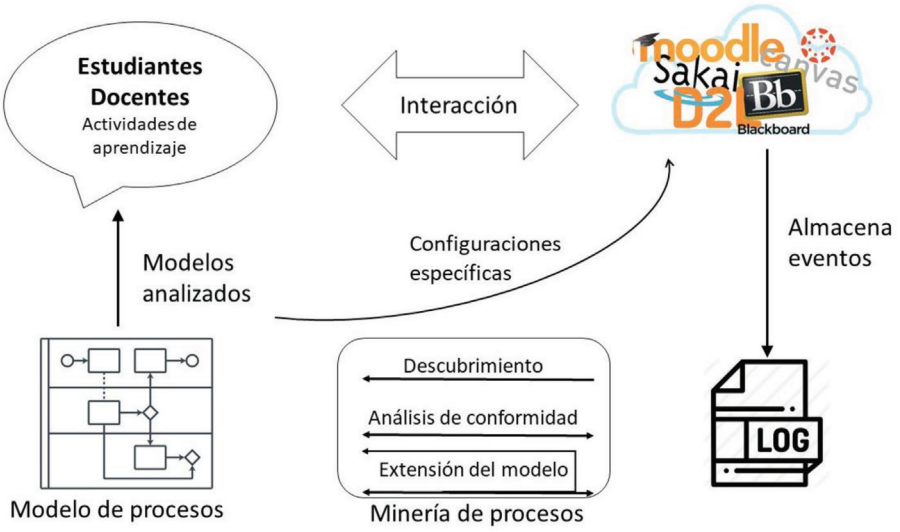


Figura 2 – Elementos de la minería de procesos

Existen diversas aplicaciones de software gratuitas y licenciadas que se utilizan para llevar a cabo la minería de procesos. En la revisión bibliográfica realizada para el presente trabajo, se han identificado tres aplicaciones que se usan específicamente en la EPM. La herramienta ProM[1] es un marco de referencia genérico de código abierto, es el más completo y el que comúnmente se ha aplicado. El software Disco[2] es una herramienta de propósito general aplicada en el área de minería de procesos, es una aplicación comercial y utilizada con frecuencia en el campo educativo. Finalmente, la herramienta SoftLearn[3] que es una aplicación de propósito específico orientada a la minería de procesos en el área educativa.

En relación con las técnicas de minería de procesos que comúnmente se han utilizado en trabajos realizados en el área educativa se han identificado las siguientes: 1) Técnicas de descubrimiento de procesos; 2) Técnicas de análisis de conformidad; y 3) Técnicas de extensión del modelo de proceso (Pechenizkiy, Trčka, Vasilyeva, Aalst, & Bra, 2009).

Las técnicas de descubrimiento de procesos, se enfocan en la creación de un modelo de procesos basado exclusivamente en un registro de eventos. Es decir, describe la información del archivo de registro de una manera compacta y precisa. Los algoritmos más utilizados en el dominio educativo son: algoritmos Alpha, algoritmos de minería heurísticos, algoritmos genéticos y algoritmos de minería difusos (Bogarín et al., 2018).

La técnica de análisis de conformidad, requiere de un modelo de proceso teórico preexistente que se pueda comparar con el modelo descubierto creado a partir de eventos

[1] http://www.promtools.org
[2] https://fluxicon.com/disco/
[3] https://tec.citius.usc.es/SoftLearn/

de registro. El propósito es confirmar si el comportamiento observado está de acuerdo con las suposiciones del modelo teórico (Tóth, Rölke, Goldhammer, & Barkow, 2017).

La técnica de extensión del modelo de proceso también involucra un modelo de procesos teórico existente. Pero a diferencia de las técnicas de análisis de conformidad, en esta técnica se busca extender o mejorar los modelos de proceso basados en los datos de comportamiento observados (Van der Aalst, 2012).

La aplicabilidad de la minería de procesos en el área educativa se ha extendido hacia diferentes aplicaciones. En el trabajo realizado por Bogarín et al., (2018) se ha propuesto una clasificación de estos aportes en las siguientes áreas: Aprendizaje colaborativo soportado por computadora, cursos de entrenamiento profesional, minería de currículum, registro de estudiantes, evaluaciones basadas en computador, entre otros. En la literatura revisada no se han encontrado trabajos relacionados con la aplicación de EPM orientada a la identificación de puntos críticos en las rutas de aprendizaje seguidas por personas con discapacidad visual.

2.4. Herramientas de analítica de aprendizaje

Las herramientas de AA pueden aportar un marco de referencia ideal para analizar los procesos de enseñanza-aprendizaje que se dan en contextos educativos mediados por las TIC. La AA se puede entender como la recopilación, medición, análisis y visualización de datos orientados a entender el comportamiento de los estudiantes y así poder optimizar su experiencia de usuario (Amo et al., 2018). Por ejemplo, las acciones como reproducir, detener, rebobinar, reenviar o repetir videos se pueden rastrear aplicando diferentes herramientas de AA.

Las herramientas de AA son usadas para recopilar datos producto de la interacción que realizan los estudiantes cuando trabajan en entornos virtuales de aprendizaje. Dependiendo del entorno, los datos recopilados se almacenan en archivos de registro y generalmente son datos sin ninguna estructura establecida. Por lo cual, sobre dichos datos es necesario realizar obligatoriamente un trabajo detallado de pre procesamiento. La elección de una herramienta adecuada es fundamental en el proceso de AA. Algunas de estas herramientas se fundamentan en el concepto de *clickstream*. De acuerdo a (Chan, Chiu, Ng, & Kwong, 2015) la definición de *clickstream* hace referencia al registro de la ruta de navegación que realiza un usuario mientras navega por la web. En los entornos virtuales de aprendizaje, los datos de *clickstream* apoyan la gestión docente, ya que les permite visualizar la interacción de los estudiantes con las diversas actividades de aprendizaje propuestas.

Tin Can API o xAPI[4] es una especificación que permite registrar una amplia gama de experiencias que un usuario puede tener al interactuar con diferentes plataformas educativas. Esta especificación es de código abierto y permite almacenar datos con alto grado de granularidad en un LRS. En este repositorio se almacenan las actividades de interacción que realiza un estudiante cuando utilizan entornos de aprendizaje mediados por las TIC (Lim, 2015). El LRS almacena declaraciones de estado (*statements*) o registros que están relacionados con las actividades de aprendizaje. Estas declaraciones siguen un patrón sencillo basado en: Sujeto + Verbo (acción) + Objeto de aprendizaje. La aplicación

[4] https://xapi.com/overview/

xAPI y el LRS pueden existir sin necesidad de una plataforma virtual de aprendizaje, este concepto nos acerca hacia la generación de cursos en línea autogestionables y actividades de aprendizaje independientes de la plataforma tecnológica.

3. Metodología

Uno de los principales inconvenientes que se tiene a la hora de identificar problemas de accesibilidad en cursos en línea es que no se han desarrollado mecanismos que permitan ubicar exactamente cuáles son los componentes de un determinado curso que presentan barreras de accesibilidad (Cooper, Ferguson, & Wolff, 2016). Este problema persiste a pesar de que existen varios trabajos orientados a realizar la evaluación la accesibilidad en cursos en línea. Por lo tanto, un objetivo que se propone en el presente trabajo es identificar y listar los componentes del curso que presentan problemas de accesibilidad, ya que son estos los que impiden avanzar a los alumnos que presentan algún tipo de discapacidad.

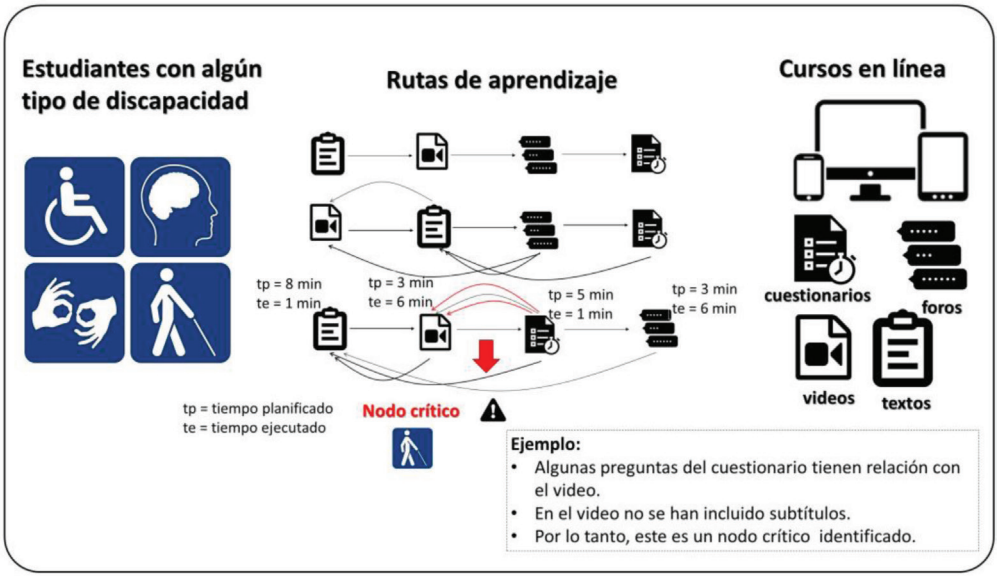


Figura 3 – Esquema que describe el problema identificado

Para cumplir con el objetivo propuesto se han implementado dos cursos en línea utilizando la plataforma Angular. Esta herramienta es un framework de código abierto que se utiliza para crear y mantener aplicaciones web de una sola página (*Single-page application*). Esta tecnología está compuesta por componentes individuales que se pueden reemplazar o actualizar de forma independiente, sin actualizar toda la página (Jadhav, Sawant, & Deshmukh, 2015). La estructura y contenidos de los cursos serán iguales. El primer curso deberá incluir criterios de accesibilidad para personas con discapacidad visual, mientras que el segundo no incluirá criterios de accesibilidad. En la Figura 4 se muestra un esquema gráfico de la propuesta.

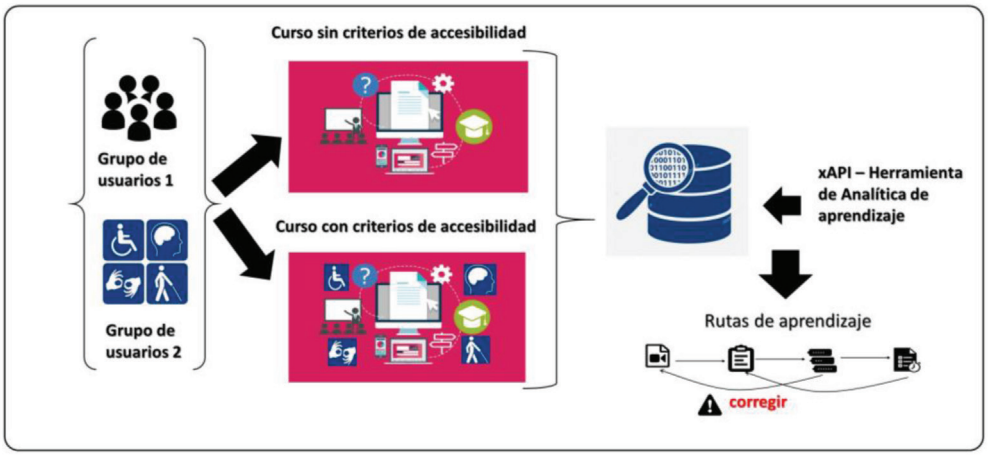


Figura 4 – Esquema de la solución planteada

Para capturar las trazas de navegación o actividades de aprendizaje del curso, es necesario configurar en xAPI declaraciones de estado (*statements*). Cada componente no interactivo de la aplicación web tiene enlazado un *statement*. Por lo tanto, dependiendo del estilo de navegación que sigue el estudiante se van generando *statements* que luego serán enviados a través de una API al repositorio LRS. Los *statements* se corresponden con la navegación del usuario (estudiante) por las pestañas de una página, se pueden registrar y luego analizar las pestañas por las que ha navegado.

Una vez que los datos hayan sido recopilados y almacenados en el LRS, se aplicarán sobre ellos técnicas de minería de procesos. El objetivo de aplicar estas técnicas es descubrir, analizar y proporcionar una representación visual del proceso educativo completo (Bogarín et al., 2018). Para el presente proyecto se plantea evaluar cuál de las siguientes técnicas de minería de procesos aporta mejores resultados al cumplimiento del objetivo planteado.

1. Técnicas de descubrimiento de procesos
2. Técnicas de análisis de conformidad
3. Técnicas de extensión del modelo de proceso

De igual manera se probará la efectividad de los siguientes algoritmos: algoritmos Alpha, algoritmos de minería heurísticos, algoritmos genéticos y algoritmos de minería difusos. Para realizar el análisis de los procesos se ha identificado en la literatura diversas aplicaciones de software gratuitas y licenciadas.

4. Conclusiones

El objetivo del trabajo es realizar un análisis cuantitativo de los posibles efectos que tienen la identificación de puntos críticos en las rutas de aprendizaje junto a la inclusión o no de criterios de accesibilidad en cursos en línea, sobre el aprendizaje de personas con discapacidad visual. El presente artículo está orientado específicamente a proponer un marco de referencia que permita llevar a cabo el trabajo de investigación mencionado.

La inclusión de estándares de accesibilidad en entornos virtuales de aprendizaje (WCAG 2.1, ATAG 2.0 y UAAG 2.0), no garantizan la consecución de un aprendizaje significativo de estudiantes con algún tipo de discapacidad. Por lo tanto, para complementar la funcionalidad que ofrecen dichos estándares se hace imprescindible realizar el análisis del comportamiento de los estudiantes cuando interactúan con plataformas virtuales. Dicho comportamiento se puede analizar a través del estudio de los puntos críticos identificados en las rutas de aprendizaje que siguen los estudiantes. Esto permitirá retroalimentar al docente en aspectos tales como: el diseño instruccional de los cursos. Al trabajar sobre estos dos elementos se espera obtener datos que permitan mejorar el aprendizaje de los estudiantes con algún tipo de discapacidad cuando se enfrenten a cursos en línea.

La aplicación de herramientas de AA permite recopilar información importante de la interacción que realiza el estudiante cuando trabaja en entornos virtuales de aprendizaje. Particularmente, la especificación xAPI ofrece una gama amplia de opciones para almacenar los rastros que deja un usuario al navegar por un sitio web, en este caso un entorno virtual de aprendizaje.

Referencias

Amo, D., Valls, A., Alier, M., Canaleta, X., García-Peñalvo, F., Fonseca, D., & Redondo, E. (2018). Using Web Analytics Tools to Improve the Quality of Educational Resources and the Learning Process of Students in a Gamified Situation. En *Proceedings of 12th annual International Technology, Education and Development Conference, INTED 2018* (pp. 5824-5829). https://doi.org/10.21125/inted.2018.1384

Bogarín, A., Cerezo, R., & Romero, C. (2018). A survey on educational process mining. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *8*(1), e1230. https://doi.org/10.1002/widm.1230

Bogarín, A., Romero, C., Cerezo, R., & Sánchez-Santillán, M. (2014). Clustering for improving educational process mining. En *Proceedins of the Fourth International Conference on Learning Analytics And Knowledge - LAK '14* (pp. 11-15). https://doi.org/10.1145/2567574.2567604

Chan, J., Chiu, R., Ng, G., & Kwong, T. (2015). How Clickstream Tracking Helps Design Mobile Learning Content. *International Journal of Humanities Social Sciences and Education (IJHSSE)*, *2*(7), 95-104.

Cooper, M., Ferguson, R., & Wolff, A. (2016). What can analytics contribute to accessibility in e-learning systems and to disabled students' learning? En *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK'016* (pp. 99-103). Recuperado de http://dl.acm.org/citation.cfm?doid=2883851.2883946

Iniesto, F., Mcandrew, P., Minocha, S., & Coughlan, T. (2008). Auditing the accessibility of Massive Open Online Courses (MOOCs). En *Proceedings of 14th Association for the Advancement of Assistive Technology in Europe Congress, AAATE 2017* (pp. 1-4).

Iniesto, F., McAndrew, P., Minocha, S., & Coughlan, T. (2014). Auditing the accessibility of Massive Open Online Courses (MOOCs). En *Proceedings of 14th annual of Congress Association for the Advancement of Assistive Technology in Europe, AAATE 2014* (pp. 169-172).

Iniesto, F., McAndrew, P., Minocha, S., & Coughlan, T. (2016). Accessibility of MOOCs: Understanding the Provider Perspective. *Journal of Interactive Media in Education*, *2016*(1), 1-10. https://doi.org/10.5334/jime.430

Iniesto, F., McAndrew, P., Minocha, S., & Coughlan, T. (2017). An investigation into the perspectives of providers and learners on MOOC accessibility. En *Proceedings of the 5th International Conference on Technological Ecosystems for Enhancing Multiculturality - TEEM 2017* (pp. 1-8). New York, New York, USA: ACM Press. https://doi.org/10.1145/3144826.3145442

Jadhav, M. A., Sawant, B. R., & Deshmukh, A. (2015). Single Page Application using AngularJS. *International Journal of Computer Science and Information Technologies*, *6*(3), 2876-2879.

Kirkpatrick, A., & Michael, C. (2018). WCAG 2.1 is a W3C Recommendation | W3C Blog. Recuperado de https://www.w3.org/blog/2018/06/WCAG21-REC/

Lim, K. C. (2015). Case Studies of xAPI Applications to E-Learning. En *Proceedings of The 12th International Conference on eLearning for Knowledge-Based Society, 2015* (pp. 11-12).

Luján-Mora, S. (2013). Web Accessibility Among the Countries of the European Union: a Comparative Study. *Actual Problems of Computer Science*, *1*(3), 18-27.

Ngubane-Mokiwa, S. (2016). Accessibility strategies for making MOOCs for people with visual impairments: A universal design for learning (UDL) perspective. En *Proceedings of Pan-Commonwealth Forum on Open Learning, 2016* (pp. 1-12).

Pechenizkiy, M., Trčka, N., Vasilyeva, E., Aalst, W. Van Der, & Bra, P. De. (2009). Process Mining Online Assessment Data. *Computer*, *10*(January), 279-288.

Sanchez-Gordon, S., & Lujan-Mora, S. (2013). Web accessibility of MOOCs for elderly students. En *2013 12th International Conference on Information Technology Based Higher Education and Training (ITHET)* (pp. 1-6). IEEE. https://doi.org/10.1109/ITHET.2013.6671024

Sanchez-Gordon, S., & Luján-Mora, S. (2014). Web Accessibility Requirements for Massive Open Online Courses Can MOOCs be really universal and open to anyone? En *Proceedings of 5th Congreso Internacional sobre Calidad y Accesibilidad de la Formación Virtual , CAFVIR 2014* (pp. 530-535).

Tóth, K., Rölke, H., Goldhammer, F., & Barkow, I. (2017). Educational Process Mining: New Possibilities for Understanding Students' Problem Solving Skills. En *The Nature of Problem Solving* (pp. 193-210).

Trcka, N., Pechenizkiy, M., & Van der Aalst, W. (2010). Process Mining from Educational Data. En *Data Mining and Knowledge Discovery Series* (pp. 123-142). https://doi.org/10.1201/b10274-11

Van der Aalst, W. (2012). Process Mining: Making Knowledge Discovery Process Centric. *SIGKDD Explorations Newsletter*, *13*(2), 45-49. Recuperado de http://dl.acm.org/citation.cfm?doid=2207243.2207251

World Wide Web Consortium. (2008). Web Content Accessibility Guidelines (WCAG) 2.0. Recuperado de https://www.w3.org/TR/WCAG20/

World Wide Web Consortium. (2015a). Authoring Tool Accessibility Guidelines (ATAG) 2.0. Recuperado de https://www.w3.org/TR/ATAG20/

World Wide Web Consortium. (2015b). User Agent Accessibility Guidelines. Recuperado de http://www.w3.org/TR/UAAG20/

World Wide Web Consortium. (2018a). W3C Accessibility Standards Overview | Web Accessibility Initiative (WAI) | W3C. Recuperado de https://www.w3.org/WAI/standards-guidelines/

World Wide Web Consortium. (2018b). Web Content Accessibility Guidelines (WCAG) 2.1. Recuperado de https://www.w3.org/TR/WCAG21/

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 26/10/2018
Aceitação/Acceptance: 06/12/2018

Modelo de educación virtual a distancia en el contexto universitario latinoamericano

Freire-Avilés Roger[1], Díaz-Nava Judith[2], Rodríguez-Suárez William[3], Vera-Lucio Néstor[4]

rfreire@uagraria.edu.ec, jdiaz@uagraria.edu.ec, william.rodriguez@urbe.edu.ve, nvera@uagraria.edu.ec

[1] Universidad Agraria del Ecuador, código postal 090112, Guayaquil, Ecuador.

[2] Universidad Agraria del Ecuador, código postal 090112, Guayaquil, Ecuador.

[3] Universidad Privada Dr. Rafael Belloso Chacín, código postal 050021, Medellín, Colombia.

[4] Universidad Agraria del Ecuador, código postal 090112, Guayaquil, Ecuador.

Pages: 1048–1061

Resumen: La presente investigación tuvo como objetivo general diseñar un modelo de educación virtual a distancia para el contexto universitario latinoamericano. El presente estudio es señalado como descriptivo, en virtud del análisis de la educación virtual a distancia en centros de educación virtual a distancia de universidades venezolanas. El diseño de la investigación desde la perspectiva cuantitativa se identifica como no experimental, transeccional y descriptiva. En tal sentido, el estudio hizo un análisis de los actores que intervienen en la educación virtual a distancia, además se conformó una población representada por coordinadores de los centros de educación virtual a distancia. Se aplicó un instrumento constituido por un cuestionario, destinado a los coordinadores de los centros de educación virtual a distancia. Los datos obtenidos fueron interpretados a través de análisis estadísticos, para proponer un modelo propio para la educación virtual a distancia en la universidad venezolana y latinoamericana.

Palabras-clave: Modelo; educación virtual a distancia; actores; centros de educación virtual.

Model of virtual distance education in the Latin American university context

Abstract: The main objective of the present research was to design a virtual distance education model for the Latin American university context. The present study is indicated as descriptive, by virtue of the analysis of distance virtual education in virtual distance education centers of Venezuelan universities. The research design from the quantitative perspective is identified as non-experimental, transectional and descriptive. In this sense, the study made an analysis of the actors involved in virtual distance education, in addition a population represented by coordinators of virtual distance education centers. An instrument consisting of a questionnaire was applied to the coordinators of virtual distance education centers. The data obtained

were interpreted through statistical analysis, to propose an own ontological model for virtual distance education in the Venezuelan and Latin American universities.

***Keywords**:* Model; distance virtual education; actors; virtual education centers

1. Introducción

Muchos de los instrumentos de formación, en gran parte hoy están disponibles gracias a las nuevas Tecnologías de Información y Comunicación (TIC), las cuales posibilitan adquirir conocimientos y competencias, incluso, fuera de las estructuras educativas tradicionales.

Este crecimiento no ha sido bajo la premisa de la revisión de las teorías y aplicación de las mismas en la partica de la educación virtual a distancia, por el contrario, las tecnologías de información y comunicación han marcado la pauta dejando a un lado elementos fundamentales de la educación virtual a distancia que deben ser considerados y ameritan atención por parte de los universidades en sus centros de educación, tanto las que existen como las que pretendan emprender el camino de la educación virtual a distancia.

En este sentido, si no se generan lineamientos que permitan unificar los criterios para la implantación de esta modalidad de estudio dentro de las universidades, se podrá caer inevitablemente en modelos no estructurados sin estándares, con recursos de formación inadecuados y sin las infraestructuras física adecuada ni la plataforma tecnológica más conveniente lo cual traerá como consecuencia inevitablemente, la no atención de la demanda de nuevo ingreso en el sistema de educación nacional, la deserción de los estudiantes en dichos programas de educación a distancia,

Partiendo de lo expuesto anteriormente, se presenta esta investigación, la cual tiene como propósito diseñar un modelo de educación virtual a distancia en las universidades venezolanas y latinoamericanas para servir como referencia para revisar los modelos actuales de la educación virtual dentro de las instituciones de educación superior.

Por último, cabe destacar que el presente artículo sirve como base para futuras investigaciones relacionadas con el área objeto de estudio, además que el análisis desarrollado puede ser utilizado no solo por los centros objetos de estudio, sino también en organizaciones empresariales pertenecientes a otros sectores de actividad económica, adaptándola a la naturaleza y necesidades de las mismas.

2. Modelos

Existen diferentes axiomas que se le pueden atribuir al término “Modelo” dependiendo del contexto donde se sitúe. A los efectos Mujica y Rincón (2011) afirman que término Modelo constituye un resultado complejo, producto de ideas, experiencias, prácticas, sentires e instituciones suscitados en un contexto determinado por la intención de conocer, cuyo dinamismo, producto de la realidad compleja y cambiante, determinan su validez en el tiempo.

Un modelo es un complejo de ideas, preceptos, precogniciones, conceptos y afirmaciones mediante las cuales se indaga y se percibe, se aprehende y se comprende. Su principal

característica es el dinamismo: las ideas que lo soportan, varían con el tiempo; sus principios, cambian; las situaciones sobre las que se fundamentan, se transforman.

Siguiendo con Ferrater (2001), desde el punto de vista epistemológico, la noción de modelo ha sido empleada en otros sentidos. Se habla de modelo como un modo de explicación de la realidad, especialmente de la realidad física; por ejemplo, se habla de modelos mecánicos, que se equivalen con el mecanicismo, y que han seguido autores como Newton y Galileo. Del mismo modo se utiliza como forma de representación de alguna realidad o serie de realidades, de algún proceso o serie de procesos; por ejemplo, un dibujo, un plano o una maqueta.

También ha sido empleado como el equivalente a una teoría; o como indicación de diferencias entre modelo y teoría, o para afirmar que una teoría puede tener diversos modelos o modelarse de varios modos. Habría que agregar un modo común de entender modelo, que es cuando se lo toma como un sistema que sirve para entender otro sistema, o cuando se toma como tal un sistema del cual se trata de presentar una teoría. En este último caso el modelo es, entonces, la realidad que la teoría trata de explicar. Puede haber, de este modo, varias teorías para explicar unas más satisfactoriamente que otras-un modelo; puede haber, además, una teoría para el cual se busque un modelo, así como una teoría que, habiendo resultado satisfactoria en la explicación de un modelo, sea capaz de aplicarse a otros modelos.

En este sentido, la profundización de los aspectos epistemológicos podrían aclarar un poco más la tarea de definir en propiedad el concepto de modelo, pues desde este punto de vista Barrera (2007), considera que el estudio de los modelos tienen la particularidad de permitir a cada quien, de acuerdo a sus intereses, a sus valores y a sus propias reflexiones, conocer un evento desde las manifestaciones que le interesan; es decir, existe una vinculación directa entre la ontología la cosa en sí y la epistemología el decir de la cosa (además de constituirse en) una herramienta importante para la generación de nuevas formas de conocer, como también distintas maneras de entender un evento que, a la manera aristotélica, se expresa de múltiples maneras, y al estilo de Morín, se evidencia como complejidad.

2.1. Modelo como representación

En este sentido, un primer intento de establecer el concepto fue el que lo restringió a modelo mecánico específicamente en el campo de la física, y obedeció a la necesidad de comprender y actuar sobre la naturaleza, la que adquiere existencia en virtud de que para ella se ha propuesto un modelo (Gallego, 2004). Para Scheler (1926), citado por Gallego (2004), un modelo lógico mecánico para dar cuenta de los fenómenos de la naturaleza, ha de pensar y percibir esos fenómenos en términos de explicaciones mecánico-formales. Un modelo de estas características ha de ser siempre factible, a partir de la elaboración previa de un pensamiento relacionante que posibilita el ordenamiento de dichos fenómenos.

Para Hanson (1958), citado por Gallego (2004), un modelo es una estructura conceptual que sugiere un marco de ideas para un conjunto de descripciones que de otra manera no podrían ser sistematizadas. Cumple esta función en virtud de que une de manera inferencial, las proposiciones que afirman algo sobre los fenómenos que en él se

integran. De esta manera, su estructura es diferente de la que se supone existe en el conjunto de fenómenos de la naturaleza. El modelo concebido en esta forma, impulsa la inteligibilidad y ayuda a la comprensión de los fenómenos, ya que proporciona los canales de interconexión entre hechos que sin la existencia de los lazos inferenciales, podrían permanecer aislados e independientes unos de otros.

En términos de Kuhn, los modelos son una serie de realizaciones que sirven durante una época de ciencia normal para definir problemas y métodos legítimos en un campo específico de investigación. Estos son siempre incompletos, ya que no abarcan todos los problemas que se espera han de ser resueltos (Gallego, 2004).

Un modelo es sustituido por otro, en la medida del desenvolvimiento progresivo del programa objeto de contrastación. Por último, Tomasi (1999), citado por Gallego (2004), explica que los modelos son, por definición, incompletos con respecto a los referentes empíricos para los cuales se construyen.

2.2. Modelo como analogía

Como se desprende, siempre se ha hecho uso de la categoría modelo. Intentando hacer una síntesis, Del Re (2000), define modelo como simplificaciones o representaciones idealizadas de los sistemas que se supone existen en la naturaleza; son las herramientas del pensamiento científico y pueden ser de dos tipos: matemáticos y físicos, siendo los modelos físicos esenciales, no sólo para describir, sino también para conocer aquello que no es directamente accesible a los sentidos.

El problema se presenta, entonces, en cómo y hasta dónde un modelo podría caracterizar e identificar fenómenos que de otra manera no serían observables. Del Re (2000) sugiere que, a fin de ser verdad, una declaración no necesita coincidir con la realidad en todos sus aspectos, relaciones y detalles, pero tiene que concordar con ella en sus aspectos esenciales; es decir, la complejidad es otra característica de la realidad de acuerdo con las entidades que deben ser clasificadas por distintas disciplinas. La inferencia principal que se extrae es que, dado que el hombre no tiene percepción directa de una gran mayoría de los objetos que la ciencia ha detectado en el universo físico, su conocimiento se basa en analogías.

En el mismo sentido, los modelos son analogías de los sistemas reales, analogías que no tienen por qué ser necesariamente un simple conjunto de afirmaciones, y cuya validación científica, teórica y epistémica vendría del análisis del estatuto lógico de cada modelo y de sus relaciones con las evidencias observables que constituyen su soporte (Caldin, 2002; Gallego, 2004).

En otras palabras, los modelos son análogos del sistema real; dos cosas se dicen que son análogas, en la terminología de la lógica moderna, si tienen algunas, pero no todas, las características en común. No se sabe si están de acuerdo en otros aspectos; sólo podemos ajustar el modelo de conformidad con nuestras pruebas, y esa evidencia es siempre incompleta. Los modelos, entonces, no deben ser tomados como descripciones exactas de la realidad, por un lado, ni como puras ficciones por el otro; sería mejor considerarlos como analogías, en las que sólo algunas, no todas, las características son las mismas o similares a las de la realidad (Caldin, 2002).

2.3. Los modelos en las ciencias sociales: los axiomas

En el campo específico de las ciencias sociales, según Armatte (2005), la noción de modelo ha estado determinada por la imposición y traslado mecánico del paradigma positivista físico-matemático y su asiento definitivo con en el programa de unificación de las ciencias sobre la base anti-metafísica del Círculo de Viena. Así, el método lógico asimila el quehacer científico a una construcción lingüística en el que la base se hace de enunciados en correspondencia con los hechos registrados, soporte sobre el que son construidas proposiciones coherentes.

Es decir, las ciencias sociales asumen, en consecuencia, el diseño de modelos en base a la formalización matemática, sobre todo en economía, como una forma de conciliar el dualismo entre el enfoque matemático hipotético-deductivo y el enfoque inductivo de la estadística económica (Armatte, 2005).

Para Sierra Bravo (2002), en la filosofía de las ciencias, el término modelo se utiliza en diversos sentidos, frecuentemente nada claros. Este hecho guarda relación con la naturaleza conceptual y derivada de la noción del modelo en la ciencia. Para el autor, los modelos son en todo caso construcciones racionales, y constructos, que fundamentalmente se forman apriorísticamente a partir de otros conceptos y no necesariamente de la observación de la realidad. Además, si la idea de modelo en el lenguaje corriente, es abstracta, por cuanto implica la operación conceptual de comparar el modelo físico con otra cosa, en la ciencia se trata normalmente de establecer dicha comparación no entre cosas sino entre construcciones conceptuales. Por otra parte, los modelos se presentan como algo intermedio, híbrido, entre las hipótesis y las teorías que hace difícil a veces su distinción de ellas, al no ofrecer frecuentemente una razón de ser tan clara y distinta como aquellas.

Marín (2004), sin embargo, afirma que los modelos hacen referencia sobre todo al mundo empírico (lo que es). En el ámbito del proceso de investigación, las hipótesis o predicciones contrastables de relaciones entre fenómenos adquieren un nuevo grado de generalidad y capacidad armonizadora y deductiva en la medida en que conecten con teorías científicas, y sus afirmaciones son plenamente significativas en el ámbito de un modelo. Se establece, de esta manera, un grado creciente de abstracción que va del fenómeno concreto a la hipótesis todavía contrastable, a la teoría y en último lugar al modelo. Y en un proceso deductivo, en sentido contrario al indicado, un modelo se puede concretar en varias teorías, que se especifican en diferentes hipótesis.

3. Variable

3.1. Definición nominal

Educación virtual a distancia.

3.2. Definición conceptual

Tomando en cuenta la opinión de Silvio (2003), la educación virtual a distancia es aquella en que los actores interactúan a través de representaciones numéricas de los elementos del proceso de enseñanza y aprendizaje, pero se encuentran en lugares y momentos de tiempo distintos.

3.3. Definición operacional

Es aquella educación, en que los actores (estudiantes, profesores y asistentes) interactúan a través de recursos de formación del proceso de enseñanza y aprendizaje mediante una plataforma tecnológica, permitiendo que los mismos se encuentren en lugares y momentos de tiempo distintos, ajustados al marco legal vigente del contexto donde se desarrolla.

Esta variable fue medida mediante un instrumento de recolección de datos elaborado por el investigador de acuerdo al siguiente cuadro de operacionalización. (Ver tabla 1).

Objetivos Específicos	Variable	Dimensiones	Indicadores
Describir el marco legal de la educación virtual a distancia en la universidad venezolana	Educación virtual a distancia	Marco Legal	• CRBV • Ley de Educación Superior • CNU – OPSU • Normativa Nacional para la EAD
Identificar los actores inmersos de la educación virtual a distancia en la universidad venezolana		Actores	• Estudiantes • Profesores • Asistentes
Revisar los recursos de formación que conforman la educación virtual a distancia en la universidad venezolana		Recursos de Formación	• Material Instruccional • Calidad • Comunicación Síncrona – Asíncrona • Evaluación de los aprendices
Determinar la infraestructura de la educación virtual a distancia en la universidad venezolana		Infraestructura Tecnológica	• Software • Hardware
Elaborar un modelo de educación virtual a distancia en la universidad venezolana		No se operacionaliza, se desarrolla con base a los resultados obtenidos por el investigador	

Tabla 1 – Operacionalización de la variable

4. Marco Metodológico

El paradigma científico de la presente investigación parafraseando a Barrientos (2008) es el de la complementariedad paradigmática, es decir, una mezcla entre el paradigma positivista y el post-positivista, donde la preocupación del investigador se centra no solo en la decisión de utilizar números o palabras en el desarrollo del estudio, sino en la forma de analizar, interpretar, describir y explicar los datos e información obtenida.

Dada la naturaleza de la variable educación virtual a distancia, esta investigación fue de tipo descriptiva, por cuanto comprende elementos metodológicos de carácter descriptivo, en función de caracterizar las actividades de detectar características,

describir funciones y definir fuente de información de la educación virtual a distancia en las Universidades. Según Hernández y otros (2006), el tipo de investigación descriptiva es aquella que mide o evalúa diversos aspectos, dimensiones o componentes del fenómeno o fenómenos a investigar.

El diseño de la investigación proporciona un modelo de verificación que permite contrastar hechos con teoría. En este sentido, el presente estudio utilizo un diseño de investigación no experimental, transeccional, descriptivo, debido a que no se manipuló deliberadamente la variable objeto de estudio, a saber, educación virtual a distancia, se recolectó en un momento o tiempo único periodo comprendido entre julio de 2016 a enero de 2018.

Para la medición/descripción de la educación virtual a distancia, se tomó la información tal y como se presentó en las universidades que poseen educación superior a distancia, correspondiéndose esto con la condición indispensable de una investigación de campo, de acuerdo a los criterios expuestos por Sabino (1998), según refiere a métodos a emplear cuando los datos de interés se recogen en forma directa de la realidad, mediante el trabajo concreto del investigador y su equipo.

4.1. Población y muestra

De acuerdo a Tamayo y Tamayo (2004), la población está determinada por sus características definitorias, por tanto, el conjunto de elementos que poseen esta característica se denomina población o universo. Población es la totalidad del fenómeno a estudiar, en donde las unidades de población poseen una característica común, la cual se estudia y da origen a la investigación.

En esta investigación la población estuvo conformada por los coordinadores de los centros de educación virtual a distancia en las universidades venezolanas. (Ver tabla 2).

Los criterios de selección de los informantes clave en dichas universidades fueron los siguientes:

- Voluntad de participar en el estudio.
- Ocupar un cargo directivo dentro del centro de educación virtual a distancia de la universidad estudiada.
- Ser coordinador de área.

Como el número de personas o informantes clave de los centros de educación virtual a distancia es reducido, no se hizo muestreo y en consecuencia se trabajó con la población completa.

4.2. Técnicas e instrumentos de recolección de información

La técnica para recolección de información son los procedimientos utilizados por el investigador para recabar información. Las técnicas pueden ser documentales y de campo. En el caso del presente estudio se utilizó esta última, la cual consiste en el procedimiento por medio de cual se obtiene y registra la información, directamente en el lugar en el que ocurren los fenómenos, hechos o situaciones objeto de investigación (De Pelekais y otros, 2007).

Para esta técnica se usó como instrumento el cuestionario con las siguientes características: de respuestas cerradas, precodificado y administrado a través de la modalidad de entrevista personal (De Pelekais y otros, 2007). En este caso el respondiente marcó una o más respuestas en función de lo que consideró más apropiado, utilizando la escala tipo Likert.

Universidad	Cantidad de personas	Cargos
Universidad Central de Venezuela - UCV	1	Coordinador
Universidad Católica Occidental Lizandro Alvarado – UCLA	1	Coordinador
Universidad Nacional Abierta - UNA	1	Coordinador
Universidad Privada Dr. Rafael Belloso Chacín - URBE	1	Coordinador
Universidad del Zulia – LUZ	1	Coordinador
Universidad Simón Bolívar - USB	1	Coordinador
Universidad Pedagógica Experimental Libertador - UPEL	1	Coordinador
Universidad Católica Cecilio Acosta - UNICA	1	Coordinador
Universidad Experimental del Táchira - UNET	1	Coordinador
Universidad del Valle de Momboy - UVM	1	Coordinador
Universidad Fermín Toro (UFT)	1	Coordinador
Universidad de Los Andes – ULA	1	Coordinador
Universidad Nueva Esparta - UNE	1	Coordinador
Universidad Nacional Experimental Politécnica de la Fuerza Armada Nacional - UNEFA	1	Coordinador
Universidad José Antonio Páez - UJAP	1	Coordinador
Universidad Marítima del Caribe - UMC	1	Coordinador
Universidad Bicentenaria de Aragua - UBA	1	Coordinador

Tabla 2 – Centros de Educación a distancia por universidades

5. Modelo de educación virtual a distancia en la universidad venezolana y latinoamericana

Luego de realizado el análisis y discusión de los resultados con la exhaustiva revisión bibliográfica llevada a cabo en relación a los diferentes modelos planteados por los distintos autores, se procederá a formular bajo un enfoque sistémico un Modelo Conceptual de Educación Virtual a Distancia el cual pueda ser tomado como punto de referencia para la aplicación en la Universidad Venezolana y Latinoamericana.

5.1. Presentación de la propuesta

La construcción de un modelo de Educación Virtual a Distancia implica la elaboración conceptual de una representación de la realidad observada sobre un conjunto de fenómenos que suceden en un entorno circundante. Lo que implica que la experiencia del investigador sobre la observación de tales fenómenos y sobre el objeto de estudio, pueden cargar al proceso de construcción del citado modelo con cierta subjetividad inherente a la óptica o visión del investigador.

En este sentido, Hanson (1958) afirma que un modelo es una estructura conceptual que sugiere un marco de ideas para un conjunto de descripciones que de otra manera no podrían ser sistematizadas. El modelo cumple esta función en virtud de que une de manera inferencial, las proposiciones que afirman algo sobre los fenómenos que en él se integran. De esta manera, su estructura es diferente de la que se supone existe en el conjunto de fenómenos de la naturaleza. El modelo concebido en esta forma, impulsa la inteligibilidad y ayuda a la comprensión de los fenómenos, ya que proporciona los canales de interconexión entre hechos que sin la existencia de los lazos inferenciales, podrían permanecer aislados e independientes unos de otros.

En concordancia con lo antes expuesto, es importante resaltar que un modelo siempre será una representación imaginaria de una realidad observada, y para la presente investigación esta realidad se encuentra limitada a los fenómenos estudiados en los centros de educación virtual a distancia de las universidades objetos de estudio donde se realizan las actividades educativas en entornos virtuales.

En tal sentido, Kunh (1972) admite igualmente, que el concepto de modelo es aquella que lo define como una serie de realizaciones que sirven durante una época de ciencia normal para definir problemas y métodos legítimos en un campo específico de investigación. Los modelos son siempre incompletos, ya que no abarcan todos problemas que se espera han de ser resueltos, de igual forma manifiesta que los paradigmas compartidos proporcionan modelos de los que surgen tradiciones coherentes de investigación científica. Los cambios paradigmáticos podrían ser asumidos como cambios de esos modelos científicos.

5.2. Elementos que conforman el modelo de educación virtual a distancia

Para la construcción del modelo de Educación Virtual a Distancia, los elementos a ser considerados se obtuvieron luego de ser aplicado los instrumentos de recolección de datos con su posterior análisis y discusión de los mismos. En tal sentido, los elementos que serán tomados en cuenta para la conformación del modelo de Educación Virtual a Distancia son los siguientes:

- Marco Legal
- Actores
- Recursos de Formación
- Infraestructura Tecnológica

5.3. El proceso de modelado

Según Ballesteros y Landazábal (2010), debe considerar también que la percepción que tiene el modelador de lo real, la realidad por él percibida, es posible que la adquiera ya sea por su propia capacidad de observación o haciendo uso de la capacidad de otros, es decir, usando las teorías que también son una interpretación consensuada y ampliamente aceptada, de lo que aquí se hadado en llamar como lo real. Luego que se ha percibido lo real esta percepción se aloja en la mente del modelador en la forma de un modelo mental que, según Senge (1992), “son supuestos hondamente arraigados, generalizaciones e imágenes que influyen sobre nuestro modo de comprender el mundo y actuar”. Tales modelos mentales se convierten, entonces, en las representaciones que cada modelador tiene de su propia percepción de lo real.

En este sentido, el primer punto de partida para el abordaje del modelo consiste en la representación del modelo mental, tal y como se muestra en la figura 1.

Figura 1 – Modelo Mental de Educación Virtual a Distancia.

Como se puede observar en la anterior figura, el modelo mental de la educación virtual a distancia dentro de los centros de educación universitaria está conformado por cuatro (4) elementos fundamentales: Marco legal, Recursos de Formación, Actores e Infraestructura Tecnológica.

Instrumento	Aporte
CRBV	Los fines, estructura y funcionamiento de la educación en Venezuela están definidos en la Constitución de Nacional (1999) y por un conjunto de leyes y reglamentos.
Ley de Educación Superior	Esta ley establece las directrices y bases de la educación como proceso integral; determina la orientación, planificación y organización del sistema educativo y norma el funcionamiento de los servicios que tengan relación con este.
OPSU Consejo Nacional de Universidades	De acuerdo a la ley de universidades, la Oficina de Planificación del Sector Universitario es una oficina técnica auxiliar del Consejo Nacional de Universidades, que se encarga de instrumentar las políticas y estrategias para la educación superior señaladas en los planes de la nación
Normativa Nacional para la educación a distancia	Esta normativa establece los lineamientos que orientan la creación, sistematización desarrollo, implantación y evaluación de la Educación Superior a Distancia en la República Bolivariana de Venezuela.

Tabla 3 – Marco Legal de la Educación Virtual a Distancia de la Universidad

En este sentido, el primer elemento denominado Marco legal está representado por la Constitución de la República Bolivariana de Venezuela, la Ley de Educación Superior, el Consejo Nacional de Universidades y la Normativa Nacional para la Educación a Distancia, la revisión de estos documentos permiten dar el contexto legal de la educación virtual a distancia dentro de las universidades venezolanas.

En este orden de ideas, el segundo elemento determina los actores que intervienen en el proceso de la educación virtual a distancia, en este sentido identificamos tres tipos de actores:

Profesores: Zambrano y otros (2010), el docente-tutor para desarrollar eficazmente el proceso enseñanza- aprendizaje debe manejar adecuadamente las tecnologías Web 2.0, contar con estrategias de enseñanzas apropiadas para la educación virtual, ser guía, motivador, facilitador y supervisor de las estrategias que utilizan los discentes para aprender, estar atento a las dificultades de los estudiantes y capacidad para planear las actividades del curso, así como para evaluarlas.

Asistentes: Para Díaz (2011), los asistentes técnicos en línea, son aquellos encargados de prestar el soporte técnico a los estudiantes de los cursos y a los tutores. Este actor representa un elemento fundamental para avanzar en la curva de aprendizaje que representa el conocimiento y la adopción de las nuevas tecnologías de comunicación e información presentes en la educación virtual.

Estudiantes: Constituyen el centro del proceso educativo y son la razón de ser de las actividades de enseñanza, son consideramos personas y ciudadanos que perciben y adquieren conocimientos de maneras distintas, presentando preferencias hacia determinadas estrategias cognitivas que les ayudan a dar significado a la nueva información

En concordancia con lo expuesto anteriormente, los actores: profesores, asistentes y estudiantes deben apropiarse de la tecnología presente en el sistema de educación virtual a distancia, a través de, procesos de transferencia tecnológica (capacitación, cursos propedéuticos, manuales, tutoriales) que permitan la gobernabilidad de la tecnología por parte de los mismos, en tal sentido, un centro de educación virtual, presente en cualquier universidad, que capacite a sus actores en el sistema de gestión de aprendizaje que utilicen tendrá más oportunidad de desarrollar las estrategias instruccionales presentes en los recursos de formación que se ofrecen, a través, del sistema de gestión de aprendizaje para el cumplimento de los objetivos académicos dispuestos en el currículo del pensum de estudio o programa que se quiera dictar.

Con respecto al tercer elemento del modelo, los recursos de formación, el proyecto nacional de educación universitaria a distancia en su Capítulo IV, Articulo 37 indica que los recursos de formación, como materiales y medios utilizados en la educación a distancia, estarán definidos según el modelo educativo institucional, los lineamientos de diseño para su entrega a distancia y contextualizados a las necesidades de formación y a las características culturales, lingüísticas, sociales y productivas de ámbitos regionales y locales en los cuales se desarrolle la educación a distancia.

En el mismo sentido, sobre la comunicación se estableció que pueden existir actividades de naturaleza síncrona y asíncrona, sin embargo, se debe fomentar las actividades síncronas, como por ejemplo las videoconferencias, que permita un acercamiento a la

individualización de los programas mediante la interacción profesor-estudiante en un ambiente colaborativo con calidad humana, que rompa con la soledad del estudiante en la educación virtual a distancia.

El cuarto elemento de modelo se refiere a la infraestructura tecnológica, los centros de educación virtual a distancia los cuales deben contar con la infraestructura tecnológica y planta física adecuada y el personal técnico idóneo que garanticen el desarrollo y funcionamiento de los medios tecnológicos, así como el respaldo y seguridad de la data académica que se produzca, en este sentido, el proyecto nacional para la normativa universitaria indica que el domicilio obligatorio y resguardo será dentro del territorio venezolano.

Adicionalmente, las universidades deben disponer de la tecnología necesaria, en todas sus localidades, que aseguren la prestación de servicios que satisfagan las necesidades de formación de los actores involucrados. En cuanto al software utilizado en los centros, deben cumplir con los lineamientos del uso de software libre estipulados por la nación, para lo cual los centros deben generar normativas en este sentido, de la misma manera, los centros deben garantizar a las y los actores involucrados en el hecho educativo, las condiciones de conectividad, navegabilidad y manejo de dispositivos físicos y técnicos para el acceso adecuado a las tecnologías de información y comunicación.

En relación al hardware, se debe implementar arquitecturas en la Nube bajo cualquiera de sus modalidades: Públicas, Privadas o Hibridas que permitan a las universidades obtener con pocos esfuerzos beneficios de implementación rápida, baja en riesgos, actualizaciones automáticas, portabilidad de información, seguridad de los datos entre otros beneficios.

6. Conclusiones

En el marco legal nacional para la Educación a Distancia, descansan los elementos con rango de ley que deben considerar las universidades al momento de iniciar el camino de la educación a distancia. El desconocimiento del marco legal por parte de las universidades y sus centros de educación virtual a distancia, implica un riesgo latente en la continuidad del proceso educativo, lo que puede acarrear sanciones contra las universidades por parte del Gobierno y los organismos regulatorios de Educación.

Del mismo modo, los recursos de formación deben cumplir con los elementos instruccionales para lograr el proceso enseñanza-aprendizaje, los cuales deben poseer una interfaz que se adapte a cualquier tipo de dispositivo móvil adicional a la de los navegadores para dispositivos de escritorio. Deben garantizar la solidez de los procesos de gestión de enseñanza – aprendizaje, ofreciendo a los participantes ventajas creativas para el adecuado desarrollo de los conocimientos adquiridos; y, brindar a los profesores ventajas organizativas para el adecuado desarrollo de los procesos administrativos y académicos en la educación virtual a distancia.

Es necesario ofrecer comunicación en tiempo real entre los actores mediante el uso de infraestructuras en la nube que incluyan recursos de formación síncronos con recursos en tiempo real; y, asíncronos que permitan la comunicación desarrollada en tiempo

diferido, con la finalidad de que los actores puedan romper las barreras de espacio y de tiempo necesarias para el proceso de enseñanza-aprendizaje en entornos virtuales. Adicionalmente, dichas plataformas deben ofrecer adaptabilidad para enriquecer la experiencia de usuario, navegación intuitiva, facilidad en la personalización de contenidos y trazabilidad del proceso educativo desde la plataforma de manera grupal e individual.

Al culminar esta investigación, se puede resaltar que es necesario que los actores se apropien de la tecnología mediante los procesos de transferencia tecnológica que les permiten gobernar la tecnología; y, no que las tecnologías gobiernen a los actores. Igualmente, deben poseer procesos de evaluación del aprendizaje mediante la utilización de talleres, foros, tareas, resolución de cuestionarios, elaboración de mapas conceptuales, cuadros sinópticos, mapas mentales, entre otros. Además, se debe evaluar la elaboración de un proyecto (producto) para que los estudiantes demuestren los conocimientos adquiridos.

La plataforma tecnológica de educación virtual a distancia debe ofrecer seguridad de datos, mecanismos de recuperación de datos, contar con las condiciones técnicas y de infraestructura necesarias para el almacenamiento de la información. Con la elaboración de este modelo de educación virtual a distancia en la universidad venezolana, se puede concluir que es necesaria la aplicación del modelo en todas las universidades venezolanas y latinoamericanas; permitiendo validar la práctica actual en las universidades que poseen esta modalidad; y, para las universidades que aún no la usan, iniciar el camino para la implementación de la educación virtual a distancia, bajo un modelo con fundamentos teóricos que les permita avanzar con paso firme hacia la búsqueda de la excelencia académica.

Referencias

Barrera M., Marcos F. (2007). Modelos Epistémicos en Educación eInvestigación. Sypal. Quirón Ediciones. (Cuarta Edición). Venezuela.

Caldin, Edward. (2002). The Structure of Chemistry in Relation to the Philosophy of Science.HYLE.International Journal for Philosophy of Chemistry.Vol.8.N 2.Universidad de Karlsruhe.Alemania.(Pp 103-121). Documento consultado el 15 de septiembre de 2017 en: http://hyle.org/journal/issues/8-2/caldin.html.

Consejo nacional de universidades, oficina de planificación del sector universitario (2009). Proyecto nacional de educación superior a distancia. Caracas, República Bolivariana de Venezuela.

De Pelekais y otros. (2007). El ABC de la Investigación. Una aproximación teórica práctica. Editorial Panapo de Venezuela, C.A. Caracas.

Del Re, Giuseppe. (2000). Models and Analogies in Science.HYLE.International Journal for Philosophy of Chemistry.Vol.6, No. 1.Universidad de Karlsruhe. Alemania. (Pp. 5-15). Documento consultado el 07 de abril de 2017 en: http://hyle.org/journal/issues/6/delre.htm.

Diaz V., Jesús E. (2011). Modelos pedagógicos en educación a distancia. Revista REDHECS. Documento consultado el 16 de febrero de 2018 en: http://www.publicaciones.urbe.edu/index.php/REDHECS/article/viewArticle/971/3002.

Ferrater M., José. (2001). Diccionario de Filosofía. Tomo III. K-P. Nueva Edición Actualizada. Ariel Referencia. Editorial Ariel. 2da reimpresión. España.

Gallego B., Rómulo. (2004). Un Concepto Epistemológico de Modelo para la Didáctica de las Ciencias Experimentales. Revista Electrónica de Enseñanza de las Ciencias. Vol. 3 N 3. Universidad de Vigo. España. (Pp. 301-319). Documento consultado el 20 de octubre de 2017 en: http://reec.uvigo.es/volumenes/volumen3/REEC_3_3_4.pdf.

Hanson, R. (1958). Observation and Explanation: A guide to Philosophyof Science. Patterns of Discovery.And Inquiry into the Conceptual

Hernández, R., Fernández C., Baptista, P. (2006). Metodología de la Investigación. McGraw-Hill Interamericana cuarta edición. México.

Kuhn, T.S. (1972). La estructura de las revoluciones científicas. México:Fondo de Cultura Económica.

Ley de Universidades, Tomado de la Gaceta Oficial N° 1.429 Extraordinario del 8 de septiembre de 1970, N° 28.262 del 17 de febrero de 1967

Marín, Jesús. (2004). Notas de Sociología Crítica. Editorial SERBILUZ. Venezuela.

Mujica N., Rincón S. (2011). Consideraciones teórico - epistémicas acerca del concepto de modelo. Revista TELOS Vol. 13 NO.1- Enero – Abril 2011. Maracaibo, República Bolivariana de Venezuela. Documento consultado el 01 de abril de 2018 en: http://publicaciones.urbe.edu/index.php/telos/article/view/1904/3244.

Sabino, Carlos (1998). El proceso de la Investigación. Editorial Panexpo. Venezuela.

Senge, Peter (1992). La quinta disciplina, Editorial Granica. Buenos Aires.

Sierra Bravo, Restituto. (2002). Ciencias Sociales: Epistemología, Lógica y Matemática. Editorial Paraninfo. España.

Silvio, José (2003) Iesalc-Unesco / Anuies (2004). La educación superior virtual en América Latina y el Caribe. México: IESALC-Unesco/ ANUIES. Documento consultado el 18 de mayo de 2018 en: www.iesalc.unesco.org.ve/dmdocuments/.../libros/EducVirtual_ALC.pdf.

Zambrano William., Medina Victor., Garcia Victor.(2010). Nuevo rol del profesor y del estudiante en la educación virtual. Dialéctica Revista de Investigación. Documento consultado el 10 de abril de 2018 en: dialnet.unirioja.es/descarga/articulo/3340102.pdf.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 13/09/2018
Aceitação/Acceptance: 07/11/2018

Metodología SCRUM y desarrollo de Repositorio Digital

Margarita Ramírez Ramírez[1], María del Consuelo Salgado Soto[1], Hilda Beatriz Ramírez Moreno[1], Esperanza Manrique Rojas[1], Nora del Carmen Osuna Millán[1], Ricardo Fernando Rosales Cisneros[1]

maguiram@uabc.edu.mx, csalgado@uabc.edu.mx, emanrique@uabc.edu.mx, ramirezmb@uabc.edu.mx, nora.osuna@uabc.edu.mx, ricardorosales@uabc.edu.mx

[1] Universidad Autónoma de Baja California, 22000, Tijuana, B.C. México.

Pages: 1062–1072

Resumen: El presente trabajo propone un modelo de gestión de información, desarrollado como un Repositorio digital, que brinda la posibilidad de mantener y accesar información actualizada sobre la productividad académica y científica de los docentes e investigadores, de la Facultad de Contaduría y Administración, de la Universidad Autónoma de Baja California. En primer lugar se muestran los antecedentes de repositorios digitales, así como la descripción detallada de las funcionalidades del mismo, se presenta la metodología ágil de desarrollo SCRUM, la cual fue implementada en la administración y el desarrollo del repositorio.

Palabras-clave: Repositorio digital; TICS; SCRUM.

SCRUM methodology and development of Digital Repository

Abstract: The present work proposes an information management model, developed as a digital repository, which provides the possibility of maintaining and accessing updated information on the academic and scientific productivity of teachers and researchers, of the Accounting and Administration Faculty, of the Autonomous University of Baja California. In the first place the background of digital repositories is shown, as well as the detailed description of the functionalities of the same, the agile SCRUM development methodology is presented, which was implemented in the administration and development of the repository.

Keywords*: Digital Repository; TICS; SCRUM.*

1. Introducción

La investigación científica y la productividad académica son de vital importancia para el desarrollo de los diferentes sectores de la sociedad, la mayor importancia de estas acciones radica en la difusión de los resultados obtenidos en la investigación y el trabajo académico, esta es una tarea imprescindible en el contexto de beneficios y de la utilidad obtenida en el trabajo desarrollado por investigadores.

En la actualidad se avanza en una nueva era, las tecnologías de información y comunicación (TIC), son el principal agente en la difusión de la ciencia y de la cultura, así como de la información generada en los medios académicos. En dirección a este avance es imprescindible que las instituciones educativas y científicas articulen estructuras que permitan albergar de manera clasificada la documentación generada por sus investigadores, docentes y estudiantes, de modo que los conocimientos generados puedan tener mayor visibilidad y accesibilidad a través de medios tecnológicos.

Alberin (2015), menciona "Nunca en la historia de la Humanidad los habitantes del planeta han tenido tan amplias posibilidades para comunicarse, para compartir y dar sentido en común a sus experiencias y conocimientos, a sus percepciones, a las imágenes de sí mismos y lo que les rodea", en los últimos años, la cantidad de información generada corresponde a la misma dimensión de la información que se ha generado en la historia de la humanidad.

Un instrumento muy relevante en las instituciones educativas son los repositorios Institucionales, los cuales han sido abordados como un tema relevante para las universidades e instituciones académicas. Desde el ámbito internacional a partir de diversos enfoques y contextos, comenzando con ArXiv1 (R. Casate, 2009).

Una Institución que genera y mantiene un repositorio digital realiza la difusión del conocimiento y permite la creación de redes de colaboración y el trabajo colaborativo lo cual trae como resultado la generación de conocimiento y la ciencia aplicada.

Existe un ranking en el que se registran a las Universidades que se apoyan de repositorios digitales Institucionales, en este ranking Web de Repositorios (CSI, 2012), es posible identificar un número reducido de universidades de México, la creación y mantenimiento de un repositorio se ha convertido en un reto para las Universidades y Centros de Investigación. La Facultad de Contaduría y Administración no es la excepción y trabaja en la creación, mantenimiento y difusión de un repositorio que permita mostrar a la comunidad académica y científica los resultados de los trabajos académicos y de investigación de sus docentes e Investigadores.

En este trabajo se presenta un modelo de gestión de información, que concentra los documentos, evidencia de la productividad académica y científica de los docentes de la Facultad de Contaduría y Administración, de la UABC. En primer lugar se muestran los antecedentes de repositorios digitales, así como la descripción detallada de las funcionalidades del mismo, se presenta la metodología ágil de desarrollo SCRUM, la cual fue implementada en la administración y el desarrollo del repositorio.

2. Antecedentes

Las bibliotecas de la actualidad están integradas por documentos que almacenan los conocimientos creados por la sociedad y han persistido por más de 25 siglos, si se toma como referencia la colección de la dinastía Song (1100 a. de C.) en China, en los últimos años han sido asociadas directamente con las escuelas donde se cuenta con libros, revistas y artículos de investigación para consulta (D. Bainbridge, 1998).

La evolución de la ciencia y tecnología hace que continuamente se realicen estudios sobre los avances logrados en cada una de las áreas de investigación, lo que provoca a

su vez la necesidad de difusión de dichas investigaciones que se producen por diferentes medios (I.H.Witten, 2001).

Hoy en día las tecnologías de Información ofrecen alternativas prácticas en la creación de almacenes y repositorios de información, considerando a un repositorio digital como un sistema que permite almacenar, preservar y dar acceso a objetos digitales. Los repositorios pueden ser de acceso abierto, esto es un repositorio que almacena y ofrece el acceso libre a una comunidad académica, científica o a público en general, en el repositorio se presentan los resultados (Publicaciones, investigaciones, datos), y estos pueden ser consultados por usuarios definidos o por cualquier persona interesada en temas afines.

Los repositorios de acceso abierto de forma común son creados y administrados por Instituciones académicas y/o de investigación, existen repositorios temáticos y repositorios de datos.

3. Características de un Repositorio.

En el caso concreto de México el Consejo Nacional de Ciencia y Tecnología (CONACYT), ha emitido lineamientos generales para los repositorios Institucionales, en los cuales se establecen los tipos de información y vías de ingreso de la información a los repositorios, así como los procesos de depósitos y depuración, las políticas de usabilidad y derechos de autor (CONACYT, 2017).

Entre las características de un repositorio están las que es un servicio institucional abierto a toda la comunidad universitaria y a todo tipo de temáticas.

3.1. Objetivos de un Repositorio Digital.

Los objetivos que se pretende lograr con un repositorio que almacene la productividad de los docentes e investigadores de la Facultad de Contaduría y Administración son:

- Reunir, preservar y dar acceso a, entre otras cosas, la producción de los investigadores y docentes en múltiples formatos.
- Difundir la producción científica, tecnológica y académica de la comunidad universitaria de la FCA.
- Facilitar el acceso a la información científica y académica.
- Propiciar la retroalimentación a la investigación.
- Facilitar la integración de redes temáticas de investigación y redes académicas.

Los repositorios digitales institucionales son los instrumentos de acceso abierto, para difusión de la producción científica y académica, así como también los recursos de académicos. Por “acceso abierto”, nos referimos a su disponibilidad gratuita en la Internet pública, que permite a cualquier usuario leer, descargar, copiar, distribuir, imprimir, buscar o añadir un enlace al texto completo de esos artículos, rastrearlos para su indización, incorporarlos como datos en un software, o utilizarlos para cualquier otro propósito que sea legal, sin barreras financieras, legales o técnicas, aparte de las que son inseparables del acceso mismo a la Internet. La única limitación en cuanto a reproducción y distribución, y el único papel del copyright (los derechos patrimoniales) en este ámbito, debería ser la de dar a los autores el control sobre la integridad de sus trabajos y el derecho a ser adecuadamente reconocidos y citados (I.H.Witten,2001).

De acuerdo a Conacyt "El Repositorio Nacional es una plataforma digital que proporciona acceso abierto en texto completo a diversos recursos de información académica, científica y tecnológica", es decir, sin requerimientos de suscripción, registro o pago. En el Repositorio pueden ser consultados, entre otros materiales: artículos de revistas científicas, tesis elaboradas en instituciones de educación superior, protocolos de investigación, memorias de congresos y patentes, así como otros documentos académicos que se producen en México con fondos públicos (CONACYT, 2017)

Existen herramientas para crear bibliotecas digitales llamados repositorios institucionales que ofrecen una ventaja sobre sólo un repositorio básico, permiten la organización mediante grupos, familias y referencias específicas entre otras ventajas.

3.2. Ventajas de un Repositorio.

Implementar un repositorio digital, en una institución educativa o centro de Investigación ofrece favorables alternativas para realizar la difusión de la productividad académica y de investigación y ventajas como las que a continuación se describen:

- Crear y compartir conocimiento.
- Facilitar aprendizaje organizacional.
- Espacio para almacenamiento e intercambio común de información en las diferentes ramas del conocimiento.
- Base como identificador de una Universidad con niveles de calidad y excelencia.
- Comunicar y difundir los resultados intelectuales a la comunidad científica.
- Preservar la información científica, académica e institucional.

Según Álvarez (2011), los beneficios que obtendrán los estudiantes es Dotarles de las herramientas e instrumentos que necesitan para el aprendizaje, estudio e investigación, Adquirir habilidades para la gestión del conocimiento y conseguir mejores resultados académicos y hacerlos más competitivos.

Analizando las características y ventajas de un repositorio, la Facultad de Contaduría definió como un objetivo prioritario, la creación de un repositorio digital, para esto se analizaron las diferentes herramientas que podía utilizarse en la creación del mismo, las cuales se mencionan enseguida:

- DSpace.
- Greenstone.
- Fedora (Flexible Extensible Digital Object Repository Architecture).
- RefBase.
- Logos Bible Software.

Si bien es cierto hay una variedad de herramientas que es posible utilizar, después de elaborar un minucioso análisis se determinó la realización del repositorio, mediante el sistema de gestión de contenidos.

El desarrollo de repositorios digitales es posible a través de la combinación de nuevas tecnologías de la información como el internet y medios electrónicos que permiten la consulta dinámica de información.

3.3. Sistema de Gestión de Contenidos

Una herramienta que se puede utilizar en el desarrollo de aplicaciones, los cuales proveen la facilidad de crear, organizar y publicar documentos y otros contenidos son los llamados Sistemas de Gestión de Contenidos (Content Management System o CMS), que se pueden conceptualizar como un conjunto de aplicaciones web que funcionan como un portal de igual forma en Internet como en la Intranet. Se puede identificar como un software que puede ser desarrollado y administrado por usuarios, además de gestionar contenidos de una web.

La diversidad que ofrece este tipo de sistemas en cuanto a la organización del sitio e inserción y administración de los contenidos permiten considerar a esta como una opción viable y adecuada.

En el diseño de la herramienta, la publicación de noticias y las estructuras principales de un sistema de gestión de contenidos, se encuentran en el FrontEnd y el BackEnd, en el cual se almacena y organizan los datos.

El Frontend, que es la parte visible del CMS para los usuarios que acceden al sitio, ya sea como invitado o usuario registrado, el Backend, es decir, el área donde se realizarán las tareas de actualización, configuración y gestión, los módulos para la configuración y personalización del sitio web, el gestor de usuarios, perfiles y derechos de acceso, los módulos para la gestión del contenido del sitio web (creación, edición y publicación de contenidos), las plantillas para personalizar el aspecto del sitio web, así como las extensiones que aumentan la funcionalidad del gestor de contenidos.

Para lograr un repositorio digital que atienda a las necesidades y requerimientos de una institución educativa, es necesario conceptualizar el diseño para la creación de una biblioteca digital, para esto se definen los usuarios que lo utilizarán, los materiales que se compartirán, la infraestructura requerida y la tecnología que se implementará, en la figura No.1, se presentan los componentes que integrarían la estructura Organizativa del repositorio académico definido para la FCA.

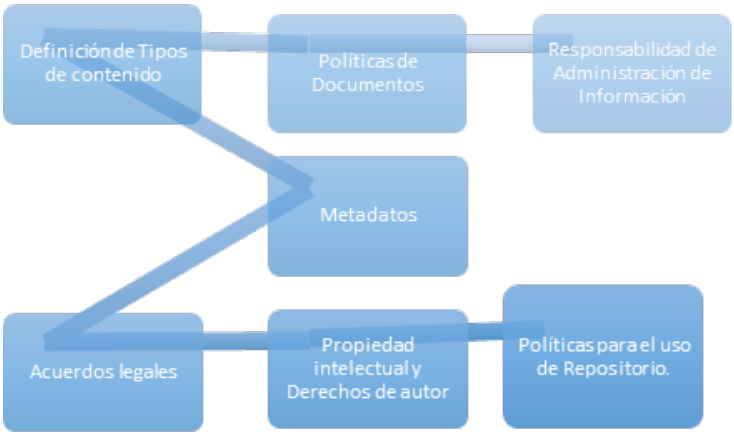


Figura 1 – Diagrama de estructura Organizativa Repositorio FCA.

3.4. Desarrollo de Repositorio Académico

El desarrollo del repositorio académico se realizó en base a la determinación de los requerimientos, los cuales de definieron en base a las funcionalidades establecidas por las autoridades académicas de la Facultad, así como a las características de la productividad generada y las necesidades de publicación.

3.4.1. Requerimientos funcionales.

Ofrecer contenidos de la productividad académico-investigativa de los docentes de la FCA en un espacio abierto, con acceso, visibilidad y preservación de la memoria docente y de investigación.

Dentro de las Entidades definidas para la creación de la base de datos se establecieron:

- Investigadores-Docentes, Estudiantes de Posgrado, Programas de Posgrado, Investigaciones, Libros, Capítulos de libro, Artículos Indexados, Artículos arbitrados, Memorias en extenso.

Cada uno de estos elementos requiere de la creación y manipulación de la información. Un aspecto importante es la definición de los metadatos de cada una de las entidades, compatibles con formatos como el Dublin Core tales como título y subtitulo de cada recurso.

Los diferentes niveles analizados en el Repositorio son:

Nivel de metadatos, Información y esquemas compatibles con estándares Internacionales.

Nivel de Contenido. Proceso para ingreso y publicación de información.

Nivel de red. Red con integración de estándares.

Nivel de catálogos. Catálogos de las entidades.

3.4.2. Productividad académica-Científica.

Se realizó una investigación preliminar para determinar las características de la productividad académica y científica de los miembros de la Facultad, las características y atributos de cada uno de estos productos y las funcionalidades en los que se veían afectados e involucrados.

Se determinaron los recursos obtenidos en el desarrollo de investigación y la práctica docente, los cuales permiten promover la generación de conocimiento y el aprendizaje

3.4.3. Definición de contenidos.

Integración de la producción académica, científica de la Institución, la cual constituye un sistema de gestión documental útil.

- Documentos de acceso libre a la comunidad.
- Tesis Doctorales y tesis de maestría presentadas en la Facultad.
- Ponencias presentadas en Congresos Nacionales e Internacionales.
- Materiales didácticos generados por los docentes y alumnos de la Facultad.

- Patentes y software desarrollado por miembros de la comunidad académica.

3.5. Metodología ágil de desarrollo SCRUM.

Para la elección de la forma de trabajar se analizaron alternativas y se determinó trabajar en la metodología ágil SCRUM, metodología que integra buenas prácticas y el trabajo colaborativo de equipo y obtener mejores resultados, mediante la colaboración de un equipo altamente competitivo. En Scrum se presentan entregas parciales del proyecto, las cuales son priorizadas de acuerdo a las aportaciones que realizan al proyecto y son valoradas por los usuarios finales del proyecto.

Scrum se recomienda en proyectos con entornos complejos, con requerimientos que se modifican y que además es necesario obtener resultados rápidos, en donde la innovación, la flexibilidad y la productividad son básicas.

Para la administración del proyecto Repositorio Digital, se determinó la utilización de la metodología ágil de desarrollo SCRUM, metodología, que considera que los involucrados en el desarrollo, pueden ofrecer mucho más que sus conocimientos técnicos, se trata de asignar y planear en un entorno de constante cambio, la toma de decisiones iterativas está basada en datos y la entrega de productos se realiza en función de los requisitos del cliente.

Existen procesos denominados: Prioritization y Time boxing, self-organization, mediante los cuales es posible determinar prioridades y tiempos para realizar cada una de las fases del proyecto (SBOOK, 2013).

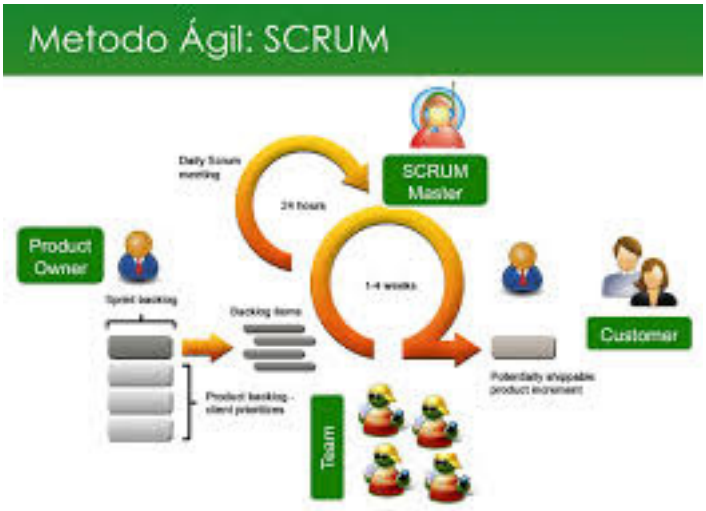


Figura 2 – Metodología ágil SCRUM.Fuente [7]

SCRUM, permite la planificación detallada por adelantado del proyecto, dar énfasis en la gestión y la solución en base a costos, horarios y gestión de parámetros.

El uso de metodología SCRUM en cualquier proyecto, ofrece ventajas como la adaptabilidad, esto permite que en el proyecto se incorpore el cambio, la transparencia, es posible utilizar herramientas que muestran los avances del proyecto como el Scrumboard y Sprint Burndown chart, esto propicia un ambiente abierto, retroalimentación continua, la cual es posible mediante Conduct Daily StandUp y Demonstrate and validate Sprint, el proyecto se divide en entregables, los cuales es posible evaluar y mejorar de manera progresiva. En los proyectos SCRUM, la resolución de problemas se da de forma más rápida, los equipos son multifuncionales y conduce a la resolución de problemas de forma más rápida (Collaboration y Colocation).

El proyecto es centrado en el cliente, se da énfasis en el valor del negocio y tener un enfoque de colaboración con los stakeholders. Los proyectos se dan en un marco de colaboración que permite a los equipos multifuncionales altamente cualificados alcanzar su potencial a una alta velocidad.

Un Equipo Scrum (Scrum Teams) está conformados de seis a diez miembros. Se desarrolló un proyecto en un ambiente de introspección, aprendizaje y capacidad de adaptación que lleva a un entorno de trabajo innovador y creativo.

Se considera que los trabajadores pueden ofrecer mucho más que sus conocimientos técnicos, se trata de asignar y planear en un entorno de constante cambio, la toma de decisiones iterativas se basa en datos, la entrega de productos que satisfacen requisitos del customer en pequeños incrementos iterativos.

El equipo para el desarrollo del proyecto se conformó con estudiantes de la Licenciatura en Informática, algunos de ellos certificados en los fundamentos de SCRUM y el liderazgo de un equipo de docentes de la misma área con la certificación de SCRUM Master. Guíados por la metodología se avanza en el proyecto atendiendo cada una de las tareas y elementos que la constituyen.

Aunado a la implementación de las reuniones diarias con el equipo que conforma el proyecto y la utilización de repositorios de información existentes como es el caso de la herramienta *Basecamp* ha facilitado la comunicación y seguimiento de avances en el desarrollo del sistema.

3.5.1. Desarrollo Tecnológico.

El desarrollo del sistema se llevó a cabo con actividades específicas en cada una de las etapas, se asignaron tareas por cada uno de los integrantes del equipo. Se realizó el análisis y la selección de las herramientas adecuadas para el óptimo avance. Una vez elegida la metodología de desarrollo y determinados los requerimientos del repositorio se realizó el modelado de la base de datos que integra el Backend del sistema, en el cual se plasman cada uno de los elementos que conforman el repositorio institucional. En la figura 3, se muestra el Diagrama Entidad Relación del repositorio de la FCA, en el que es posible observar los elementos principales del repositorio.

4. Conclusiones

La elaboración y puesta en marcha de un repositorio en el cual se integra la productividad de un sector, de manera particular de los docentes, investigadores y

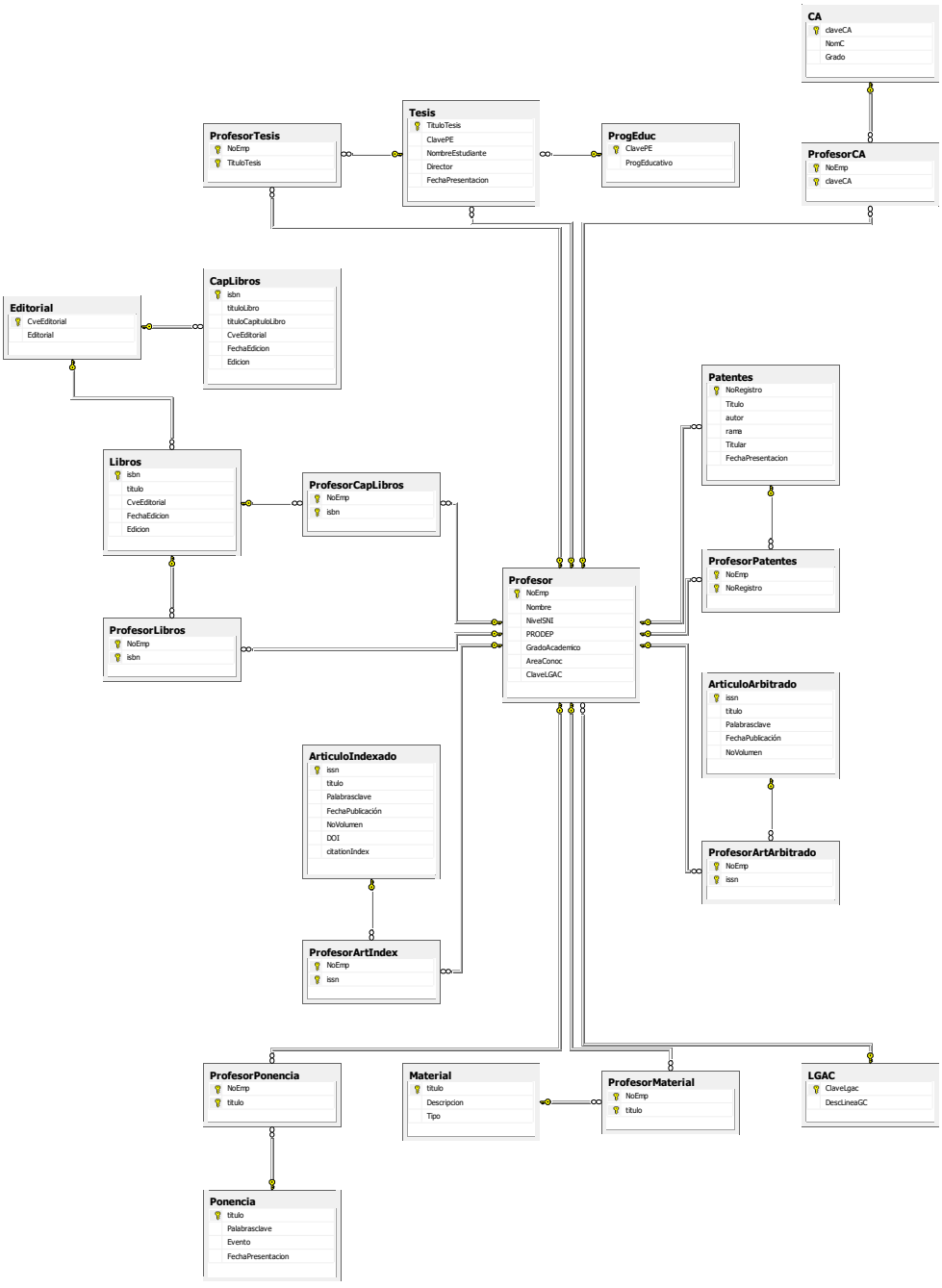


Figura 3 – Diagrama Entidad-Relación Repositorio FCA.

estudiantes de la FCA, ha permitido avanzar en el campo de la investigación aplicada como es la Administración las tecnologías de la información y la comunicación y áreas específicas de la Contaduría y los Negocios Internacionales, presentando los resultados de proyectos que se desarrollan en conjunto con organismos regionales, nacionales e internacionales.

Una vez planteado el objetivo de la creación de un repositorio, se integran actividades y procesos encaminados a lograr esta meta, es importante la documentación de cada una de estas fases incluidos la selección de herramientas y tecnologías que se utilizan para la creación de repositorios de instituciones educativas o centros de investigación, de modo que sirva de base para nuevos desarrollos y en el que se reconocen las características de todos los recursos incluidos los de capital humano.

Cada fase tiene implícitas tareas concretas, actores involucrados, reglas y restricciones por lo que es de vital importancia para un proyecto de esta magnitud la selección adecuada de metodologías y de herramientas que faciliten el resultado.

El uso de SCRUM permitió realizar una distribución adecuadas de responsabilidades y tareas entre los miembros del equipo, dar seguimiento puntual a los avances de las metas (Sprint), y poder realizar los ajustes necesarios en caso de así requerirse.

La fusión de tecnologías permite mayor eficiencia en los tiempos y recursos. Por otra parte la integración de materiales académicos y de investigación es un reto importante que vale la pena realizar, de manera que estos sean aprovechados por los estudiantes, docentes, público, por cualquier generador de conocimiento.

Referencias

Alvarez, J. y Alvarez, M. (2011). La importancia de los repositorios institucionales para la educación y la investigación. Synthesis, volumen (57), 43-48. Recuperado 12 septiembre 2018 de http://www.epn.edu.ec/wp-content/uploads/2017/03/la_importancia_de_los_repositorios_institucionales.pdf

Casate, R. (2009). Propuesta de perfeccionamiento de los servicios de la red cubana de la ciencia en correspondencia con los principios y fundamentos Tecnologicos del Acceso Abierto. Cuba: Tesis Doctoral.

CONACYT. (2017). CONACYT. Obtenido de UNA GUIA PARA EL SCRUM CUERPO DE CONOCIMIENTO (GUIA SBOKTM): www.conacyt.gob.mx

CSIC. (2012). Ranking web de Repositorios. Obtenido de Repositories.webometrics.info: http://repositories.webometrics.info/es

IH. Witten, M. L. (2001). La promesa de bibliotecas digitales en lols paises en desarrollo . En M. L. IH. Witten, Comunicaciones de la ACM (págs. 82-85). Vol.55.

IH.,Witten, M.,Loots, M.F. Trujillo and D.Bainbridge,(2001) “The promise of digital libraries in developing countries”.Communications of ACM, vol 55, No.5, pp. 82-85.

JP Alperin, D. B. (2015). Open Access indicators and scholarly communications ONU LatinoAmerica. Conbsejo Latinoamericano de CIencias Sociales (págs. 237-246).

Liberty, J., Horovitz, A. (2008). Programming .NET 3.5. California : O'reilly Media Inc.

SCRUM Study (2013). "Una guía para el conocimiento de SCRUM", Guía SBOK.

SCRUMstudy Targeting sucess, SMC WORK BOOK Student, 2014.

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 10/10/2018
Aceitação/Acceptance: 12/12/2018

Complejidad y gestión del conocimiento e innovación

María del Consuelo Salgado Soto[1], Ricardo Fernando Rosales Cisneros[1], Nora del Carmen Osuna Millán[1], Margarita Ramírez Ramírez[1], Hilda Beatriz Ramírez Moreno[1], Maricela Sevilla Caro[1].

csalgado@uabc.edu.mx, ricardorosales@uabc.edu.mx, nora.osuna@uabc.edu.mx, maguiram@uabc.edu.mx, ramirezmb@uabc.edu.mx, mary_sevilla@uabc.edu.mx.

[1] Universidad Autónoma de Baja California, 22000, Tijuana, Baja California, México

Pages: 1073–1082

Resumen: Las organizaciones requieren practicar una forma de aprendizaje colaborativo como una estrategia de mejora continua. Generar innovación construye capacidades y competencias nuevas traducidas en ventajas competitivas para las organizaciones, ya que, en mundo complejo, estas ventajas son claves para diferenciarse en el mercado global. El presente trabajo aborda la gestión del conocimiento y la innovación desde un enfoque de la complejidad. Se discute como un modelo de auto-organización puede enfrentar la complejidad, y como juega un papel importante para la mejora continua en las organizaciones.

Palabras Clave: Complejidad, Gestión, Conocimiento, Innovación, Sociedad

Complexity and management of knowledge and innovation

Summary: Organizations require a form of collaborative learning as a strategy for continuous improvement. Generate innovation, skills and new competencies, translated into the best competencies for organizations, since in a complex world, these advantages are key to differentiate in the global market. The present work approaches knowledge management and innovation from a complexity approach. It is discussed how a model of self-organization can face complexity, and how it plays an important role for continuous improvement in organizations

Keywords: Complexity, Management, Knowledge, Innovation, Society

1. Introducción

Estamos viviendo una revolución en los sistemas de gestión del conocimiento donde las personas y organizaciones hacen frente a su necesidad de aprender, comunicarse y tomar decisiones con ayuda de medios de sistemas informáticos, en donde el conocimiento de la naturaleza de los sistemas complejos puede ayudar a convertir la gestión en algo manejable y llevarlo a la práctica.

Las organizaciones por si solas viven en un entorno complejo que está cambiando todo el tiempo y con múltiples elementos interactuando de forma dinámica. Con la inclusión de la tecnología llego para simplificar procesos, actividades, comunicaciones, tareas, sin embargo, en ocasiones puede significar un nivel de complejidad, cuando está no es adoptada, aplicada o entendida correctamente. En muchos casos, los humanos son capaces de percibir una verdadera situación emergente sin ayuda de la tecnología, pero en otras, está permite simplificarle el nivel complejidad.

La creación de modelos puede ayudar a gestionar la complejidad que se presenta. sin embargo, no pueden decir más de lo que necesita saber y deben hacerlo de manera eficaz y menos ambigua que antes. Las personas interactúan todo el tiempo en entornos dinámicos con el fin de disminuir el nivel de complejidad de sus contextos de interacción (Etkin, 2014).

1.1. Complejidad

La complejidad es aquella que conjuga la razón, emoción, ciencia, filosofía, imaginación y realidad; permite entender que lo que se ha creado es real, pero esto puede ir más allá y convertirse en otro entender. La complejidad normalmente es tomada como sinónimo de dificultad, puede ser considerada como algo que no puede ser codificado, lo que significa que es lo contrario de fácil (Agazzi, 2002). Wegener (2015) menciona que tiene la meta de probar que los problemas difíciles no pueden ser resueltos con recursos modestos. Así mismo puede ser definida como la medición del entendimiento de la incertidumbre con el fin de alcanzar los requerimientos funcionales como lo comenta Suh, (2005).

Cuando las partes actúan independientemente entre sí, el comportamiento a escala final es más complejo, porque habría que describir por separada cada elemento, pero en cambio no hay comportamiento organizado a gran escala. En el límite a gran escala no solo el sistema no es complejo, no puede ser que no haya un sistema. En cambio, cuando las partes actúan conjuntamente, la complejidad a pequeña escala puede ser más pequeña, pero hay un comportamiento organizado a escala mayor por lo que esta escala si puede tener una gran complejidad (Bar-Yam, 2002).

1.1.1. Sistemas complejos

El enfoque sistémico inicia en examinar y comprender las relaciones entre los diferentes elementos del sistema. Este enfoque es una teoría general donde sus principios se pueden aplicar a cualquier disciplina o área. Una de las maneras más comunes para abordarlo es al concebir que “el que un todo es solo más que la suma de sus partes” (Miller, 2007).

El campo de los sistemas complejos cuestiona la idea de comprender perfectamente el comportamiento de cada uno de los componentes de un sistema, sino comprender el sistema en su conjunto. Los sistemas complejos están compuestos de un gran número de elementos que interactúan, además, dos de sus propiedades son consideradas como complicadas: emergencia y auto-organización.

- La emergencia es la aparición de un comportamiento que no se puede prever a partir del conocimiento de los componentes del sistema por si solos. Las propiedades recién emergentes dan retro-alimentación de información a las

entidades originales de nivel inferior, entrando en un ciclo de retroalimentación, donde cada elemento (micro/macro nivel) interactúa (Csiro, 2008).
- La auto-organización significa que no hay un controlador externo o planificador que ayude a la aparición de las características emergentes, sino que aparecen de forma espontánea (Csiro, 2008). Cada uno de estos dominios se compone de un conjunto de diversas entidades y actores que interactúan de forma dinámica y se sumergen en un mar de información (Miller, 2007).

1.2. Gestión del Conocimiento

La gestión del conocimiento es la identificación y aprovechamiento del conocimiento colectivo, puede ayudar a que se tenga una mayor competitividad. El desempeño general de las organizaciones depende de la extensión de los administradores para movilizar todos los recursos retenidos por los individuos y equipos y que estos recursos se tornen con el fin de crear actividades de valor (Von Krogh, 1998). La gestión del conocimiento es el desarrollo de las actividades involucradas tales como el descubrir, capturar, compartir y aplicar el conocimiento para mejorar hacia un costo efectivos para alcanzar la meta (Becerra, 2004). Para gestionar el conocimiento es importante apoyarse de la tecnología con la cual se pueden crear distintos sistemas tales como:

- Sistemas que descubran conocimiento a través de la implementación de algoritmos inteligentes.
- Sistemas de captura de conocimientos que preserven y formalicen el conocimiento de expertos y que este sea compartido con otros.
- Sistemas de compartimiento del conocimiento, organizando y distribución del conocimiento en repositorios de conocimiento.
- Sistemas de aplicación de conocimiento, asisten en la resolución de problemas. (Becerra, 2004).

1.3. Gestión de la Innovación

La innovación es una idea, objeto y/o práctica que esta es percibida como algo nuevo o como una unidad adopción. Esto implica que la percepción de la innovación puede o no ser objetivamente nueva (Pocket, 2013). Cuando la innovación está presente es importante gestionar la misma, por ejemplo, la gestión de la innovación es el conjunto de técnicas que permite la identificación del potencial y los problemas de la empresa y/o organización, se aplica con el fin de elaborar e implementar planes de innovación y mejora continua, a efectos de mejorar su efectividad (Solleiro, 2008). En la Figura 1.0 muestra los elementos correlacionados del impacto de la innovación.

Otra definición de gestión de la innovación es la que aporta Corre, (2006) y que menciona que es la organización, dirección de los recursos, humanos económicos, tecnológicos, de procesos, que tienen el objetivo de aumentar nuevos conocimientos; generar ideas técnicas que permitan obtener nuevos productos, procesos, y servicios o mejorar los ya existentes.

Esta gestión es un aspecto fundamental debido al impacto directo en las áreas generadoras de valor. Realizar esta actividad adecuadamente implica conocer el mercado, tendencia tecnológica y la capacidad de los competidores, además, evaluar resultados y conseguir la

optimización de los procesos productivos como menciona Solleiro, (2008). En la figura 2.0. muestra que la innovación continua depende de la creación de nuevo conocimiento, una innovación continua desarrolla ventajas competitivas.

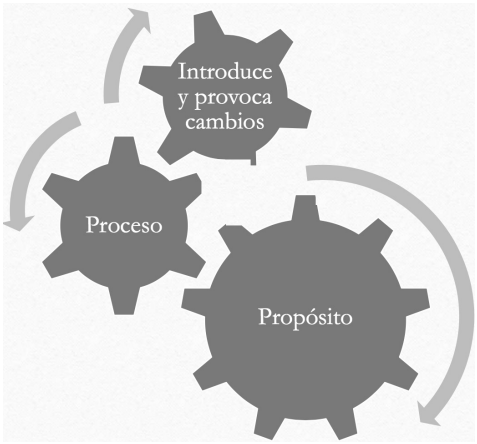


Figura 1 – Impacto de la Innovación

Figura 2 – Relación entre conocimiento e innovación

2. Gestión de la Complejidad en las organizaciones

Ya se ha definido a la organización como un sistema complejo de carácter socio técnico y adaptativo, además, en ella se conjugan distintos factores culturales, políticos, económicos y sociales, sin embargo, aun cuando estos factores están relacionando entre si y de manera complementaria al mismo tiempo presentan oposiciones y procesos paralelos. El tiempo y la influencia de los distintos grupos que operan en el sistema y los cambios repentinos en el ambiente hacen que la organización se oriente en varios sentidos a la vez.

Un factor clave para alcanzar el éxito en las organizaciones es comprender la complejidad que le afecta para hacer frente a la misma en un mundo tan caótico y competitivo, aunque alguno de los recursos que se tienen tanto operativos como intelectuales pudieran ser insuficientes para gestionar y reducir la complejidad.

Los sistemas complejos son imprescindibles y esto pone en un predicamento al tipo de gestión basada en el orden y el control. Hoy en día las organizaciones necesitan manejar la complejidad reduciéndola o absorbiéndola, deben de ser coparticipes de la complejidad y hacerla parte de ella. Para lograr esto se debe definir un balance de todas las áreas según sus requerimientos, por ejemplo, el área de ventas debe estar a la par de las necesidades del cliente, es decir debe estar lista para lo que demande el mercado permitiéndole ser altamente competitiva. Por otra parte, la organización también debe entender todas las interacciones que existen entre los elementos que la componen, así como aquellas que se realizan con otros actores (proveedores, clientes, socios), esto es importante porque se debe adaptar a los requerimientos de ellos, se deber de absorber la complejidad y simplificarla (Etkin, 2014).

2.1. Abordar la complejidad

Para abordar la complejidad es importante apoyarse a través del aprendizaje y comunicación, por ejemplo, el dialogo se asemeja al sistema adaptativo complejo ya que la emergencia se produce en base a sus reglas y relaciones iniciales adecuadas (López, 1997), es importante que las organizaciones escuchen a sus clientes teniendo canales de comunicación que arrojen distintos puntos de vista para producir mejores productos y servicios.

Un aspecto importante es considerar las perspectivas de las organizaciones que admitan la posibilidad que más de unas cosas sea cierta a la vez, esto permite tener distintas interpretaciones sobre la racionalidad de los conceptos y de la perspectiva de cada uno. Otro punto importante, es abrir un espacio a la posibilidad, las organizaciones necesitan explorar espacios de distintas opciones e ideas que pueden llevar a ser realidad, esto fomenta que las organizaciones mejoren su capacidad de inventar y reinventar continuamente, es importante que se hagan exploraciones a los entornos escuchando a todos los involucrados en los distintos niveles de la organización con el fin de que participen en una búsqueda continua (Etkin, 2003). En la figura 3.0 se muestra la zona de la posibilidad en una organización.

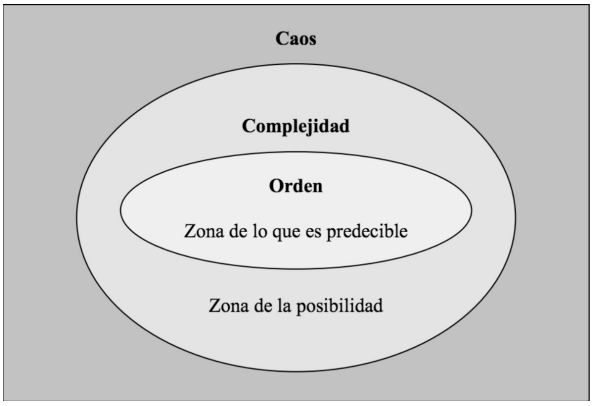


Figura 3 – Zona de la posibilidad

2.2. Auto organización ante la complejidad

La organización necesita ser sensible a su contexto ya que la complejidad lo exige y fomentar la auto organización permitiéndole mayor estabilidad y flexibilidad, para repercutir en una dinámica interna, y lograr estar preparada para las interacciones cambiantes del entorno.

Esta auto organización permite formar estructuras disipativas que emergen del resultado de los procesos auto organizativos en continua interacción con el entorno. Las organizaciones necesitan gestionar las complejidades emergentes y las mismas debe de dar respuestas a las posibilidades por medio del aprendizaje y la comunicación. También permite entender mejor el comportamiento complejo, así como las nuevas estructuras derivadas del desequilibrio y del caos progresivo (Cornejo, 2004).

Derivado de la gran complejidad y caos que existe hoy en día en el entorno a nivel mundial, las organizaciones que sobresalen de sus competidores y que están a la vanguardia son aquellas que ya no tienen estructuras rígidas y verticales, estas tienen estructuras adaptativas y colaborativas, que están en red buscando siempre un control y aprendizaje que les permite tener un equilibrio en todos los niveles de la misma.

3. Modelo de una sociedad del conocimiento como solución de problemas complejos

La representación abstracta representada en un modelo es por si solo un proceso complejo y a que varios factores deben ser considerados. La complejidad emerge de la necesidad de tener en cuenta no solo los factores que contribuyen a una persona el conocimiento de la situación sino también a los factores que contribuyen cuando este pertenece o es miembro de una sociedad del conocimiento, lo cual conlleva a un conocimiento de la situación compartida. Así mismo dentro de cada uno de estos elementos existen múltiples factores que afectan el conocimiento de la situación, tales como la distribución geográfica, el liderazgo, el uso de herramientas de colaboración, la proximidad de la red, las experiencias de antecedentes similares y familiaridad (Carballo, 2009).

El crear un modelo que permita acceder a servicios y/o información de forma continua e interrumpida, que proporcione acceso instantáneo inclusive hasta en situaciones de emergencia, que evolucione para satisfacer las necesidades de información, cada vez más amplia en base al requerimiento, acciones, comportamiento, o rendimiento por parte de los seres humanos, que soporten una sociedad del conocimiento, deben ofrecer las siguientes características:

- Información za`àr que esté disponible en todas partes, en cualquier momento, y para todos, independientemente de sus habilidades.
- Acceso a la información en diferentes contextos.
- Tener la misma información presentada distintamente para los distintos involucrados en la sociedad.
- Proporcionar el mismo contenido a distintos dispositivos.
- Interactuar con esta información usando una variedad de distintos dispositivos.

- Satisfacer las necesidades de los usuarios desde el acceso, manipulación, análisis y control de la información.
- Gestión integral de la información del entorno.

Un modelo para una sociedad del conocimiento aborda cuestiones respecto a los usuarios en sus actitudes, percepciones, aceptación del uso de tecnología, interacción con la información y tareas. La figura 4.0 representa una vista general de algunos componentes considerados en una modelo para una sociedad del conocimiento.

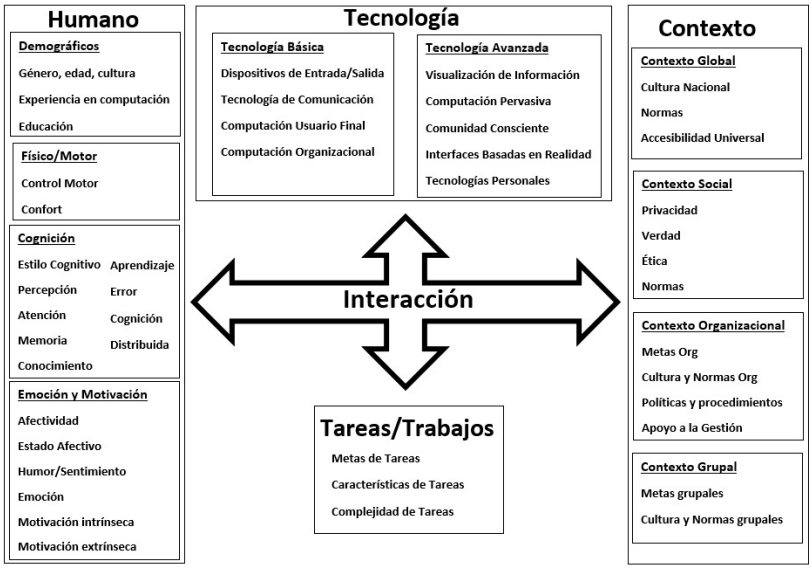


Figura 4 – Componentes de modelo para una sociedad del conocimiento

El desarrollo de un modelo para una sociedad del conocimiento, implica que este inmerso en un campo interdisciplinario debido a que no basta una solo disciplina para que pueda completamente abordar un ambiente complejo. Es importante considerar a los humanos como un componente clave del modelo independientemente de la perspectiva disciplinaria, dado que el análisis de los seres humanos en este contexto se basa en gran medida en ideas sobre la psicología humana, comportamiento, acciones, percepciones, y también han sido estudiados por algunas perspectivas que se mencionan a continuación:

- Individuos o grupos con diferencias individuales, tales como rasgos generales, específicos, estilos cognitivos y personalidades (Agarwal, 1998).
- Individuos o grupos como actores sociales en el diseño, desarrollo y uso de tecnologías de la información y la comunicación. Se puede argumentar que la mayoría de las personas que utilizan aplicaciones TIC utilizan múltiples aplicaciones, en varios roles, como parte de sus esfuerzos por producir bienes y servicios, mientras interactúan con otras personas a menudo en múltiples contextos sociales (Lamb, 2003).
- Individuos o grupos como agentes económicos cuyas preferencias, comportamientos, personalidades, y en última instancia, el bienestar económico está estrechamente vinculados con el diseño de sistemas de información (Bapna, 2004).

En general, es interesante examinar como el concepto individuos o grupos ha evolucionado a partir de una sociedad difusa con propósito y pertenencia dinámica.

Esta inevitable evolución está desafiando el diseño y temas de investigación respecto a la interacción entre estas sociedades como lo es el conocimiento de la situación compartida. El tener un modelo que ayude a soportar una sociedad del conocimiento con conciencia de la situación mejora la interacción entre el usuario y su ambiente recibiendo contenido, información, servicios que realmente requiera para elevar el nivel y calidad de experiencia en la interacción, así como el interés sobre el contenido ofrecido, repercutiendo en relaciones con otros usuarios para intercambio de opiniones y experiencias dentro de una sociedad.

Por otra parte, un modelo de una sociedad del conocimiento para enfrentar un mundo lleno de complejidad con conciencia de la situación, se debe considerar a los involucrados como el centro del desarrollo, sus capacidades y sus tareas. Se debe tener bien claro: ¿Quiénes son los involucrados potenciales? ¿Qué tan variados son? ¿Cuál es su comportamiento actual? ¿Qué aspectos de su comportamiento son difíciles de entender? ¿Cuáles son sus necesidades?, así mismo, se debe considerar algunas estrategias claves tales como involucrar a los interesados, realizar un diseño iterativo, conocer el rol que desempeñan los involucrados (Olson, 1991).

Un modelo que permite entregar información y/o servicios adecuados en base a las peticiones solicitadas o censadas del contexto, puede beneficiar a individuos, grupos, organizaciones y/o sociedades aumentando su rendimiento. La idea clave es que la tecnología utilizada para entregar esta información tenga un impacto positivo si se ajusta a la tarea que soporta.

4. Conclusiones

El desarrollo de modelo de una sociedad del conocimiento sensible de la situación ayuda a la detección de varios tipos de contextos que pueden llegar presentarse, permite el acceso de conocimientos entre los distintos individuos, utilizando mecanismos de aprendizaje masivos, reforzando el conocimiento individual, maximizándolo con un conocimiento colectivo.

Se debe tener presente que actualmente se vive en un mundo caótico totalmente competitivo lo que implica que la complejidad está en todas partes, esta presenta distintas situaciones que no están pronosticadas o planeadas, por lo que las organizaciones deben de auto organizarse para afrontar circunstancias emergentes, aquí en esta situación, el conocimiento compartido será clave para las soluciones de las posibles problemáticas. El generar innovación es conocimiento, por lo que los individuos involucrados son innovadores, estos en conjunto, hacen que la organización tenga una estrategia clara, que cumpla sus objetivos, que controle su gestión, que tenga un conjunto de procesos definidos y asegurados, un sistema de comunicación externa e interna, sistema de evaluación de resultados y de mejora es decir una cultura corporativa propia. El crear una sociedad del conocimiento es una herramienta clave para la gestión ya que garantiza un ritmo de aprendizaje que se acerca al ritmo de los cambios del mercado, esto permite maximizar el conocimiento, tener claridad del proyecto, el saber que queremos hacer y a dónde queremos ir, tener claro un método definir la información

que se tiene que recolectar y que procesar, tener grupos redes, nos permite tener cooperación, colaboración, aprendizaje y desarrollo, así mismo el tener la identificación de las cualidades personales de cada integrante beneficia el crear nuevas estrategias aumentando el desempeño colaborativo.

La acumulación de conocimientos a través de la realización de distintos proyectos, repercute en el aumento de la experiencia de los involucrados, va formando un capital intelectual de innovación en una organización que permite obtener competencias esenciales en la cadena de valor que representa tener ventajas para el presente y para el futuro.

Así mismo, es importante considerar que deben crearse herramientas que permitan abordar la complejidad con una gestión del conocimiento e innovación efectivas, para esto es importante que esta herramientas sean sensibles al contexto y proporcionar mecanismos genéricos para la soluciones de problemáticas emergentes, además, deben inferir en contextos de todos los niveles y rangos posibles, deben ser distribuidas, heterogéneas, y con noción de todos los contextos con el fin de interactuar con las entidades involucradas

La concentración de la información emergente es un factor clave para crear sociedades del conocimiento que permitan maximizar el capital humano de las organizaciones, fortalecer a las organizaciones para incursionar en procesos de innovación constante como clave para alcanzar las metas u objetivos trazados, así mismo permitir la creación de empresas vanguardistas para quedar fuera del mercado.

Es importante concientizarse que una organización con bases de conocimientos solidas en toda su estructura mejorara constantemente en sus procesos, productos y servicios generados, siempre estará en constante evolución, identificando ideas y que a su vez se transfieran, que se desarrollen y que se cristalicen en nuevos producto o servicios. Eso permite que un mundo tan globalizados y de servicios tan diferenciados, la innovación ofrezca ventajas competitivas. Por último, la innovación es la clave para abordar la complejidad y la dificultad que esta representa, por lo tanto, hace que se revierta hacia la simplicidad.

Referencias

Agarwal R and Prasad J (1998). A conceptual and operational definition of personal innovativeness in the domain of information technology. Information Systems Research,9-2:204–215.

Agazzi, E and Motecucco (2002). Complexity and Emergence: Proceedings of the Annual Meeting of the International Academy of the Philosophy of Science, Bergamo. World Scientific. ISBN: 9812381589

Bapna R, Goes P, and Gupta A (2004). User heterogeneity and its impact on electronic auction market design: an empirical exploration. MIS Quarterly, 28-1:21–43, 2004.

Bar-Yam, Y. (1997). Dynamics of Complex Systems, Perseus Press, Reading, MA,1997

Becerra-Fernández, et al. (2004). Knowledge Management: challenges, Solutions and Technologies. Upper-Saddle River, NJ: Pearson Education.

Carballo C.R (2009) Ediciones Díaz de Santos. ISBN: 8479789247

Cornejo A.A. (2004). Complejidad y Caos. Guía para la Administración del Siglo XXI. Juan Carlos Martínez Coll. ISBN: 8468883190

Corre A and Mischke G (2006). A New Approach to Innovation Management and R&D. Springer Science & Business Media. ISBN: 9780387238432.

CSIRO Publishing (2008). Complex or just complicated: what is a complex system? Recuperado de: http://www.csiro.au/ast

Etkin J (2003). Gestión De la Complejidad en un Entorno Competitivo la Complicada Relación entre la Eficacia y los Valores Sociales. Revista de Ciências da Administração v.5, n.10.

Etkin J (2014). Gestión de la complejidad en las organizaciones: La estrategia frente a lo imprevisto y lo impensado. Ediciones Granica. ISBN: 9506418225

Lamb R and Kling R (2003). Reconceptualizing users as social actors in information systems research. MIS Quarterly, 27-2:197–235, 2003.

López, F. (1997). Complejidad y educación, Revista Española de Pedagogía, (206): 106–112.

Miller J and Page S (2007). Complex adaptive systems: an introduction to computational models of social life. Princeton University Press.

Olson G and Olson J (1991). User-centered design of collaboration technology. Journal of Organizational Computing, 1-1:41–60, 1991.

Pocket Mentor (2013). Executing Innovation: Expert Solutions to Everyday Challenges, Harvard Business School Press. ISBN: 9781422129555

Solleiro J.L. and Castañón R (2008). Gestión Tecnológica, Conceptos y Practicas. Plaza y Valdez, S.A. de C.V. ISBN: 978-970-722-763-7.

Suh, N.P. (2005). Complexity: Theory and Applications. Oxford University Press. ISBN: 0195178769

Von Krogh (1998). Care in knowledge creation. California Management. Review. v40. 133–154.

Wegener I, (2005). Complexity Theory: Exploring the Limits of Efficient Algorithms. Springer Science & Business Media. ISBN: 3540210458

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 16/09/2018
Aceitação/Acceptance: 07/11/2018

Sistema de Gestión Académica a través del desarrollo de Modelo-Vista-Controlador

Esperanza Manrique Rojas[1], Margarita Ramírez Ramírez[2], Hilda Beatriz Ramírez Moreno[3], María del Consuelo Salgado Soto[4], Nora del Carmen Osuna Millán[5], Luis Manuel Cerda Suarez[6].

emanrique@uabc.edu.mx, maguieram@uabc.edu.mx , ramirezmb@uabc.edu.mx, csalgado@uabc.edu.mx, nora.osuna@uabc.edu.mx. luismanuelcerdasuarez@gmail.com

[1] Universidad Autónoma de Baja California, 22000, Tijuana Baja California, México.

2 Universidad Autónoma de Baja California, 22000, Tijuana Baja California, México.

[3] Universidad Autónoma de Baja California, 22000, Tijuana Baja California, México.

[4] Universidad Autónoma de Baja California, 22000, Tijuana Baja California, México.

[5] Universidad Autónoma de Baja California, 22000, Tijuana Baja California, México.

[6]Universidad Internacional de la Rioja, 26006, Madrid, España

Pages: 1083–1093

Resumen: El presente proyecto de desarrollo de los módulos base de un sistema de gestión académica para la Facultad de Contaduría y Administración o por sus siglas SIGAF, a través de una arquitectura de desarrollo Modelo-Vista-Controlador y obedeciendo las mejores prácticas, metodologías y técnicas del desarrollo de software construido sobre tecnologías web modernas como PHP, Javascript, jQuery, Ajax, HTML5 y CSS3, utilizando marcos de trabajo como Laravel con la finalidad de realizar código sustentable y mantenible para un proyecto de desarrollo sustentable y continúo.

Palabras-clave: Backend, Programación, Sistemas, Horarios, MVC.

Academic Management System through the development of Model-View-Controller.

Abstract: The present project of development of the base modules of an academic management system for the Faculty of Accounting and Administration or by its abbreviations SIGAF, through a Model-View-Controller development architecture and obeying the best practices, methodologies and techniques of the software development built on modern web technologies such as PHP, Javascript, jQuery, Ajax, HTML5 and CSS3, using frameworks such as Laravel with the aim of making sustainable and maintainable code for a sustainable development project and continuous.

Keywords*:* Backend, Programming, Systems, Schedules, MVC.

1. Introducción

En las organizaciones, incluyendo las educativas, poseen sistemas de información los cuales alimentan a grandes bases de datos. Estas bases de datos con gran volumen de información tienden a convertirse en un gran problema para las organizaciones que no poseen una arquitectura de aplicaciones adecuada y de datos correctamente definida. El manejo de estos grandes volúmenes de datos alimentados por varios sistemas informáticos lleva a crear dificultades en las tareas de comunicación e intercambio de información entre aplicaciones (Narvaes & Piedra, 2018).

Actualmente en la Facultad de Contaduría y Administración de la Universidad Autónoma de Baja California, campus Tijuana, México, se imparten 4 Licenciaturas, se ofertan 240 asignaturas, imparten clases 300 a 350 docentes distribuidos en 120 grupos al semestre. Todos estos datos son organizados y combinados por el personal administrativo para crear los horarios de cada grupo semestralmente. Este proceso de generación de horarios se realizaba durante los últimos meses de cada semestre y organizado por un sistema de información local con múltiples limitaciones para el mantenimiento de módulos.

Algunas instituciones educativas han implementado sistemas de horarios a través de aplicaciones informáticas y han logrado de esta manera optimizar la gran mayoría o la totalidad los procesos de generación y gestión de horarios, estos sistemas reducen significativamente largos tiempos de trabajo y horas hombre invertidas en estas actividades. Instituciones educativas tanto regionales como extranjeras, han solventado esta problemática mediante la aplicación de la tecnología con el uso de los sistemas de información.

Con la introducción de los sistemas de información se automatizan los procesos claves en las organizaciones e instituciones al hacer más eficiente las operaciones relativas al giro en el que se desenvuelven, así la generación y gestión de los horarios tiene un área de oportunidad para adoptar este avance tecnológico, como los son los sistemas de información, en la actualidad existen soluciones informáticas y sistemas que agilizan todo este proceso de logística.

Existen instituciones donde grupos de investigadores ha propuesto iniciativas tecnológicas en el campo de la educación, con el objetivo de mejorar los procedimientos comunes de aprendizaje en las instituciones, a través del uso de tecnologías semánticas para hacer que los datos del programa de estudios sean comprensibles tanto para los humanos como para las computadoras reutilizando modelos ontológicos existentes de la mejor manera posible para facilitar la inclusión y la interoperabilidad de los nuevos datos en la red de datos (Sequicela, 2018).

Es las instituciones de educación tiene importancia los sistemas que impactan el proceso de enseñanza-aprendizaje como aquellos sistemas de gestión administrativa. Por esta razón se realizo un proyecto donde se desarrolló el software: Sistema de Gestión Académica de la Facultad de Contaduría y Administración (SIGAF), utilizando la metodología ágil SCRUM, dividiéndolo en FRONTEND, BACKEND y Diseño de Base de Datos. En este escrito se aborda las herramientas utilizadas en desarrollo BACKEND.

Backend se define como el área que soluciona procesos de información, fluidez de datos, arquitectura interna y usualmente está basado en servidores remotos (Vega, 2013).

El desarrollo de plataformas web que agilicen y sean eficientes al momento de formar los horarios, y con la aplicación correcta de las técnicas y herramientas de programación se logra construir la arquitectura adecuada que de soporte y estructura a los sistemas web. Debido a la especialización que requieren las aplicaciones, han surgido los profesionistas en BACKEND. En la actualidad, este tipo de profesión es la encargada de desarrollar la arquitectura de programación de cualquier aplicación web, además, el profesionista Backend necesita interactuar con un equipo especializado en diferentes áreas del desarrollo web, como el Frontend, Base de Datos, Análisis, Pruebas y Documentadores. Una de las actividades principales de esta área es unir el modelo de negocio (base de datos) de las aplicaciones con las interfaces donde el usuario interactúa en las plataformas y sistemas web. (Valdés, 2007). Para el desarrollo del proyecto, se diseñó un marco conceptual que se especifica en la figura 1.

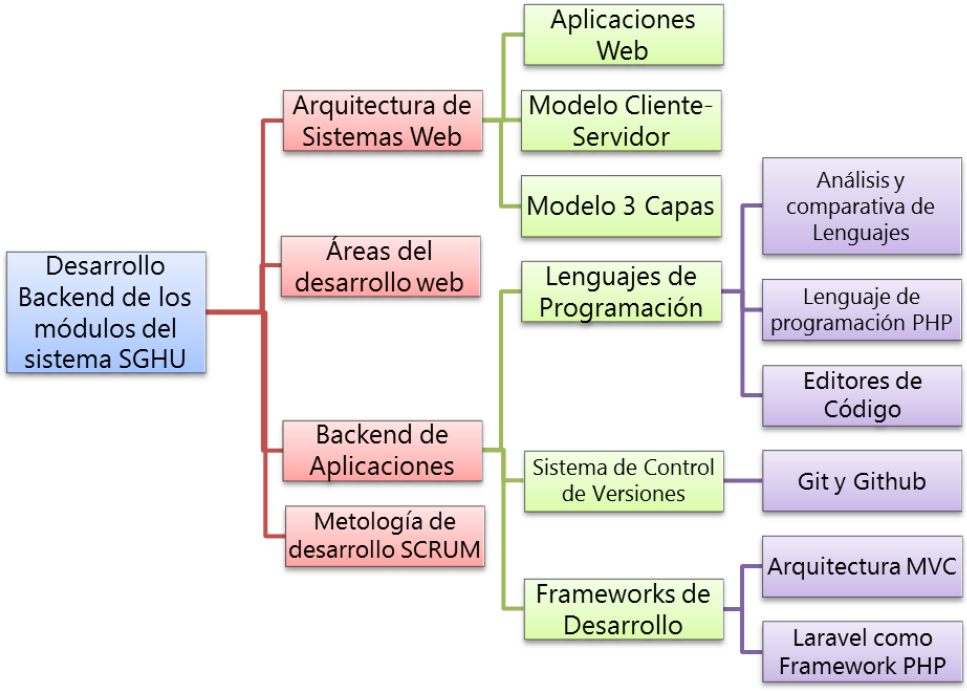


Figura 1 – Diagrama marco conceptual. Fuente Propia.

Existen lenguajes especializados para el desarrollo web, el solo hecho de programar en estos lenguajes, no asegura un orden o una estructura eficaz para el procesamiento de la información, para esto, existen modelos de programación probados y definidos, uno de los más utilizados en la actualidad es la Arquitectura N-Capas,(Paredes,2011) este estilo de programación tiene como objetivo principal, separar los diferentes aspectos del desarrollo, tales como la presentación, lógica de negocio y mecanismos de almacenamiento. Esto surge, a raíz de la desorganización de código que existía en la programación de los sitios web, comúnmente denominado como “código espagueti”, una analogía a como se presenta el desorden en la pasta como alimento (Garrido, 2004).

Los lenguajes especializados para generar aplicaciones web son generalmente permisivos, esto crea una fácil integración con otro tipo de lenguajes, este tipo de característica permite separar una aplicación web completa en varios aspectos o puntos de vista del desarrollo. Desde el punto de vista de la interfaz e interacción existen lenguajes como HTML, CSS, y Javascript, estos forman parte del desarrollo en área de denominada Frontend que solo es visible por el cliente, lo que se denomina como capa de presentación, por otro lado, el área encargada de Bases de Datos permite que la información tenga permanencia en las aplicaciones, comúnmente se denomina como capa de persistencia o datos, y por último, el Backend de una aplicación, área encargada de conectar las bases de datos con el Frontend. La aplicación de este modelo, permite tener una arquitectura robusta, separada que ayuda a la corrección y organización del código que se genera y ejecuta. (Libros Web, 2009).

El proyecto que a continuación se describe, tuvo como objetivo desarrollar la arquitectura de programación y la codificación necesaria para la implementación de cada uno de los tres módulos base que contempla el sistema SIGAF (Sistema de Gestión Académica de la Facultad de Contaduría y Administración). Estos modulo son:

- Manejo de información de los planes de estudios de la facultad que resulte a partir de la creación y gestión de las unidades de aprendizaje asignadas en un período determinado.
- Realizar la carga académica de cada programa educativo del ciclo escolar vigente, basándose en los planes de estudio previamente establecidos por el personal administrativo.
- Capturar y actualizar la disponibilidad docente, formación académica e información relevante a la institución a lo largo de su trayectoria educativa en la facultad, la cual será esencial al momento de la asignación de docentes en la creación de horarios.

Una aplicación web es un programa informático, que es accedido y ejecutado en una red como internet o intranet, mediante una interfaz comúnmente por un navegador web (Alegsa, 2013).

Una aplicación web se define en el sitio Diccionario de informática (Alegsa, 2013) “En general, el término también se utiliza para designar aquellos programas informáticos que son ejecutados en el entorno del navegador (por ejemplo, un applet de Java) o codificado con algún lenguaje soportado por el navegador (como JavaScript, combinado con HTML); confiándose en el navegador web para que reproduzca (rende-ricé) la aplicación” (Herts, 2011).

Una de las ventajas de las aplicaciones web cargadas desde internet (u otra red), es la facilidad de mantener y actualizar, sin la necesidad de distribuir e instalar un software en potencialmente miles de clientes. También tiene como ventaja la posibilidad de ser ejecutadas en múltiples plataformas.

Las aplicaciones web a diferencia de otro tipo de programa tienen las características tal como el usuario puede acceder fácilmente a estas aplicaciones mediante un navegador web (cliente) o similar con acceso a internet se actualizar y mantiene una única aplicación y todos sus usuarios verán los resultados inmediatamente (Alegsa, 2013).

La interfaz gráfica de una aplicación web puede ser compleja y funcional, gracias a las variadas tecnologías web que existen: Java, JavaScript, DHTML, Flash, Silverlight, Ajax, entre otras (Lapuente, 2013). Ajax es un ejemplo de una tecnología de desarrollo web que le da gran poder de interactividad a las aplicaciones web (Microsoft, 2013).

2. Metodología

El desarrollo y gestión del proyecto SIGAF sigue la metodología de SCRUM, esta a su vez se gestiona mediante la herramienta RALLY que esta específicamente diseñada para trabajar sobre SCRUM. En esta herramienta gratuita online (también existen versiones de pago) cada rol del equipo registra su trabajo durante cada sprint, así como el registro de las tareas necesarias para completar cada historia del usuario definida en los sprint´s, Referencia (Cohn, 2014).

A continuación se listan las actividades generales del rol de desarrollo Backend de la aplicación (Rally Software, 2015):

- Investigación CVS (Git, Github).
- Investigación del framework LARAVEL para programación MVC.
- Codificación.
- Integración de interfaz y base de datos.
- Pruebas unitarias, depuración, verificación de requerimientos (Pruebas de Caja Blanca).

2.1. Metodología de desarrollo Backend

SCRUM es utilizado como línea general para el ciclo de vida del desarrollo de software, dentro de la metodología empleada para desarrollar el Backend de los módulos del sistema SIGAF, se enumeran las actividades generales para el desarrollo de forma secuencial y sistemática para la obtención del código usable, que mantendrá la arquitectura del sistema en cada módulo descrito por el análisis, para el funcionamiento del sistema en general, la metodología está enfocada a un estudio de caso, por lo que, está dirigida a comprender la dinámicas presente en un contexto singular según (Piedad, 2006). Esta particularidad hace que la metodología del estudio de caso investigado varíe acorde a la disciplina en este caso las ciencias computacionales. Para el desarrollo del sistema SIGAF se seguirá la pauta del modelo propuesto en el fundamento teórico del MVC o Modelo Vista Controlador.

Para optimizar la estructura de la metodología y mejorar la presentación del desarrollo, se divide en las siguientes fases los procedimientos y actividades a realizar en cada módulo del sistema SIGAF.

2.2. Fases de la metodología

2.2.1. Fase de preparación del entorno de desarrollo.

Contempla la preparación e instalación del entorno de desarrollo, esto se realiza solo una vez durante toda la vida del ciclo de desarrollo, por lo tanto, se ejecuta esta acción al inicio del proyecto, que consta de las siguientes actividades:

- Instalar un servidor web de prueba para peticiones http donde se harán las pruebas locales.
- Instalar el módulo de php para que interprete el código generado.
- Instalar el servicio de base de datos MYSQL para el almacenamiento y lectura de datos.
- Instalar el framework de desarrollo LARAVEL el cual tiene las siguientes dependencias: Composer como gestor de paquete, php como lenguaje base y un editor de código para la escritura, configuraciones extra como extensiones openssl para comunicación y la activación del módulo de reescritura de las url del servidor web para permitir direcciones amigables.
- Instalar un navegador web para pruebas.
- Probar la página de inicio del LARAVEL para ver que todo esté integrado y funcione correctamente.

2.2.2. Fase de análisis de insumos o entradas necesarias para la codificación.

Para codificar y unificar las interfaces de la aplicación con el modelo de negocio (Base de datos del sistema), en el Modelo-Vista-Controlador, es necesario contar con las siguientes fuentes o entradas:

- El análisis de la lógica de negocio o los requerimientos del sistema enunciados en el documento oficial de recopilación de requerimientos, así como los diagramas de proceso y casos de uso que modelan el sistema, recopilados durante la fase de análisis.
- Las interfaces o vistas de la aplicación, como es un sistema web se refiere a los documentos fuentes HTML que contienen la maquetación del diseño gráfico en plantillas funcionales para que los usuarios interactúen con el sistema.
- Los scripts del modelo de negocio (Base de datos), así como también los diagramas entidad relación que describen el comportamiento y las relaciones de las entidades dentro del contexto del sistema.
- Minutas de cambios y acuerdos realizados en las reuniones periódicas que describe la metodología SCRUM.
- Una vez obtenidos los insumos se procede al análisis de cada uno, para relacionarlos mediante el modelo descrito con anterioridad, esta actividad de recolección de insumos se realiza por cada módulo durante el proceso de desarrollo.

2.2.3. Fase de emparejamiento y cotejo de los insumos o entrada de proceso.

Las siguientes actividades aseguran que la información de los insumos mantenga una consistencia y coherencia para unificarlos en el modelo de programación:

- Verificar que los datos descritos en el análisis de requerimientos se cumplan tanto a nivel de modelo (base de datos) como a nivel de vista (interfaces), esto quiere decir, que se verifica si existe un tipo de dato en un requerimiento específico, cumpla con un campo en las entidades de la base de datos y que de igual manera sea provisto un control para el usuario en la interfaz de la vista.

- Verificar que las validaciones descritas en el análisis estén contempladas, en primer plano, durante la ejecución de acciones en la interfaz del usuario y en segundo plano, durante la integridad y mecanismos de control de la base de datos, para asegurar el almacenamiento de la información correctamente.
- Si existe alguna discrepancia, comunicar directamente con el encargado del área y resolver las posibles dudas, errores hasta obtener el resultado óptimo antes de empezar con la codificación.

2.2.4. Fase de codificación y desarrollo.

Durante cada módulo se seguirán las siguientes actividades para generar y mantener código sustentable y que tendrá como resultado la funcionalidad de la aplicación:

- Revisar los diagramas de secuencia y verificar mientras se codifica.
- Crear la estructura de organización de los archivos dependiendo el modulo.
- Escribir la parte lógica de la aplicación (controladores) y documentar, los cuales recibirán los datos y serán los encargados de procesar la información.
- Escribir en la parte del Frontend los resultados que arrojen los controladores.
- Escribir las funciones y documentar, estas funciones ejecutarán las acciones del lado del Frontend.
- Hacer pruebas de cada petición y ruta ejecutada por el usuario.
- Depurar la funcionalidad de la aplicación.
- Verificar funcionamiento con equipo de desarrollo y retroalimentar.
- Corregir y verificar de nuevo con equipo de desarrollo hasta que cumpla con el requerimiento aceptado.

2.2.5. Fase de publicación y control de versiones.

Una vez codificadas las funcionalidades de la aplicación se procedió con las publicaciones y fusiones de código con los otros miembros del equipo de desarrollo, esto con el fin de mantener actualizado en todo momento la última versión funcional del sistema, para lograr esto se seguirán las siguientes actividades:

- Cada vez que se logre la funcionalidad especificada en el análisis, se hace una confirmación al repositorio principal del proyecto a través del sistema de control de versiones GIT-GITHUB, esto con el fin de tener una organización a nivel de código del proyecto.
- Si el equipo de desarrollo sube actualizaciones en colaboración al repositorio principal del proyecto, se revisa, analiza y se aprueba para hacer una fusión respecto al código, esto provee una mejora, avance y actualización a la aplicación en desarrollo.
- Revisar periódicamente el repositorio principal para tener actualizada la aplicaciones tanto remotamente en un servidor en la nube como localmente en los equipos de cómputo.

2.3.6. Fase de retroalimentación para la metodología Scrum con usuarios finales y equipo de desarrollo.

Una vez hechas las confirmaciones al repositorio principal del proyecto, se realizan las siguientes actividades para dar seguimiento al proyecto principal:

- Reportar en la aplicación RALLY el avance de la codificación para retroalimentar al equipo.
- Revisar en juntas programadas el avance de la aplicación y los problemas que se presentan durante el avance del desarrollo del proyecto con el equipo de desarrollo y los usuarios del sistema (retroalimentación).
- Las fases descritas con anterioridad se desarrollaran por módulo hasta concluir el sistema, al final, la documentación de las funcionalidades describirá un API el cual podrá consultar el usuario para su posterior actualización o mejoras del sistema.

3. Desarrollo

Para la realización del código se utilizó un entorno de desarrollo adecuado a la arquitectura que se estableció al inicio del proyecto. Mediante el framework Laravel para PHP se puede utilizar el modelo-vista-controlador para centrar los esfuerzos solo en la codificación de la funcionalidad requerida separando apropiadamente las capas del sistema.

El framework Laravel es la médula espinal de la arquitectura MVC utilizando como lenguaje base PHP, su correcta implementación presupone agilidad, desarrollo ágil y código ordenado de acorde a SCRUM, por lo tanto su puesta en marcha y correcto funcionamiento es esencial para continuar con el desarrollo de la codificación.

Una vez analizado y cotejado los insumos se procede a codificar de la manera en que el patrón de la arquitectura indica, para esto es necesario tener en cuenta que se programan 3 capas en el modelo.

1. Los modelos: entidades del modelo relacional.
2. Los controladores: se encargan de manejar la lógica de programación.
3. Las vistas: La interfaz de usuario con que el usuario interactúa con el sistema.

El framework laravel permite crear la abstracción de las entidades de la base de datos, mediante el ORM (Object-Relational-Mapping) ELOQUENT, este permite convertir datos entre el modelo relacional(tablas, relaciones, tipo de datos, etc.) a un sistema de datos orientado a objetos, transfiere las entidades de la base de datos a clases para utilizarlas en un sistema de objetos omitiendo la parte de las conexiones y el SQL propios del lenguaje de consulta, que son necesarios para manipular las entidades del modelo.

La codificación de la inserción del número del Plan de Estudios se procede a crear los modelos en el framework lo que permitirá la interacción entre la aplicación y la base de datos.

Al término del desarrollo se tiene la estructura MVC en Laravel del módulo con la funcionalidad descrita en el análisis de Disponibilidad Docente y la funcionalidad de la interfaz interactuando con la base de datos.

4. Resultados

4.1. Resultados de Funcionalidad del Backend

Una vez desarrollado las fases de la metodología, y teniendo la retroalimentación del equipo de los problemas encontrados en la funcionalidad, el código en ejecución cumple con la

funcionalidad requerida descrita en el análisis, y el usuario final puede interactuar con los módulos mediante los controles de la interfaz, se anexa el código para este resultado.

El código generador provoca que a través de la interfaz el usuario pueda:

- Registrar un Plan de Estudios
- Crear unidades aprendizaje en los planes de estudio.
- Consultar los planes de estudio y las unidades aunadas ahí.
- Modificar los atributos de las unidades de aprendizaje.
- Eliminar de los planes de estudios una unidad de aprendizaje.
- Asignar unidades de aprendizaje a diferentes programas educativos.
- Agregar coordinaciones, números de plan de estudios, etapas y demás catálogos relacionados.
- Consultar las unidades de aprendizaje de un plan y visualizar en forma de listado cada una.
- Consultar los diferentes planes de estudios por carrera, etapa, unidad, tronco común y coordinación.

4.2. Resultados generales de la arquitectura

Durante el desarrollo de la aplicación se desarrolló un API para retornar información en formato JSON a una petición HTTP, esto con el fin de reutilizar la información en cualquier lenguaje o plataforma que no sea SIGAF.

El framewok Laravel mapea todas la tablas de la BD a clases para tener un entorno de desarrollo más limpio y no mezclar consultas sql en el código de producción (Gilmore, 2015). También al termino del desarrollo se concluye con una estructura clara y concisa que no permite visualizar los elementos de programación como fragmentos de código individuales, capaz de sufrir modificaciones sin alterar el flujo de la aplicación.

La interactividad de la aplicación fue desarrollada con Javascript y jQuery para solucionar de forma práctica la dinámica entre usuario-sistema. Al término de los módulos el sistema quedó con la siguiente estructura que permite toda la funcionalidad requerida en el análisis.

5. Conclusiones

Durante el desarrollo de la aplicación se construyó un API (Interfaz de Programación de Aplicaciones) para retornar información en formato JSON (Notación Objeto de JavaScript) a una petición HTTP mediante el framework Laravel y tecnologías como Javascript, HTML5, CSS3, esto brinda a futuros desarrollos del sistema SIGAF la posibilidad de crear diferentes aplicaciones para diversos fines, con el añadido de utilizar cualquier tecnología Frontend para consumir los datos. Esto es posible gracias al equipo completo de desarrollo y a las habilidades de cada miembro.

El uso del Modelo Vista Controlador (MVC) que se utilizo, permitió que la arquitectura de software que separa los datos de una aplicación, la interfaz de usuario, y la lógica de control en tres componentes distintos, se desarrollaran de forma independiente y muy coordinada intercambio la información necesaria para estar en completa armonía logrando el funcionamiento perfecto al articular cada sección.

Como oportunidad de aprendizaje es importante la integración del equipo de desarrollo, el presente trabajo solo forma una parte del sistema en su conjunto y es el resultado de la construcción de otras áreas como la creación de la base de datos, Frontend, por ello la comunicación entre los miembros del equipo es vital y necesaria, las metodologías describen un área de oportunidad muy grande para facilitar la creación de un sistema, sin embargo, el que ejecuta dicha metodología son personas por lo que la convivencia es necesaria entre los miembros del equipo, Cabe aclarar que durante el desarrollo de los módulos se presentaron cambios radicales por parte del cliente que repercutieron en el avance del mismo, sin embargo, el uso de la metodologia SCRUM permitio la comunicación e hizo posible solventar dichos cambios de forma adecuada.

Los módulos realizados cumplen con las expectativas del proyecto y son la base para la creación de horarios del sistema SIGAF.

Referencias

Alegsa. (2013, 12 5). *Alegsa*. Retrieved 11 16, 2010 from Aplicaciones Web: http://www.alegsa.com.ar/Dic/aplicacion%20web.php

Cohn, M. (2014, Febrero 24). *User Stories & Back-End Systems*. Retrieved Julio 8, 2014 from Scrum Alliance: https://www.scrumalliance.org/community/spotlight/mike-cohn/february-2014/user-stories-back-end-systems?feed=Scrum-Alliance-Spotlight

Garrido, J. S. (2004). Arquitectura y diseño de sistemas web modernos. *Revista de Ingeniería Informática del CIIRM*, 1-6.

Garrido, J. S. (2004, Julio 10). *Arquitectura y diseño de sistemas web modernos*. Recuperado el Marzo 20, 2014 de http://pegaso.ls.fi.upm.es/~sortega/html_css/files/Arquitectura_y_disenyo_de_sistemas_web_modernos.pdf

Gilmore, W. J. (2015). *EASY LARAVEL 5*. LEANPUB.

Hearts, C. V. (2012, Marzo 15). *Lenguajes de Programación*. From Maestros del Web: http://www.maestrosdelweb.com/editorial/los-diferentes-lenguajes-de-programacion-para-la-web-c106224I/

Herts, C. V. (2011, Febrero 8). *Maestros del web*. Recuperado Abril 5, 2014 de http://www.maestrosdelweb.com/editorial/frontend-backend-duplicando-programacion/

Lapuente, M. J. (8 de Diciembre de 2013). *Hipertexto*. Recuperado el 15 de Febrero de 2014, de La interfaz gráfica: http://www.hipertexto.info/documentos/interfaz.htm

Libros Web. (2009, Marzo 7). *Javascript Vanilla*. Retrieved Octubre 3, 2013 from http://librosweb.es/libro/jobeet_1_4/capitulo_4/la_arquitectura_mvc.html

McCarthy, P. (s.f.). Obtenido de AJAX for developers: Build dynamic java applications: http://www.ibm.com/developerworks/library/j-ajax1/

Microsoft. (2013, Diciembre 2). From ASP .net: Enhanced Interactivity and Responsiveness. Microsoft: http://www.asp.net/ajax

Narvaez, E., & Piedra, N. (2018) Un enfoque de Linked Data para garantizar la interoperabilidad semántica e integridad de datos académicos universitarios, Recuperado November 30, de http://ceur-ws.org/Vol-2096/paper5.pdf.

Paredes, M. A. (2011, Agosto 10). Arquitectura en Capas. RecuperadoFebrero 20, 2014 de http://arquitecturaencapas.blogspot.mx/2011/08/arquitectura-3-capas-programacion-por.html

Piedad Cristina Martínez Carazo. (2006, Marzo 12). *El método de estudio de caso*. Recuperado Septiembre 15, 2014 de http://ciruelo.uninorte.edu.co/pdf/pensamiento_gestion/20/5_El_metodo_de_estudio_de_caso.pdf

Rally Software. (2015, Enero 22). From https://www.rallydev.com/

Saquicela, V., Orellana, G., Piedra, N., Orellana, M., & Espinoza, M. (2018) Similarity Detection among Academic Contents through Semantic Technologies and Text Mining. Recuperado Noviembre 30, de http://ceur-ws.org/Vol-2096/paper1.pdf.

Valdés, D. P. (2007, Noviembre 2). *Los diferentes lenguajes de programación para la web*. Recuperado Septiembre 1, 2013 de MAESTROS DEL WEB: http://www.maestrosdelweb.com/los-diferentes-lenguajes-de-programacion-para-la-web/

Vega, F. (5 de 12 de 2013). *Backend Profesional*. Recuperado de https://cursos.mejorando.la/cursos/backend-online/material/un-backend-profesional/

Vega, F. (5 de 12 de 2013). *Cristalab*. Recuperado de http://www.cristalab.com/blog/que-significa-backend-y-frontend-en-el-diseno-web-c106224l/

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 03/10/2018
Aceitação/Acceptance: 13/12/2018

Procesamiento de las opiniones de un personaje público en el Ecuador

Boris Enrique Herrera Flores[1]

bherrera@uce.edu.ec

[1] Universidad Central del Ecuador, Campus Universitario, 170521, Quito, Ecuador.

Pages: 1094–1102

Resumen: La presente investigación es el estudio de las técnicas de minería de opiniones, enfocada a obtener información de un personaje público en el Ecuador, determinando indicios de polaridad a su gestión en forma positiva, negativa o neutra, resultado que le permitirá a dicho personaje público tomar decisiones sobre su accionar en función de una imagen de servicio a la comunidad. La extracción de las opiniones en redes sociales y técnicas basadas en Tecnologías del Lenguaje Humano posibilitó la interpretación de los datos polarizados precisando parámetros de relevancia a la opinión resultante enfocados a la toma de decisiones, procesamiento que se adapta a los nuevos formatos de comunicación logrando la interpretación y valoración de la opinión. Las redes sociales fue la plataforma para la captura de textos por medio de un API, que luego del procesamiento del lenguaje natural se obtuvieron resultados de indicios de la popularidad del personaje.

Palabras-clave: Procesamiento del lenguaje natural, minería de opiniones, análisis de sentimientos, tecnologías del lenguaje humano, clasificación de opiniones.

Processing of the opinions of a public person in Ecuador

Abstract: The present investigation is the study of opinion mining techniques, focused on obtaining information from a public figure in Ecuador, determining signs of polarity for your management in a positive, negative or neutral way, a result that will allow said character public to make decisions about their actions based on an image of service to the community. The extraction of opinions in social networks and techniques based on Human Language Technologies enabled the interpretation of polarized data by specifying parameters of relevance to the resulting opinion focused on decision making, processing that adapts to the new communication formats achieving the interpretation and assessment of opinion. Social networks was the platform for the capture of texts by means of an API, which after the processing of the natural language obtained results of indications of the popularity of the character.

Keywords*:* Natural language processing, opinion mining, sentiment analysis, human language technologies, rating of opinions.

1. Introducción

Actualmente la red social es la plataforma Web 2.0 para que los usuarios expresen sus opiniones, puntos de vista en forma inmediata con una participación frecuente usando un lenguaje informal, identifico la necesidad de procesar estos párrafos, oraciones y palabras precisando grados de relevancia a la informalidad al definir polaridad positiva, neutra o negativa, usando las Tecnologías del Lenguaje Humano.

Este estudio se enfoca a la comprensión de las expresiones subjetivas y el lenguaje informal usado por el Ecuatoriano al opinar sobre un personaje público a su accionar de servicio a la comunidad, donde las actuales tecnologías de lenguaje humano no son directamente aplicables y es necesario incorporar métodos y herramientas para el procesamiento del lenguaje natural por medio de la experimentación de diferentes técnicas y sistemas genéricos de análisis de sentimientos que extraigan primero la fuente de información para luego dar un tratamiento con métodos que determinen la polaridad de la opinión.

El objetivo es la mejora de los recursos, técnicas y herramientas que modelan el lenguaje subjetivo e informal que genera la red social Twitter, con el tratamiento del lenguaje emocional en un entorno de subsistemas inteligentes de procesamiento para la recuperación, tratamiento, compresión y descubrimiento de la información apto para la toma de decisiones.

2. Bases Conceptuales

Twitter, es una aplicación web gratuita de microblogging que recibe un estimado de 313 millones de visitas al mes, que reúne las ventajas de los blogs, las redes sociales y la mensajería instantánea, esta nueva forma de comunicación permite a sus usuarios estar en contacto en tiempo real con personas de su interés a través de mensajes breves de texto a los que se denominan Tweets, por medio de una sencilla pregunta: ¿Qué estás haciendo?. Los usuarios envían y reciben Tweets de otros usuarios a través de breves mensajes que no deben superar los 140 caracteres, vía web, teléfono móvil, mensajería instantánea o a través del correo electrónico; e incluso desde terceras aplicaciones.

Una API de Twitter (siglas en inglés “Application Programming Interface”), es un conjunto de reglas y especificaciones que las aplicaciones pueden seguir para comunicarse entre ellas, sirviendo de interfaz entre programas diferentes de la misma manera que la interfaz de usuario facilita la interacción humano-software. Twitter tiene tres tipos de APIs: REST API, Search API y Streaming API:

- REST API, ofrece a los desarrolladores el acceso al core de los datos de Twitter. Todas las operaciones que se pueden hacer vía web son posibles realizarlas desde el API, soporta los formatos: xml, json, rss, atom.
- Search API, suministra los tweets con una profundidad en el tiempo de 7 días que se ajustan a la consulta demandada. Es posible filtrar por, cliente utilizando, lenguaje y localización, no requiere autenticación y los tweets se obtienen en formato json o atom.
- Streaming API, establece una conexión permanente por usuario con los servidores de Twitter y mediante una petición http se recibe un flujo continuo de tweets en formato json.

Con Search API y en el REST API existe una limitación de 150 peticiones a la hora por usuario o por IP si la llamada no estuvo autenticada.

El análisis de sentimientos, de textos en las redes sociales es el proceso que determina el tono emocional que hay detrás de una palabra determinada, si una frase contiene una opinión positiva, neutra, o negativa sobre un producto, marca, institución, organización, empresa, evento o persona, el objetivo es extraer aquellos términos semánticos que expresen un sentimiento en particular para conocer la opinión, las actitudes y las expectativas sobre un tema en concreto así como para analizar el comportamiento de los usuarios ante algún mensaje y, por tanto, determinar su impacto o poder anticipar su reacción.

La determinación de la polaridad, consistente en determinar cuándo una opinión es positiva, negativa o neutra con respecto a la entidad a la cual se está refiriendo desde dos enfoques diferentes dependiendo del tipo de método que se utiliza, como son:

- Método de clasificación supervisada: Los rasgos extraídos del texto y el método de aprendizaje determinan cuando este pertenece a la clase positiva o negativa.
- Método de clasificación no supervisada (clustering): Tienen en cuenta la presencia de palabras con orientaciones conocidas que son obtenidas de diccionarios o corpus como, por ejemplo, las palabras excelente o alegre, que son representativas de expresiones polares positivas.

Existen varias técnicas al momento de estimar un texto y su posible polaridad positiva, negativa o neutra. Al utilizar software en la minería de opiniones se encuentran dos funciones: classify polarity y score.sentiment

classify polarity, utiliza un marco probabilístico, tiene sus fundamentos en el Teorema de Bayes para calcular la probabilidad de una clase (positiva, negativa). Es una técnica de clasificación y predicción supervisada, entre sus ventajas está que su implementación es muy fácil y obtiene buenos resultados de clasificación en la mayoría de los casos:

$$p(CjF1; :::; Fn)$$

Donde C representa un valor positivo (+) o negativo (-) y F1; F2; :::; Fn factores que representa las palabras de un tweet, donde se busca establecer la probabilidad a priori de que un tweet sea positivo o negativo.

score.sentiment, es una función muy simple de predicción no supervisada que asigna una puntuación simplemente contando el número de ocurrencias de positivos y negativos en un tweet comparando con un diccionario de palabras conocido como corpus.

$$score = sum(PalabrasPositivas) - sum(PalabrasNegativas)$$

El éxito de esta función es tener la mayor cantidad de palabras positivas y negativas, existen algunos diccionarios ya cargados con estas palabras uno de ellos es MPQA, LIWC.

3. Metodología

La metodología a utilizar es SEMMA creada por SAS Institute, que fue propuesta especialmente para trabajar con el software SAS Enterprise Miner. Si bien en la comunidad científica se conoce a SEMMA como una metodología, en el sitio de la empresa SAS se aclara que este no es el objetivo de la misma, sino más bien la propuesta de una organización lógica de las tareas más importantes del proceso de minería de datos. SEMMA establece un conjunto de cinco fases para llevar a cabo el proceso de minería:

Sample (Muestreo), Explore (Exploración), Modify (Modificación), Model (Modelado) y Assess (Evaluación).

Para el esquema general al detectar la polaridad de opiniones, se debe tomar en cuenta las palabras considerando su sentido correcto en un contexto determinado:

- Identificar Tokens: En esta primera etapa se realiza un pre- procesamiento del texto de las opiniones y devuelve los términos que aportan información útil, o definido como limpieza de los datos como por ejemplo duplicados generados por trolls
- Desambiguar lexicalmente cada token: En la segunda etapa se parte de cada término que aporta información útil, este le da forma, lematiza y desambiguan lexicalmente.
- Obtener todas las acepciones de cada palabra: Esta etapa parte del listado de términos lematizados y desambiguados y devuelve todas las acepciones del término en el idioma en que se esté realizando el análisis.
- Clasificar cada token en positivo y negativo: Una vez obtenidas todas las acepciones de cada término, en esta etapa se propone determinar la polaridad del término de acuerdo a los algoritmos que se empleen, finalmente la polaridad de la opinión se determina por el número de palabras positivas y negativas que contiene.
- Evaluar la opinión: Si el número de palabras positivas que contiene una frase es mayor se considera la frase como positiva y en caso contrario es negativa.

Los índices de relevancia identificados en twitter son:

- Ratio seguidores/seguidos: Los usuarios que tienen un ratio de seguidores/ seguidos cercano al 1 suelen contar con el efecto followback: muchas cuentas les siguen para ganar seguidores fácilmente. Para encontrar su relevancia usaremos una métrica basándonos en el ratio de seguidores/seguidos, pero esta será solo una parte, la parte fundamental para encontrar la relevancia se encuentra en el número de retweets que ha tenido el tweet asignándole un factor de multiplicidad de 6.5 al valor original, así mismo al campo favoritos que indica cuantas veces el tweet ha sido marcado como favorito, se le asignó un factor de multiplicidad de 3.5.
- Retweets/nº seguidores: Los retweets son un indicador universal de que lo que dices es interesante o digno de atención. Cuantos más retweets tengas, más interesante o relevante es lo que dices o quien eres. El número de tweets es un factor importante, ya que no es lo mismo conseguir 100 retweets con un solo tweet que hacerlo a lo largo de un mes con 300.

El software para la extracción de datos utilizada es #TAGS, una aplicación que funciona con Google Docs, que permite entre otras cosas conocer los tweets y las conversaciones

que ha tenido un usuario determinado ofreciendo características adicionales del tweet como ubicación geográfica, id usuario, seguidores, favorito, retweets, etc.

Del estudio realizado aplicando la metodología indicada anteriormente, el personaje público obtuvo 71.845 tweets en el periodo.

4. Resultado Experimentales

Del estudio realizado de 71.845 tweets usando el software de JMP de SAS se aplicó la limpieza a los datos, eliminando links en el texto, tweets vacios, duplicados aplicando la fórmula de la Figura 1

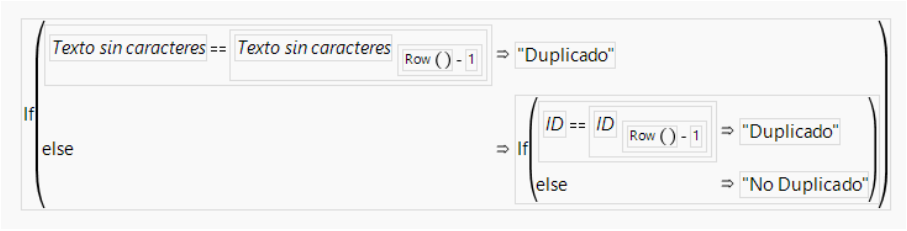


Figura 1 - Regla de JMP para limpieza de datos.

La polaridad se ensayó con la siguiente definición de un corpus básico de polaridad en la Figura 2:

No	POSITIVO	NEGATIVO
1.	Mejor	Ratas
2.	Felicitaciones	Ladrón
3.	Felicidades	Mentira
4.	Compañeros	Malo
5.	Bien	Peor
6.	Bienvenido	Rechazo
7.	Confianza	Nunca
8.	Confiable	Vago
9.	Respaldo	Lame botas
10.	Bueno	Tirano
11.	Solidario	Corrupto
12.	Hermoso	Borrego
13.	Lindo	Chavista
14.	Éxito	Traición
15.	Ético	Traidor
16.	Aprobado	Inepto
17.	Adelante	Decepción
18.	Ganador	Perdedor

Figura 2 – Corpus básico de polaridad.

Se determina si es positivo con el valor de 1, neutro con el valor de 0 y negativo con el valor de -1.

Las siguientes son las métricas utilizadas para el estudio:

Número de seguidores, A esta métrica la hemos definido de tal manera que si nuestro personaje público tiene más de 200 seguidores en cada uno de sus tweets es aceptado por parte de los twitteros en la red social.

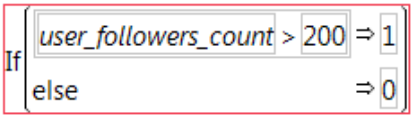


Figura 3 – Regla de JMP deNúmero de seguidores.

Aceptabilidad, esta métrica nos da a conocer que tan aceptado es el mensaje compartido por nuestro personaje por los twitteros en la red social.

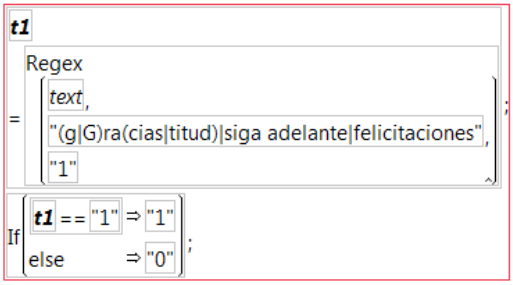


Figura 4 – Regla de JMP de Aceptabilidad.

Ratio de iteraciones /fans, esta fórmula nos permite determinar cuan relevante es una cuenta en relación al número de fans.

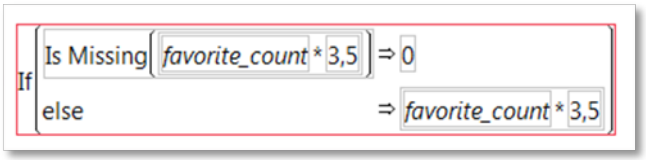


Figura 5 – Regla de JMP de Valoración de favoritos.

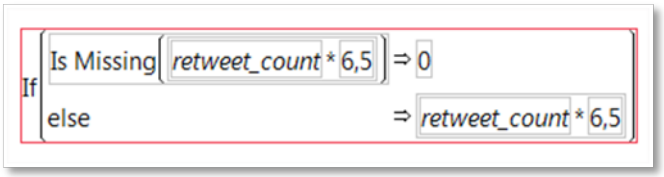
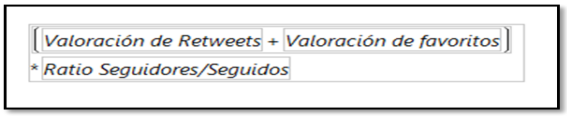


Figura 6 – Regla de JMP de Valoración de retweets.

Figura 7 - Regla de JMP de Ratio seguidores/seguidos.



Valoración de favoritos	Valoración de Retweets	Ratio Seguidores/Seguidos	Relevancia
0	0	0,952681388	0
0	0	0,952681388	0
0	6,5	11,757731959	76,425257732
84	149,5	124,77484183	29134,925568
0	3425,5	0,5596868885	1917,2074364
24,5	39	19,728401192	1252,7534757
0	0	111,77002967	0
28	208	17,157360406	4049,1370558
0	0	17,756722151	0
0	26	17,698209719	460,15345269
0	0	13,534482759	0
73,5	143	124,77484183	27013,753256
0	0	6,8653846154	0
0	962	0,023255814	22,372093023
0	6,5	2,1091674765	13,709588597
3,5	13	2,1091674765	34,801263362
0	253,5	2,1080467229	534,38984426
0	6,5	2,1080467229	13,702303699
0	0	2,1102106969	0
24,5	52	4,6952736318	359,18843284
0	97,5	4,6514143095	453,51289517
0	188,5	4,6514143095	876,79159734
38,5	84,5	4,6952736318	577,51865672
0	97,5	4,6514143095	453,51289517
7	13	4,6514143095	93,02828619
35	58,5	4,6952736318	439,00808458
0	110,5	4,6514143095	513,9812812
0	0	4,6514143095	0
10,5	39	2,1091674765	104,40379009
0	182	2,1080467229	383,66450357
0	260	2,1080467229	548,00214796

Figura 8 – Regla de JMP de Relevancia y resultados

En lo referente a la polaridad del ejercicio se pude identificar que:

3% Negativo

92% Neutro

4% Positivo

Las opiniones sobre el personaje público son en su mayoría neutras, definiendo que su accionar tampoco en negativo pero que tampoco da mucho que celebrar con opiniones positivas.

Al aplicar el score.sentiment, que es la función de predicción no supervisada se obtiene una puntación positiva en los tweets, esto es comparando con un diccionario de palabras corpus.

Al establecer las métricas de relevancia a las polaridades positiva y neutral, se puede identificar lo siguiente:

14% Nada relevante

84% Poco relevante

0% Algo relevante

2% Normalmente relevante

Por medio del conteo de la métrica de relevancia, la información del cuadro indica que las opiniones sobre el personaje público fueron de poca relevancia, por los bajos niveles de retweets, y de poca marcación en favoritos.

5. Conclusiones

Existen retos en el procesamiento del lenguaje natural que deben ser resueltos en la detección automática de la polaridad de un texto, esto es el corpus en español que entienda el léxico ecuatoriano o latinoamericano, este trabajo de investigación lo detecta e identifica una línea de estudio que puede ser trabajada en la comunidad de investigadores.

La definición de polaridad de sentimientos dentro de este estudio de opinión de un personaje público, determina indicios o tendencias al tipo de impacto, no tiene margen de error como una encuesta tradicional pero si provee un resultado que ayuda a reforzar la toma de decisiones.

La contribución del usuario en el lenguaje informal es determinante en una temporalidad cual puede variar en opinión, dichos textos deben ser preprocesados para obtener una data que aporte a un resultado relevante.

Los niveles de retweets y marcación en favoritos determinan la relevancia de una opinión, métrica utilizada para el estudio de una opinión en twitter, pueden ser también identificadas otras como ubicación geográfica, número de seguidores, etc.

El trabajo futuro es trabajar en otros dominios como el educativo que permitiría identificar el manejo del contenido por el Docente, el seguimiento a las unidades o temas planteados y la relación interpersonal Docente Estudiante.

Referencias

Borras-Morell, Jose Enrique (2015) "Data Mining for Pulsing the Emotion on the Web" Colección: Methods in Molecular Biology Volumen: 1246 pp: 123-130

Chen, Xin; Vorvoreanu, Mihaela; Madhavan, Krishna(2014) "Mining Social Media Data for Understanding Students' Learning Experiences" IEEE TRANSACTIONS ON LEARNING TECHNOLOGIES Volumen: 7 Número: 3 pp: 246-259

Khan, Farhan Hassan; Bashir, Saba; Qamar, Usman (2014) "TOM: Twitter opinion mining framework using hybrid classification scheme" DECISION SUPPORT SYSTEMS Volumen: 57 pp: 245-257

Akaichi, Jalel; Dhouioui, Zeineb; Lopez-Huertas Perez, Maria Jose (2013) "Text Mining Facebook Status Updates for Sentiment Classification" Conferencia: 17th International Conference System Theory, Control and Computing (ICSTCC) Ubicación: Sinaia, ROMANIA pp: 640-645

Colace, F.; De Santo, M.; Greco, L. (2014) "SAFE: A Sentiment Analysis Framework for E-Learning"International Journal of Emerging Technologies in Learning Volumen: 9 Número: 6 pp: 37-41

Colace, Francesco; Casaburi, Luca; De Santo, Massimo (2015) "Sentiment detection in social networks and in collaborative learning environments" COMPUTERS IN HUMAN BEHAVIOR Volumen: 51 Número especial: SI pp: 1061-1067

Baldominos Gomez, Alejandro; Luis Mingueza, Nerea; Cristina Garcia del Pozo, Ma (2015) "OpinAIS: An Artificial Immune System-based Framework for Opinion Mining" INTERNATIONAL JOURNAL OF INTERACTIVE MULTIMEDIA AND ARTIFICIAL INTELLIGENCE Volumen: 3 Número: 3 pp: 25-34

Critérios Editoriais

A RISTI (Revista Ibérica de Sistemas e Tecnologias de Informação) é um periódico científico, propriedade da AISTI (Associação Ibérica de Sistemas e Tecnologias de Informação), que foca a investigação e a aplicação prática inovadora no domínio dos sistemas e tecnologias de informação.

O Conselho Editorial da RISTI incentiva potenciais autores a submeterem artigos originais e inovadores para avaliação pelo Conselho Científico.

A submissão de artigos para publicação na RISTI deve realizar-se de acordo com as chamadas de artigos e as instruções e normas disponibilizadas no sítio Web da revista (http://www.risti.xyz).

Todos os artigos submetidos são avaliados por um conjunto de membros do Conselho Científico, não inferior a três elementos.

Em cada número da revista são publicados entre cinco a oito dos melhores artigos submetidos.

Criterios Editoriales

La RISTI (Revista Ibérica de Sistemas y Tecnologías de la Información) es un periódico científico, propiedad de la AISTI (Asociación Ibérica de Sistemas y Tecnologías de la Información), centrado en la investigación y en la aplicación práctica innovadora en el dominio de los sistemas y tecnologías de la información.

El Consejo Editorial de la RISTI incentiva autores potenciales a enviar sus artículos originales e innovadores para evaluación por el Consejo Científico.

Lo envío de artículos para publicación en la RISTI debe hacerse de conformidad con las llamadas de los artículos y las instrucciones y normas establecidas en el sitio Web de la revista (http://www.risti.xyz).

Todos los trabajos enviados son evaluados por un número de miembros del Consejo Científico de no menos de tres elementos.

En cada número de la revista se publican cinco a ocho de los mejores artículos enviados.

Os associados da AISTI recebem a RISTI gratuitamente, por correio postal. Torne-se associado da AISTI. Preencha o formulário abaixo e envie-o para o e-mail aistic@gmail.com

Los asociados de la AISTI reciben la RISTI por correo, sin costo alguno. Hazte miembro de la AISTI. Rellena el siguiente formulario y remítelo al e-mail aistic@gmail.com

Formulário de Associado / Formulario de Asociado

Nome/Nombre: _____

Instituição/Institución: _____

Departamento: _____

Morada/Dirección: _____

Código Postal: _____ Localidade/Localidad: _____

País: _____

Telefone/Teléfono: _____

E-mail: _____ Web: _____

Tipo de Associado e valor da anuidade:

☐ Individual - 25€

☐ Instituição de Ensino ou I&D/Institución de Educación o I&D - 250€

☐ Outro (Empresa, etc.) - 500€

NIF/CIF: _____

Data/Fecha: __/__/_____ Assinatura/Firma: _____

risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies
©AISTI 2019 http://www.aisti.eu