



ISSN: 1646-9895

Revista Ibérica de Sistemas e Tecnologias de Informação
Revista Ibérica de Sistemas y Tecnologías de Información

M a r ç o 2 3 • M a r c h 2 3



©RISTI 2023 <http://www.risti.xyz>

Nº 49

Edição / Edición

Nº 49, 03/2023

Tiragem / Tirage: 1000

Preço por número / Precio por número: 17,5€

Subscrição anual / Suscripción anual: 30€ (2 números)

ISSN: 1646-9895

Depósito legal:

Indexação / Indexación

Academic Journals Database, CiteFactor, Dialnet, DOAJ, DOI, EBSCO, GALE, IndexCopernicus, Index of Information Systems Journals, ISI Web of Knowledge, Latindex, ProQuest, QUALIS, SciELO, SCImago, Scopus, SIS, Ulrich's.

Publicação / Publicación

RISTI – Revista Ibérica de Sistemas e Tecnologias de Informação

Rua Pedro Álvares Cabral 520 - 2º Esq. Nasc., 4435-386 Rio Tinto, Portugal

E-mail: aistic@gmail.com

Web: <http://www.risti.xyz>

Director

Álvaro Rocha, ISEG, Universidade de Lisboa, PT

Coordenadores da Edição / Coordinadores de la Edición

Jezreel Mejía, Centro de Investigación en Matemáticas - Unidad Zacatecas, MX

Yadira Quiñonez, Universidad Autónoma de Sinaloa, MX

Giner Alor-Hernández, Tecnológico Nacional de México, MX

Abraham Dávila, Pontificia Universidad Católica del Perú, PE

Conselho Editorial / Consejo Editorial

A. Patrícia Oliveira Universidade de Aveiro, PT

Abel Méndez, Instituto Tecnológico de Costa Rica, CR

Alejandro Peña, Universidad EAFIT, CO

Alma Gómez-Rodríguez, Universidade de Vigo, ES

Ana Melro, Universidade de Aveiro, PT

Ana Paula Afonso, P.PORTO/ISCAP.IPP/CEOS.PP, PT

Ana Rita Calvão, ESTGA, Universidad de Aveiro, PT

Ania Cravero, Universidad de La Frontera, CL

António Abreu Silva, ISCAP, Instituto Politécnico do Porto, PT

Antonio Garcia, Universidad de Santiago de Compostela, ES

António Godinho, ISEC, Instituto Politécnico de Coimbra, PT

Antonio Jiménez-Martín, Universidad Politécnica de Madrid, ES

Arturo J. Méndez, Universidad de Vigo, ES

August Climent, Ramon LLull University, ES

Beatriz Rodríguez, Universidad de la Republica, UY

Borja Bordel, Universidad Politécnica de Madrid, ES

Brenda L. Flores-Rios, Universidad Autónoma de Baja California, MX

Bruno Miguel Ferreira Gonçalves, Instituto Politécnico de Bragança, PT

Carlos Alexandre Silva, Instituto Federal de Minas Gerais, BR

Carlos Carreto, Instituto Politécnico da Guarda, PT

Carlos Morais, Instituto Politécnico de Bragança, PT

Carlos Rompante Cunha. UNIAG & CeDRI & Instituto Politécnico Bragança, PT

Carlos Vaz de Carvalho, Instituto Politécnico do Porto, PT

Ciro Martins, Universidade do Algarve, PT

Cristina Caridade, Instituto Politécnico de Coimbra, PT
Daniel Polónia, Universidade de Aveiro, PT
Dante Carrizo, Universidad de Atacama, CL
Edwin Cedeño Herrera, Universidad de Panamá, PA
Fábio Marques, ESTGA, Universidade de Aveiro, PT
Felipe Vasquez, Universidad de La Frontera, CL
Fernando Moreira, REMIT, IJP, Universidade Portucalense &
IEETA, Universidade de Aveiro, PT
Fernando Ribeiro, Instituto Politécnico de Castelo Branco, PT
Fernandp Bandeira, PT
Filipe Cardoso, Instituto Politécnico de Viseu, PT
Flor Gomes de María Sánchez Aguirre, Universidad César Vallejo, PE
Gabriel Guerrero-Contreras, Universidad de Cádiz, ES
Gerardo González Filgueira, Universidade da Coruña, ES
Gloria Valencia, Universidad de las Fuerzas Armadas, EC
Guilllhermo Davila, Universidad de Lima, PE
Hélder Gomes, Universidade de Aveiro, PT
Helder Zagalo, Universidade de Aveiro, PT
Hélia Guerra, Universidade dos Açores, PT
Henrique Gil, Instituto Politécnico de Castelo Branco, PT
Henrique S. Mamede, Universidade Aberta, Portugal
Inês Domingues, ISEC, Instituto Politécnico de Coimbra, PT
Isaias Bianchi, Universidade Federal de Santa Catarina, BR
Isidro Navarro, Universidad Politécnica de Cataluña, ES
Ismael Etxeberria-Agiriano, Universidad del País Vasco, ES
Ivan Garcia, Universidad Tecnológica de la Mixteca, MX
Jeimy Cano, Universidad de los Andes, CO
João Balsa, LASIGE, Universidade de Lisboa, Portugal
João Emílio Almeida, ISTE - Porto, PT
João Paulo Ferreira, Instituto Politécnico de Coimbra, PT
Joao Roberto de Toledo Quadros, CEFET/RJ, BR
Joao Tavares, Universidade do Porto, PT
Joaquim Reis, ISCTE - Instituto Universitário de Lisboa, PT
Jorge Eduardo Ibarra Esque, Universidad Autónoma de Baja California, MX

Jorge Hochstetter, Universidad de La Frontera, CL
José Álvarez-García, Universidad de Extremadura, ES
Jose Felipe Cocon Juarez, Universidad Autónoma del Carmen, MX
José Lousado, Instituto Politécnico de Viseu, PT
José Luis Pastrana Brincones, Universidad de Málaga, ES
Jose M. Molina, Universidad Carlos III de Madrid, ES
José Ribeiro, Politécnico de Leiria, PT
Jose Silvestre Silva, Academia Militar, PT
Josep Maria Marco-Simó, Universitat Oberta de Catalunya, ES
Juan Angel Contreras Vas, Universidad de Extremadura, ES
Juan M. Santos-Gago, Universidad de Vigo, ES
Juan Pablo DAmato, Universidad Nacional del Centro de
la Provincia de Buenos Aires, AR
Leila Weitzel, Universidade Federal Fluminense, BR
Leonardo Bermón Angarita, Universidad Nacional de Colombia, CO
Lilia Muñoz, Universidad Tecnológica de Panamá, PA
Lucila Romero, Universidad Nacional del Litoral, AR
Luis Álvarez Sabucedo, Universidad de Vigo, ES
Luís Bruno, Instituto Politécnico de Beja, PT
Luis Enrique Sánchez Crespo, Universidad de Castilla-la Mancha, ES
Luisa María Romero-Moreno, Universidad Sevilla, ES
Luz María Hernández Cruz, Universidad Autónoma de Campeche, MX
Luz Sussy Bayona Oré, Universidad Nacional Mayor de San Carlos, PE
Marcelo Marciszack, Universidad Tecnológica Nacional, AR
Marcelo Zambrano Vizuite, Universidad Tecnica Del Norte
Marco Painho, NOVA IMS, PT
Margarita Ramirez Ramirez, Universidad Autónoma de Baja California, MX
Maria Cristina Marcelino Bento, UNIFATEA, BR
María de la Cruz del Río-Rama, Universidad de Vigo, ES
Maria de los Milagros Gutierrez, Universidad Tecnológica Nacional, AR
Maria do Rosário Bernardo, Universidade Aberta, PT
Maria Hallo, Escuela Politécnica Nacional, EC
Maria João Ferreira, Universidade Portucalense, PT

Maria João Gomes, Universidade do Minho, PT
Maria Sousa, ISCTE - Instituto Universitário de Lisboa, PT
Marisol B. Correia, ESGHT - Universidade do Algarve & CiTUR, PT
Maristela Holanda, Universidade de Brasília, BR
Martin Llamas Nistal, Universidad de Vigo, ES
Miguel Casquilho, Universidade de Lisboa, PT
Miguel Ramón González Castro, Centro Tecnológico Aimen, ES
Mirna Ariadna Muñoz Mata, CIMAT, MX
Nelson Rocha, Universidade de Aveiro, PT
Nuno Melão, Instituto Politécnico de Viseu, PT
Nuno Ribeiro, Universidade Fernando Pessoa, Portugal
Patricia Dias, Universidade do Estado de Minas Gerais, BR
Paula Prata, Universidade da Beira Interior, PT
Paulo Martins, Universidade de Trás-os-Montes e Alto Douro, PT
Paulo Pinto, Universidade Nova de Lisboa, PT
Paulo Rurato, Universidade Fernando Pessoa, Portugal
Pedro Araújo, Universidade da Beira Interior, PT
Pedro R. Palos- Sanchez, Universidad de Sevilla, ES
Pedro Sanz-Angulo, Universidad de Valladolid, ES
Pedro Sobral, Universidade Fernando Pessoa, Portugal
Pedro Sousa, Universidade do Minho, Portugal
René Faruk Garzozi-Pincay, Universidad Estatal Península de Santa Elena, EC
Rita Oliveira, DigiMedia, Universidade de Aveiro, PT
Rita Santos ESTGA/DigiMedia, Universidade de Aveiro, Portugal
Ruben Pereira, ISCTE – Instituto Universitário de Lisboa, PT
Rui Pedro Marques, Universidade de Aveiro, PT
Rui Silva Moreira, Universidade Fernando Pessoa, PT
Samuel Sepúlveda, Universidad de La Frontera, CL
Santiago Raul Gonzales Sanches, Universidad Cesar Vallejo, PE
Sara Balderas-Díaz, Universidad de Cádiz, ES
Sergio Araya Guzmán, Universidad del Bío-Bío, CL
Sergio F. Lopes, Universidade do Minho, PT
Sergio Gálvez Rojas, Universidad de Málaga, ES

Solange N Alves-Souza, Universidade de São Paulo, BR
Telmo Silva, DigiMedia, Universidade de Aveiro, PT
Teresa Guarda, Universidad Estatal Península de Santa Elena, EC
Thiago Dias, Centro Federal de Educação Tecnológica de Minas Gerais, BR
Vitor Carvalho, Instituto Politécnico do Cávado e do Ave, PT
Vitor Santos, Universidade Nova de Lisboa - NOVA IMS, PT
Wagner Tanaka Botelho, Universidade Federal do ABC, BR
Yamid Hernández Julio, Universidad del Sinú, CO

Índice / Index

EDITORIAL

Nuevos desafíos en Tecnologías de Información y Comunicación.....	1
<i>Jezreel Mejía, Yadira Quiñonez, Giner Alor-Hernández, Abraham Dávila</i>	

ARTIGOS / ARTICULOS / ARTICLES

Gamificación para atender los desafíos de la enseñanza Ingeniería de Software en instituciones de educación superior	5
<i>Mirna Muñoz, Gloria Piedad Gasca-Hurtado</i>	

Diseño de videojuegos para el análisis de habilidades personales	22
<i>Adriana Peña Pérez Negrón, David Bonilla Carranza, Mirna Muñoz, Raúl Aguilar</i>	

Automatización de los procesos del estándar ISO/IEC 29110 a través de la Adopción de DevOps	37
<i>Josefina García, J Jesús Minero, Elvia Lara</i>	

Propuesta de mejoras para la implementación del estándar ISO/IEC 29110 utilizando lógica matemática.....	52
<i>Victor Terron, Jezreel Mejía</i>	

Aplicación móvil para monitoreo del consumo de materia seca en rumiantes usando la tecnología bluetooth	66
<i>Jesús Sierra Martínez, Juan Carlos Elizondo Leal, Daniel López Aguirre, Yadira Quiñonez, José Hugo Barrón Zambrano, Alan Díaz Manríquez, Vicente Paul Saldívar Alonso, José Ramon Martínez Angulo</i>	

Determinación del mejor algoritmo de detección de matrículas en ambientes controlados y no controlados	83
<i>Elias Ccoto Hualpa, Abel Angel Sullon Macalupu, Jorge Eddy Otazu Luque, Jorge Sánchez-Garces</i>	

Análisis automático de micrografías SEM mediante aprendizaje profundo.....	100
<i>Iker Toscano, Diego Alejandro Morales Bravo, Miguel De-la-Torre, Brenda Acevedo Juárez, Gabriel Alberto García Mireles</i>	

Características de las Bases de Datos para el Cómputo en la Nube: Un Estudio Secundario	115
<i>Raúl A. Aguilar Vera, Naomi García Sánchez, Julio C. Díaz Mendoza, Brenda L. Flores Ríos</i>	

Tendencias y Desafíos de la Computación de Alto Rendimiento en la Nube	131
<i>Ezra F. Parra-González, Uziel Jaramillo-Avila, Pablo Salazar-Linares, Carlos A. Lara-Álvarez</i>	

Contrato Inteligente para la Gestión de Requerimientos en la Construcción de Software.....	147
<i>Wilson Soto</i>	

Nuevos desafíos en Tecnologías de Información y Comunicación

New challenges in Information and Communication Technologies

Jezreel Mejía¹, Yadira Quiñonez², Giner Alor-Hernández³, Abraham Dávila⁴

jmejia@cimat.mx; yadiraqui@uas.edu.mx; galor@ito-depi.edu.mx;
abraham.davila@pucp.edu.pe

¹Centro de Investigación en Matemáticas- Unidad zacatecas, Parque Quantum, Ciudad del Conocimiento, Avenida Lasec, andador Galileo Galilei, manzana 3 lote 7, C. P. 98160, Zacatecas, México.

²Facultad de Informática Mazatlán, Universidad Autónoma de Sinaloa, Mazatlán, 82000, México.

³Tecnológico Nacional de México/I. T. Orizaba, Av. Oriente 9 852, Col. Emiliano Zapata, Orizaba C.P. 94320, Veracruz, México

⁴Departamento de Ingeniería, Pontificia Universidad Católica del Perú, Lima, Perú

DOI: 10.17013/risti.49.1-3

1. Introducción

Los artículos presentados en este número fueron seleccionados de los artículos aceptados en el 11vo., Congreso Internacional de Mejora de Procesos Software (CIMPS) 2022, con sede local en la Ciudad de Acapulco, Guerrero, México del 19 al 21 de octubre del 2022. Este congreso ha tenido una tasa de aceptación del 30% de los artículos enviados.

Los artículos publicados en este número de la Revista Ibérica de Sistemas y Tecnologías de la Información (RISTI), aborda temas relacionados a “Nuevos Desafíos en Tecnologías de Información y Comunicación”. Los artículos presentan versiones extendidas y mejoradas de las versiones originales que han sido presentados en el congreso CIMPS 2022.

2. Estructura

El primer artículo toma como base los desafíos para la enseñanza de estándares en instituciones de educación superior obtenido de una muestra de 36 profesores sugiriendo

un conjunto de elementos de gamificación para diseñar estrategias que contribuyan a la educación en estándares de Ingeniería de Software, relacionados con los desafíos identificados.

El segundo artículo presenta un diseño de videojuegos para el análisis de habilidades personales ya que el estudio de las habilidades personales ha ido tomando mayor relevancia en los últimos años en todos los ámbitos; probablemente debido a que son fundamentales para el trabajo en equipo, además de ser clave para el éxito en la vida profesional a nivel individual. En este contexto, los videojuegos presentan la ventaja de abstraer al participante, haciéndolo olvidar que está siendo observado, generando un ambiente en el que se espera un desempeño más natural

El tercer artículo tiene como objetivo demostrar cómo la adopción de DevOps permite la automatización de los procesos del estándar ISO/IEC 29110 del perfil básico a través de un caso de estudio en el Centro de Desarrollo de Software del Instituto Tecnológico Superior de Nochistlán, Zacatecas, México. Los resultados obtenidos muestran el 78% de los productos de trabajo de los procesos del estándar automatizados incrementando la productividad del equipo y reduciendo el tiempo de entrega del software.

En el cuarto artículo tras realizar un análisis de las estructuras de los procesos dentro del estándar ISO/IEC 29110 en su perfil básico los autores identificaron falta de información acerca de los estados de los productos de trabajo que son necesarios para desarrollar las tareas, lo que dificulta su correcta implementación. Por lo tanto, los autores presentan tautologías de lógica matemática con el objetivo de presentar propuestas de mejora en el establecimiento de estados a productos de trabajo que ayudan en el control de la versión de los productos de trabajo y su documentación asociada.

El quinto artículo presenta el desarrollo de un sistema de monitoreo para la recolección de datos, tales como la temperatura y la humedad, así como el peso de los alimentos a lo largo del experimento nutricional, utilizando una aplicación móvil desarrollada en Android Studio que envía y recibe la información vía tecnología Bluetooth a un dispositivo electrónico basado en el microcontrolador Arduino Nano.

En el sexto artículo los autores proponen el mejor algoritmo de detección de matrículas en ambientes controlados y no controlados. Para ello, aplicaron una serie de algoritmos de inteligencia artificial para automatizar dicha detección obteniendo los resultados con imágenes tomadas con distintos ángulos y en distintas luminosidades (ambiente no controlado) donde el mejor Overall accuracy con 86 % y el segundo grupo fueron imágenes tomadas con un ángulo similar y distancias similares (ambiente controlado), obtuvo un Overall accuracy de 95.5 %.

El séptimo artículo presenta un mapeo sistemático de la literatura referente al uso de técnicas de aprendizaje profundo (AP) para la detección y clasificación de nanopartículas contenidas en micrografías SEM. Los resultados reflejan que las redes neuronales convolucionales (CNN) son las técnicas más utilizadas para analizar micrografías, obteniendo una alta precisión en los proyectos realizados en las publicaciones revisadas. Como prueba de concepto, se presentan ejemplos del uso de los enfoques más comunes en micrografías SEM de CaCO₃, incluyendo el uso de herramientas como OpenAI. Los resultados revelan ventajas y retos que se presentan al utilizar técnicas de aprendizaje profundo en el análisis de micrografías SEM.

El noveno artículo presenta las principales tendencias y desafíos de la Computación de Alto Rendimiento (HPC) en la nube. Para abordar estos desafíos, los autores revisan trabajos relacionados que permiten identificar los desafíos como la gestión de recursos y virtualización, lo que afecta su desempeño de la computación en la nube.

Finalmente, el décimo artículo propone el diseño y construcción de un contrato inteligente desplegado en una red blockchain para la gestión de los requerimientos dentro de una metodología ágil de construcción de software. Incluso el contrato inteligente propuesto puede ser usado para los pagos contractuales de cada requerimiento aceptado por el cliente.

3. Agradecimientos

Agradecemos al director de la Revista Álvaro Rocha por otorgar la oportunidad para dar a conocer los trabajos mas relevantes presentados en el congreso CIMPS 2022. Asimismo, expresamos nuestra gratitud a los autores de los trabajos publicados, lo cual hizo posible este número en RISTI. Damos la enhorabuena a todos aquellos autores cuyos artículos han sido publicados en este número de RISTI.

Gamificación para atender los desafíos de la enseñanza Ingeniería de Software en instituciones de educación superior

Mirna Muñoz¹, Gloria Piedad Gasca-Hurtado²

mirna.munoz@cimat.mx; gpgasca@udemedellin.edu.co

¹ Centro de Investigación en Matemáticas A.C.- unidad Zacatecas, Zacatecas, México.

² Universidad de Medellín, Carrera 87 No. 30-65, 50026, Medellín, Colombia.

DOI: [10.17013/risti.49.5-21](https://doi.org/10.17013/risti.49.5-21)

Resumen: El desarrollo profesional de los ingenieros de software es uno de los intereses de la industria, ya que las habilidades profesionales contribuyen al desarrollo de software de calidad. Por lo tanto, los procesos de enseñanza aprendizaje requieren transformaciones importantes que fortalezcan la educación en estándares de Ingeniería de Software, estos estándares son utilizados para el desarrollo de software de calidad. Con su implementación se logra resolver los retos de la industria del software en cuanto a la producción de software de calidad que satisfaga sus necesidades. El desarrollo de estas habilidades se hace más crítico en muy pequeñas organizaciones, siendo la principal fuente de trabajo de recién egresados. En este contexto, la gamificación está siendo utilizada como estrategia útil en la academia para aumentar las capacidades de los futuros profesionales con altos índices de éxito. Tomando como base los desafíos para la enseñanza de estándares en instituciones de educación superior, obtenido de una muestra de 36 profesores, en este artículo se sugiere un conjunto de elementos de gamificación para diseñar estrategias que contribuyan a la educación en estándares de Ingeniería de Software, relacionados con los desafíos identificados.

Palabras-clave: entidades muy pequeñas, estándares de calidad, desafíos en la enseñanza, entidades de educación superior, estrategias de enseñanza, gamificación.

Gamification for addressing the challenges of teaching international Software Engineering standards in higher education institutions

Abstract: The professional development of software engineers is one of the industry's interests since professional skills contribute to developing quality software. Therefore, the teaching-learning processes require significant transformations that strengthen education in Software Engineering standards used to create quality software. Its implementation helps solve the software industry's challenges regarding the quality of its processes. The development of these skills becomes more critical in very small entities. In this context, gamification is being used as a helpful strategy in academia to increase the skills of future professionals

with high success rates. Based on the challenges that arise in the process of teaching standards in higher education institutions from a sample of 36 teachers, this article suggests a set of gamification elements to design strategies that contribute to education in Software Engineering standards, specifically related to the identified challenges.

Keywords: very small entities, quality standards, teaching challenges, higher education institutions, teaching strategies, gamification.

1. Introducción

La industria del software está teniendo una gran oportunidad como proveedora del software para distintos entornos industriales. Este hecho favorece a las empresas de desarrollo de software grandes, medianas, pequeñas y muy pequeñas, en especial para las muy pequeñas (EMPs), debido al lugar que éstas ocupan en la cadena de suministro de software para las pequeñas y medianas empresas (Muñoz, Mejía & Laporte, 2019) (Ibarra, Vullingsh & Burgos, 2021).

Este hecho resalta la responsabilidad que las EMPs, deben afrontar para desarrollar software con la calidad requerida por sus clientes. En este contexto, el uso de estándares de Ingeniería de Software puede apoyar a las entidades muy pequeñas para lograr desarrollar software de calidad. Desafortunadamente, para la mayoría de éstas, se complica la implementación de estándares de calidad, debido a sus características propias como (Muñoz, et. al., 2021):

1. Tienen poca o nula experiencia en la implementación de mejoras en sus procesos, así como en la implementación de modelos y estándares.
2. Deben trabajar muy duro para sobrevivir en el mercado del software.
3. Tienen muy pocos empleados, y la mayoría de las veces son recién egresados contratados con poca o nula experiencia.
4. Tienen un mínimo o nulo presupuesto para invertir en actividades relacionadas con la mejora de sus procesos de desarrollo de software.

Esta investigación se enfoca en la poca o nula experiencia de los ingenieros que son contratados, debido a que las EMPs contratan recién egresados. En una publicación previa (Muñoz et al., 2016), se aborda este tema resaltando dos desafíos que enfrentan los recién egresados al integrarse en la industria de software. Por un lado, haber adquirido el conocimiento para desempeñar de una manera adecuada el rol que le sea asignado (e.g. gestor de proyecto, miembro del equipo de desarrollo). Por otro lado, que los egresados sean capaces de trabajar bajo un modelo o estándar requerido por la organización en la que se integran a laborar.

En este contexto, la gamificación, entendida como una estrategia que facilita la creación de experiencias lúdicas en un contexto no lúdico (Caponetto et al. 2014), está siendo muy usada para el aprendizaje dando excelentes resultados tanto a nivel de mejoras en el proceso enseñanza - aprendizaje como en la mejora de factores como la motivación y el compromiso. Además, la incorporación de elementos de gamificación en estrategias de enseñanza puede reforzar habilidades como la resolución de problemas, la colaboración y la comunicación (Rabah et al., 2018).

Este trabajo toma como punto de partida los resultados de un análisis realizado con una muestra de 36 profesores, para identificar un conjunto de desafíos a los que se enfrentan los profesores para enseñar estándares internacionales a sus alumnos durante su formación en las Instituciones de Educación Superior (IES) (Muñoz, 2021), para realizar una propuesta que integra elementos de gamificación utilizados para diseñar estrategias para abordar los desafíos que enfrentan las IES para la incorporación de los estándares internacionales de Ingeniería de Software en sus currículas y en la integración de nuevas estrategias de enseñanza que demanda el mundo.

Después de la introducción, este artículo está organizado como sigue: en la sección 2 se muestran trabajos relacionados con estándares de Ingeniería de Software y de Gamificación para la enseñanza en Ingeniería de Software. En la sección 3 se describe el análisis del estudio exploratorio realizado para determinar los desafíos en la enseñanza de estándares de Ingeniería de Software. La sección 4 detalla los desafíos identificados y presenta la propuesta de elementos de gamificación sugeridos para abordarlos. Finalmente, la sección 5 presenta las conclusiones y trabajo futuro de esta propuesta.

2. Trabajos relacionados

En esta sección se presentan trabajos relacionados con la enseñanza de estándares de Ingeniería de Software en la academia y el uso de gamificación para la enseñanza de Ingeniería de Software.

2.1. Enseñanza de estándares de Ingeniería de Software en la academia

Los estándares encapsulan un cuerpo de conocimiento común de la Ingeniería de Software representando un conjunto sólido de prácticas de desarrollo de software para la industria (Jovanovic & Andres, 1999)

Muñoz et al., en el artículo “*Análisis de la cobertura de las currículas de universidades de México para la industria de software en México*” (Muñoz et al., 2016), presentan un análisis de cobertura de currículas de universidades y la industria del software en México. El artículo analiza el nivel de cobertura que tienen carreras relacionadas con Ciencias de la Computación e Informática en referencia de la norma Moprosoft usada por la industria mexicana, el análisis se realizó para identificar el grado de cumplimiento de requisitos de conocimiento solicitados por la industria de software versus la formación recibida por los alumnos de carreras relacionadas con el desarrollo de software en las universidades. Como resultado se identificó un grado de cobertura medio entre el conocimiento requerido por el estándar y el proporcionado por las currículas de las universidades, resaltándose hallazgos referentes a la gestión de requisitos, gestión de riesgos, la validación, la medición y el análisis, y la gestión de la configuración como conocimiento que debe ser reforzado en las currículas.

Laporte, Muñoz and Bruel en el artículo “*La educación de estudiantes en el estándar ISO/IEC 29110 y su implementación en entidades muy pequeñas*” (Laporte, Muñoz & Gerançon, 2017) abordan las estrategias para implementación del estándar ISO/IEC 29110 que se han adoptado en diferentes países como Colombia, Brasil, Haití, Jordania, Malasia, México, Perú and Tailandia. En este artículo resalta la importancia de la

involucración de la academia para la enseñanza del estándar 29110 en países de América y Asia.

Claude y Muñoz en el artículo “*No enseñar estándares de ingeniería de software a futuros ingenieros, ¿es una mala práctica?*” (Laporte & Muñoz, 2021) analizan las consecuencias de la falta de uso de estándares en las organizaciones de desarrollo de software y de la falta de enseñanza de estándares a los ingenieros en formación y sus posibles consecuencias. Además, abordan cómo se podría apoyar a los profesores para la enseñanza de los estándares.

Vives, Melendez y Dávila en el artículo “*ISO/IEC 29110 and Software Engineering Education: A systematic Mapping Study*” (Vives, Melendez & Dávila, 2022) muestran los resultados de un estudio realizado con la enseñanza del estándar 29110 a los futuros ingenieros en la academia. Como resultado identifican dos técnicas didácticas utilizadas, por un lado, el aprendizaje orientado a proyectos, que habilita un espacio para adquirir experiencia en habilidades como son trabajo en equipo, visión holística del proyecto, pensamiento crítico y habilidades de análisis en situaciones reales. Por otro lado, las técnicas de gamificación para motivar de manera extrínseca el cumplimiento de los objetivos del curso.

2.2. Gamificación para la enseñanza en Ingeniería de Software

Se ha evidenciado que la gamificación brinda beneficios para promover el aprendizaje, el desempeño de los empleados, la participación del cliente e incluso iniciativas de crowdsourcing (Caponetto et al. 2014). En el contexto educativo, es considerada una estrategia interesante por su efectividad, ya que demuestra resultados positivos para la mejora del aprendizaje por medio de promoción de factores como la motivación y el compromiso. Además, tiene la capacidad de apoyar el refuerzo de habilidades como la resolución de problemas, la colaboración y la comunicación (Rabah et al., 2018).

A continuación se listan un conjunto de contribuciones que promueven el uso de la gamificación en el contexto de la enseñanza. Algunas de las contribuciones identificadas utilizan elementos como desafíos, que generan en los participantes motivación, mejora la capacidad de contribución al equipo, entre otros factores (Zhang et al., 2021; Yusri et al., 2020; Cheng et al., 2020; Norsanto & Rosmansyah, 2018). También es posible identificar propuestas donde la curiosidad, el deseo por descubrir y el elemento sorpresa (Palmas & Niermann, 2022; Gasca-Hurtado & Gomez-Alvarez, 2021; Lee et al., 2021; Gough et al., 2021; Zhang et al., 2021) son considerados importantes para el diseño de estrategias de gamificación enfocadas a la educación.

En términos generales, se logra identificar que hay un interés particular por el desarrollo de propuestas centradas en gamificación. Por lo tanto, se identifican en la literatura algunas propuestas que buscan evaluar la gamificación y arrojan resultados positivos, sobre todo en el campo de la enseñanza (Palmas & Niermann, 2022) al lograr motivar e involucrar a los alumnos, apoyando los procesos de aprendizaje diarios.

Otra propuesta muestra resultados relacionados con la enseñanza en Ingeniería de Software utilizando microaprendizaje y gamificación, donde se evidencia la asimilación

y fijación de conceptos por parte de los estudiantes en el área de conocimiento aplicada (Gasca-Hurtado & Gomez-Alvarez, 2021).

Finalmente, se menciona la propuesta con resultados positivos que reflejan la satisfacción de los alumnos durante una estrategia de gamificación en Ingeniería de Software relacionada con el algoritmo de emparejamiento (Yusri et al., 2020).

3. Análisis de los desafíos en la enseñanza de estándares de Ingeniería de Software

Este análisis se realizó a partir de los resultados de una encuesta para recopilar los datos. El instrumento se compartió por medio de invitaciones directas a profesores de Instituciones Educación Superior (IES) de México, Colombia, España y Chile. Además, se compartió a grupos de investigadores por WhatsApp y por LinkedIn.

La encuesta estuvo abierta dentro del periodo del 18 de enero al 1 de marzo del 2021 y recibieron un total de 36 respuestas de profesores de diferentes IES. De acuerdo con los datos recopilados por el instrumento, se identificó la participación de profesores de 4 países: 30 de México, 3 de España, 2 de Ecuador y 1 de Panamá.

3.1. Asignaturas impartidas

La Tabla 1 muestra una lista de 30 asignaturas en las que se podrían impartir temas relacionados con estándares de Ingeniería de Software. Éstas fueron seleccionadas de un total de 123 asignaturas mencionadas por los profesores (lista completa de asignaturas pueden consultarse en (Muñoz, 2022)).

Asignaturas	Profesores
Ingeniería de Software I y II	Prof2, Prof3, Prof6, Prof7, Prof8, Prof9, Prof11, Prof12, Prof16, Prof18, Prof19, Prof21, Prof23, Prof25, Prof36
Administración de proyectos, administración de proyectos ágil, administración de proyectos TSP, gestión cuantitativa de proyectos de software	Prof3, Prof11, Prof12, Prof15, Prof16, Prof17, Prof24, Prof30, Prof32
Sistemas de calidad en TI, calidad de software, métricas del software, evaluación de proyectos de software	Prof3, Prof1, Prof12, Prof13, Prof17, Prof27, Prof30, Prof34
Modelados de procesos de negocio	Prof3, Prof16,
Requisitos de software, Ing. de requisitos	Prof8, Prof11, Prof16, Prof19, Prof21
Gestión de la TI en una empresa	Prof9,
Modelos y métodos de mejora	Prof9
Arquitectura de software, arquitectura orientada a servicios	Prof10, Prof15,
Procesos de desarrollo, proceso personal de software	Prof10, Prof11, Prof34, Prof16
Administración de empresas de software, cultura empresarial	Prof12, Prof35

Asignaturas	Profesores
Modelado conceptual, modelado orientado a objetos, modelado de datos	Prof13, Prof23
Construcción del software, programación taller	Prof15, Prof26, Prof28, Prof29
Desarrollo ágil del software	Prof16,
Desarrollo de software	Prof17
Análisis y diseño de programas	Prof26, Prof33
Seguridad	Prof28, Prof29

Tabla 1 – Materias impartidas en las que se pueden impartir estándares de Ingeniería de Software

Del conjunto de materias reportadas por los profesores se identifican que las asignaturas de Ingeniería de software I y II son las que han sido mencionadas como impartidas con un total de 15 profesores, seguida por gestión de proyectos, mencionada por 9 profesores y calidad del software mencionada por 7 profesores.

3.2. Caracterización de la enseñanza de estándares

En análisis de las razones por las cuales existen deficiencias en la enseñanza de estándares se muestra en la Tabla 2.

ID	Razones	profesores
1	No se incluyen las normas en las asignaturas de las carreras	Prof1, Prof3, Prof6, Prof10, Prof25, Prof28, Prof29
2	No existen estándares para la asignatura	Prof4
3	Falta de reconocimiento de su utilidad	Prof7, Prof21
4	Los profesores no cuentan con preparación en Ingeniería de Software	Prof8, Prof10, Prof12, Prof13, Prof16, Prof18, Prof22, Prof32
5	Se centran en el qué más no el cómo	Prof9, Prof30
6	Son muy costosos	Prof10, Prof13
7	Requieren una actualización constante de versiones	Prof13, Prof22, Prof31, Prof35
8	Por desconocimiento	Prof14, Prof20, Prof22, Prof28, Prof29
9	Son aburridos	Prof23
10	Toman exceso de tiempo y papeleo que no aporta conocimiento	Prof23
11	Son complicados de enseñar cuando la formación de los ingenieros se centra en marcos de trabajo y lenguajes de programación	Prof24, Prof27, Prof30, Prof32, Prof34
12	Son absurdos los tiempos programados para enseñarlos	Prof26, Prof36

ID	Razones	profesores
13	No se usan en la industria	Prof12
14	Otra respuesta no relacionada a la pregunta	Prof2, Prof5, Prof15, Prof17, Prof19, Prof33, Prof11

Tabla 2 – Razones por las que no se enseñan estándares

Como se observa en la Tabla 2, la razón con más respuestas es la *falta de preparación de los profesores en Ingeniería de software* con 8 respuestas de profesores, seguido por la razón de que *no se incluyen estándares como parte de las asignaturas de las carreras* con 7 respuestas de profesores. Aunado a estas dos razones, podemos agregar al *desconocimiento por parte de los profesores*, debido a la complejidad que representa enseñarlos cuando los alumnos carecen de conocimiento previo, ya que su formación se centra en marcos de trabajo y lenguajes de programación respondido por 5 profesores cada razón.

3.3. Estándares comúnmente enseñados

Se identificaron 28 recursos diferentes, no son únicamente estándares. Entre las respuestas se mencionan metodologías, modelos y estándares, como se lista a continuación:

- Estándares: IEEE 1028 (1); IEEE 739(1); ISO 25000 (1); ISO/IEC 19761 COSMIC (1); IEEE 829 (1); ISO 21500(2); ISO 25000 (SQUARE) (2); ISO/IEC 20000 (2); ISO/IEC 330XX (SPICE) (2); ISO/IEC 12207 (2); ISO 9001 (4); IEEE Std. 830-1996 (5); MOPROSOFT, NMX-1-059-NYCE (6); ISO/IEC 29110 (13).
- Modelos: ciclo de vida cascada, espiral (1), ITIL (1), modelo relacional (2); CMMI (3).
- Metodologías y métodos: SCRUM (2); TSP (3); PSP(3).
- Otros: UML (1); SWEBOK (1); RUP (1); relacionados con IoT, IEEE, ETSi, CMU (3); PMBOK (4).

Como se observa en la lista los estándares que más se enseñan por parte los la muestra de profesores son el ISO/IEC 29110 con 13 respuestas, seguido de Moprosoft con 6 respuestas y el IEEE Std. 830-1998 con 5 respuestas.

Cabe resaltar que la mayoría de los profesores de la muestra son mexicanos y Moprosoft es el modelo para la industria mexicana y además es la base del estándar ISO/IEC 29110.

3.5. Beneficios de enseñar estándares

Con respecto a los beneficios que los profesores tienen identificados sobre la enseñanza de estándares internacionales, se pueden clasificar en dos grupos, por una parte, el beneficio que las empresas pueden tener al contratar ingenieros recién egresados formados en éstos, y por otro lado, los beneficios que los alumnos obtienen al ser formados en estos estándares. Cabe resaltar que del total de las 36 respuestas analizadas únicamente se encontró que una respuesta no estaba relacionada con la pregunta realizada.

- *Beneficios enfocados en las empresas:* (1) creación de software con calidad, (2) aseguramiento del producto de acuerdo al estándar, buenas prácticas demostradas, conocimiento sobre certificaciones en esos estándares, (3) promover la mejora de procesos al valorar la importancia del uso de éstos, y (4) desarrollo de proyectos TI más profesionales, relacionado con el seguimiento y aseguramiento de calidad debido al conocimiento de buenas prácticas.
- *Beneficios enfocados en los alumnos:* (1) los alumnos adquieren conocimiento integral y de regulación internacional de acuerdo a la temática tratada, (2) los alumnos egresan de las universidades mejor preparados para desarrollar software usando buenas prácticas de gestión e ingeniería, (3) desarrollan un mejor conocimiento y entendimiento de prácticas de Ingeniería de Software enfocadas a la calidad, (4) adquieren más conocimiento para el desarrollo de proyectos propios de emprendimiento, (5) mejoran su habilidad de trabajo en equipo teniendo la habilidad para distribuir responsabilidades y asumir compromisos, (6) mejoran el entendimiento en relación a la existencia de cuerpos de conocimiento elaborados, revisados y aplicados de forma global, en organizaciones de todos los países, y (7) mejor colocación de los estudiantes en empresas.

Es importante resaltar que ambos beneficios están relacionados, debido a que al adquirir conocimiento y un mejor entendimiento de la utilidad de los estándares, los recién egresados pueden promover su uso para la definición y/o mejora de procesos en las organizaciones de desarrollo de software.

3.5. Necesidades detectadas para lograr la enseñanza de estándares en instituciones de educación superior

Con las respuestas recibidas para esta pregunta se identificaron 5 categorías de necesidades que a continuación se describen:

1. *Capacitación a los profesores:* 21 respuestas hacen énfasis en la necesidad de capacitar y actualizar a los profesores en los estándares. Se resalta que la formación en estándares cubra tanto la parte teórica como la parte práctica para que los profesores puedan enseñar los estándares de una manera adecuada a los alumnos.
2. *Experiencia práctica:* 7 respuestas hacen referencia a la necesidad de que los profesores adquieran experiencia práctica en la implementación del o los estándares que se pretenden enseñar, tal que, los profesores tengan identificados los beneficios de su implementación y estos beneficios puedan ser transmitidos a sus alumnos.
3. *Incluir la enseñanza de los estándares como parte de las currículas de las IES:* 8 respuestas hacen énfasis en la necesidad de incluir a los estándares como parte de las currículas universitarias. Sin embargo, resaltan, por un lado, la necesidad de hacer una selección de los estándares que estén relacionados con la formación del alumno, y por otro lado, que sean los que están requiriendo las empresas.

4. *Presupuesto para adquisición de los estándares y certificaciones*: 3 respuestas hacen énfasis en la necesidad de contar con presupuesto en las IES para adquisición de estándares y el pago de certificaciones relacionadas con los mismos, debido a que los profesores no pueden tener acceso libre a los estándares por sus costos elevados, y además, no pueden optar por las certificaciones asociadas a estos estándares.
5. *Que los estándares sean requeridos por las empresas*: 2 repuestas hacen énfasis en la importancia de que sea la industria la que solicite a las IES la enseñanza de estándares requeridos por ser aplicados en su actividad diaria para garantizar la calidad de sus productos, procesos o servicios, debido a que el uso de los estándares no es constante en las empresas de desarrollo de software, debido a que su uso depende del tipo de cliente que atienden y la madurez de las mismas.

4. Propuesta de elementos de gamificación para abordar los desafíos identificados

Los elementos de gamificación son considerados la base para el diseño de estrategias de gamificación. Por esta razón, existen estudios donde se describen los elementos utilizados en el diseño de estrategias de gamificación, específicamente estrategias enfocadas al ambiente de aprendizaje (de Monterrey, 2016).

Existen diversos marcos de referencia para el diseño de la Gamificación tales como: Mechanics, Dynamics & Aesthetics - MDA (Hunicke, LeBlanc y Zubek, 2004; Zichermann y Cunningham, 2011), Mechanics, Dynamics & Emotions - MDE (Robson et al., 2015) y Dynamics, Mechanics & Components - DMC (Werbach & Hunter, 2015), entre otros.

Aunque hay diversidad de clasificaciones y definiciones de los elementos de juego (Dicheva et al., 2015), es posible identificar algunas que se destacan por su popularidad en la literatura, es el caso de mecánicas, dinámicas, componentes, estéticas y emociones (Wiklund & Wakerius, 2016).

A partir de estas diversas propuestas, en este trabajo se realiza una trazabilidad con los desafíos identificados a partir del estudio exploratorio. La caracterización de elementos de gamificación establecida por el Observatorio de Innovación Educativa del Tecnológico de Monterrey (de Monterrey, 2016), es la base para la trazabilidad realizada en este artículo.

Con la trazabilidad mencionada se pretende proponer un conjunto de elementos de gamificación sugeridos para ser utilizados en el diseño de estrategias de gamificación e instrumentos pedagógicos (Gómez-Alvarez et al., 2016).

A continuación, se mencionan los desafíos identificados a partir de los resultados de la encuesta y de las necesidades identificadas en la sección anterior. Los elementos de gamificación recomendados para abordar cada desafío son enunciados a continuación.

1. Compromiso de las IES.
 - Descripción del desafío: se requiere un análisis de currículas en conjunto con un grupo empresas de desarrollo de software. Este análisis debe dar

como resultado un listado de necesidades de las organizaciones frente a las habilidades requeridas por los profesionales de Ingeniería de Software. Este listado de necesidades permitirá acotar la brecha entre la formación recibida por los estudiantes en las IES y los requisitos de la industria del software.

- Elementos de gamificación: la trazabilidad que se puede establecer entre este reto identificado y los elementos de juego está relacionada a partir de la necesidad del trabajo colaborativo. Por lo tanto, los elementos de gamificación de la categoría metas y objetivos, tales como retos, misiones y desafíos épicos podrían ser el pilar fundamental de una estrategia de gamificación. Además, para es necesario promover en los participantes el logro de los objetivos. Por lo tanto, se sugiere incorporar elementos de la categoría libertad para elegir. Estos elementos están relacionados con proporcionar al participante diferentes rutas o casillas para llegar a una meta, opciones de usar poderes o recursos, que pueden ser incorporados a la estrategia debidamente articulados con la categoría metas y objetivos. Tanto las empresas como las IES trabajan por un *statu quo*, por lo que puede ser interesante incluir en la estrategia una categoría de elementos de gamificación como el estatus visible que se materializa con elementos como insignias, puntos, logros, resultados obtenidos, tablero de posiciones.
2. Formación de profesores.
 - Descripción del desafío: este reto está asociado a la necesidad del diseño de estrategias de adquisición de conocimiento teórico-práctico para profesores. Este reto permitirá fortalecer las habilidades para enseñar estándares internacionales de Ingeniería de Software. Los requerimientos para lograr estrategias que enfrenten este reto están asociados con la generación de la suficiente confianza del profesor o facilitador, a partir de estrategias donde la retroalimentación sea un pilar fundamental de la formación.
 - Elementos de gamificación: la trazabilidad que se puede establecer con los elementos de gamificación está dada a partir del uso de elementos de juego como las pistas visuales, señalizaciones de respuesta o conducta correcta o incorrecta, barras de progreso, advertencias sobre riesgos que se tienen al realizar cierta acción, estadísticas del desempeño del jugador. Además, en un proceso de formación es indispensable permitir que el participante sienta libertad de equivocarse, por lo que es necesario que una estrategia de gamificación enfocada a la formación de profesores incluya elementos como vidas múltiples, puntos de restauración o reinicio, número ilimitado de posibilidades. Otra alternativa para consolidar la estrategia de gamificación es incluir elementos de categoría progreso, que facilita medir el nivel de formación adquirido, por ejemplo, a través de elementos de juego como los tutoriales para desarrollo de habilidades iniciales, puntos de experiencia, niveles, barras de progreso y acceso a contenido bloqueado.
 3. Entendimiento del conocimiento asociado al estándar.
 - Descripción del desafío: se requieren recursos que faciliten el entendimiento del conocimiento codificado en los estándares, y que éste además se establezca como un punto de apoyo para la enseñanza.

- Elementos de gamificación: la trazabilidad con elementos de gamificación se recomienda utilizando la categoría narrativa que puede materializarse a partir de elementos de gamificación como identidades, personajes o avatares; mundos, escenarios narrativos o ambientes tridimensionales. Con estos elementos se puede conseguir que la comprensión o el entendimiento de un tema particular sea más efectiva al considerar elementos donde hay una identificación del personaje con el actor real de la estrategia de gamificación. Al igual que la categoría retroalimentación a partir elementos de juego como pistas visuales, señalizaciones de respuesta o conducta correcta o incorrecta, barras de progreso, advertencias sobre riesgos que se tienen al realizar cierta acción, y estadísticas del desempeño del jugador.

4. Recursos económicos.

- Descripción del desafío: este desafío está asociado con los costos de las certificaciones asociadas con los estándares y las inversiones de alto nivel que deben realizar las IES para recursos de laboratorios de práctica de los estándares.
- Elementos de gamificación: a pesar de que es una restricción externa a una estrategia de gamificación, es posible sugerir la inclusión de elementos de gamificación como parte de la trazabilidad. Específicamente aquellos asociados con la categoría recompensas que se materializan con elementos como las monedas o recursos virtuales, vidas, poderes limitados. Estos elementos pueden ser combinados con las estrategias que se diseñen para afrontar los retos antes mencionados, donde los bienes recibidos en el juego estén relacionados con el logro de los objetivos, es decir con temas como el costo de la certificación, entre otros. Este recurso motiva la competencia y el sentimiento de logro de los participantes.

No obstante, esta propuesta de trazabilidad con los elementos de gamificación se considera que este reto debe ser intervenido desde una perspectiva diferente a la gamificación, ya que los recursos financieros de implementación y formación en estándares internacionales supone altos costos.

5. Concientización de empresas.

- Descripción del desafío: se requiere adquirir el compromiso formal por parte de las empresas para concientizar a las IES de la necesidad, importancia y beneficio del uso de estándares. Específicamente los beneficios asociados con la producción de software de calidad que implica establecimiento de procesos de calidad.
- Elementos de gamificación: una estrategia de gamificación con elementos de tipo Metas y objetivos como Desafíos, misiones, desafíos épicos, donde la IES y la EDS logren gestionar y diseñar sus propias misiones y desafíos en las que pretende involucrar sus participantes, así como establecer desafíos particulares para ser alcanzados durante el proceso de formación o implementación de las estrategias de gamificación con indicadores de logro de los participantes.

Finalmente, se sugiere que la categoría recompensa se considere como una categoría transversal. Esta categoría está compuesta por elementos tales como monedas o recursos virtuales, vidas y poderes, entre otros. La recompensa, además de ser una categoría de elementos de gamificación como los enunciados, se considera un principio básico de la gamificación ya que se ha demostrado que el uso de recompensas virtuales o reales aumenta la adopción de nuevas iniciativas (Perryer et al., 2012; Whitton, 2012; Dominguez et al., 2013). Por lo tanto, autores como Oprescu, et al., 2014 incluyen las recompensas basadas en logros como un principio clave para la gamificación al considerarlo un apoyo fundamental por sus beneficios orientados al bienestar (Perryer et al., 2012; Dominguez et al., 2013).

A continuación se presenta una matriz de trazabilidad, d a manera de resumen de la propuesta de este trabajo, a partir de retos identificados de un estudio exploratorio y elementos de gamificación sugeridos por el Observatorio de Innovación Educativa del Tecnológico de Monterrey (Tabla 3).

Reto	Metas y objetivos	Libertad para elegir	Estatus visible	Retroalimentación	Libertad para equivocarse	Progreso	Narrativa	Recompensa
1. Compromiso	Retos Misiones Desafíos épicos	Rutas o casillas para llegar a una meta Opciones de usar poderes o recursos	Insignias Puntos Logros Resultados obtenidos Tablero de posiciones					
2. Formación				Pistas visuales Señalizaciones de respuesta o conducta correcta o incorrecta Barras de progreso	Vidas múltiples Puntos de restauración o reinicio Número ilimitado de posibilidades	Tutoriales para desarrollo de habilidades iniciales Puntos de experiencia Niveles Barras de progreso		Monedas o recursos virtuales Vidas Poderes limitados
3. Entendimiento				Pistas visuales Señalizaciones de respuesta o conducta correcta o incorrecta Barras de progreso			Identidades, personajes o avatares Mundos Escenarios narrativos o ambientes tridimensionales	
4. Concientización	Retos Misiones Desafíos épicos							

Tabla 3. Matriz de trazabilidad de retos y elementos de gamificación.

En la matriz que se presenta en la Tabla 3 es posible evidenciar una categoría adicional de elementos de gamificación que hace referencia a monedas, recursos virtuales como vidas, poderes entre otros. Esta categoría se sugiere como una categoría común para cualquier estrategia que se diseñe en el marco de los retos identificados e reto asociado con el financiamiento ha sido eliminado por considerarse un reto externo que debe ser abordado desde otra perspectiva diferente a la gamificación.

Como parte del análisis que se propone en este artículo, a continuación, se presentan ejemplos de estrategias de gamificación previamente diseñadas que utilizan los elementos de gamificación de matriz de trazabilidad. Estas estrategias se presentan a manera de ejemplo y pueden ser utilizadas como punto de partida para la creación de nuevas estrategias, o bien para abordar el reto.

1. Compromiso de las IES: un ejemplo herramienta que implementa una estrategia de gamificación para abordar este reto es el Sistema de Exploración Planetaria.
 - Propósito: disminuir causas de resistencia al cambio en empresas de desarrollo de software.
 - Descripción: utiliza los datos reales para interactuar con plataformas de sistemas de gestión de incidentes y gestiones de proyectos. Esta herramienta software cuenta con un panel administrativo para configurar exploraciones, mundos, misiones y asociar los jugadores con la configuración dada. A partir de la integración del Sistema de Exploración Planetaria con sistemas de gestión de incidentes, gestiones de proyectos y herramientas de apoyo de procesos de mejora como los mencionados, es posible asignar valores ponderables a Teckitz. Los Teckitz es el equivalente al combustible de las naves, necesario para recorrer y cumplir las misiones en los planetas. En la estrategia propuesta también se define una escala de asignación de Teckitz relacionada con las actividades realizadas en la plataforma. De esta manera se busca la mejora del proceso de gestión de incidentes en una empresa de software.
 - Objetivos perseguidos: incrementar el conocimiento del equipo sobre el proceso mejorado a través de una definición clara de las responsabilidades de cada miembro. Además, con la estrategia se logra motivar al equipo a realizar las diferentes actividades del proceso mejorado mostrando sus beneficios. Otro de los objetivos está orientado a facilitar la integración de los diferentes actores del proceso a través de una visión compartida del propósito y alcance de la mejora esperada del proceso (Gasca-Hurtado, Gómez-Álvarez & Herrera, 2020).
2. Formación de profesores: un ejemplo herramienta que implementa una estrategia de gamificación para abordar este reto es la aplicación móvil denominada “Mr. Lottery”.
 - Propósito: Proporcionar conocimientos y habilidades de una manera más dinámica que la tradicional utilizando elementos de gamificación para diseñar una estrategia educativa involucrando el micoraprendizaje. El objetivo es dar mayor agilidad y brevedad con lecciones y unidades de aprendizaje proporcionadas a partir la asociación de conceptos y definiciones.

- Descripción: constituye una evolución del juego de mesa Lotería o Bingo que consta de un conjunto de cartas y un número de tarjetas distribuidas en tableros de manera aleatoria. En esta herramienta en lugar de asociar conceptos e imágenes en los tableros guía, los participantes deben asociar definiciones y conceptos sobre un área de conocimiento específica (configurable).
 - Objetivos perseguidos: Esta herramienta está presentada como estrategia de microaprendizaje diseñada con elementos de gamificación sugeridos como parte de las estrategias útiles para enfrentar los desafíos caracterizados en la categoría de formación de profesores (Gasca-Hurtado & Gomez-Alvarez, 2021).
3. Entendimiento del conocimiento asociado al estándar: un ejemplo herramienta que implementa una estrategia de gamificación para abordar este reto es un videojuego diseñado a partir de principios de gamificación para generar habilidades de trabajo en equipo y cooperación entre un grupo de estudiantes.
- Propósito: Se trata de una herramienta de realidad virtual colaborativa que fomenta que los participantes se involucren en la tarea que deben llevar a cabo.
 - Descripción: En el videojuego los participantes forman parte de una tripulación de un barco que después de una batalla quedó destruido y deben volver a armarlo. El barco y sus partes tienen un esquema y unos colores determinados para identificar las restricciones que maneja cada uno de los roles de los jugadores. Los participantes visualizan las islas donde se ubican ciertas piezas del barco. Los jugadores tendrán la oportunidad de ver una descripción de los posibles roles a desempeñar junto con sus limitaciones físicas (Gasca-Hurtado, et al., 2015).
 - Objetivos perseguidos: La interacción entre usuarios en un entorno virtual colaborativo es un valioso sustituto de una experiencia en la vida real, proporcionando a los participantes una práctica en primera persona que ayudará a generar un comportamiento espontáneo basado tanto en el involucramiento como en la sensación de co-presencia. Dentro de las reglas del juego está el fomentar la colaboración entre cuatro participantes.

5. Conclusiones

La industria del software, en especial muy pequeñas organizaciones EMPs, demanda profesionales del área de la Ingeniería de Software con altos niveles de cualificación. En este contexto, identificar y reconocer los desafíos que actualmente se tienen frente a la formación y entrenamiento de los futuros profesionales, se convierte en una responsabilidad tanto de la industria del software como de la académica, representada principalmente por las IES.

En este artículo, se presentan los desafíos y elementos de gamificación para abordar dichos retos. La principal contribución de la propuesta es su utilidad para el desarrollo de recursos de gamificación para la enseñanza de estándares internacionales de Ingeniería de Software. Estos estándares están siendo considerados como una de las principales

fuentes de información veraz y verificable que facilita la producción de productos y servicios de software de calidad. La principal limitación del análisis exploratorio base para esta propuesta, es el tamaño de muestra de 36 profesores, ya que no permite generalizar los resultados obtenidos.

Para este análisis se tomo como referencia la categorización de elementos de gamificación realizada por el Observatorio de Innovación Educativa del Tecnológico de Monterrey (de Monterrey, 2016). Además, A partir de la matriz de elementos de gamificación y desafíos identificados, fue posible mapear estrategias de gamificación diseñadas con elementos como los trazados para los desafíos.

Como trabajo futuro se pretende refinar el instrumento de recolección de datos y expandir la muestra de profesores tanto de México como de otros países de habla hispana para lograr la generalización de resultados. A partir de lo cual, será posible generar estrategias de gamificación específicas para los desafíos y validar los resultados de las posibles implementaciones que se realicen en grupos de individuos donde se puedan lograr resultados del uso de dichas estrategias.

Referencias

- de Monterrey, T. T. (2016). Observatorio de Innovación Educativa. Gamificación.
- Dicheva, D., Dichev, C., Agre, G., & Angelova, G. (2015). Gamification in Education: A Systematic Mapping Study. *Educational Technology & Society*, 18(3), 75–88. Recuperado de: http://www.ifets.info/journals/18_3/6.pdf
- Gasca-Hurtado, G. P., et al. (2015). Realidad virtual como buena práctica para trabajo en equipo con estudiantes de ingeniería= Virtual Reality as good practice for teamwork with engineering students. *RISTI - Revista Ibérica de Sistemas y Tecnologías de la Información* (16),76-91.
- Gasca-Hurtado, G. P., & Gomez-Alvarez, M. C. (2021). Mobile application based on gamification to promote microlearning in Software Engineering. 2021 16th Iberian Conference on Information Systems and Technologies (CISTI). <https://doi.org/10.23919/cisti52073.2021.9476569>
- Gasca-Hurtado G. P., Gómez-Álvarez M. C. & Herrera S. (2020) Gamified tool to mitigate change resistance causes in software process improvement. 2020 15th Iberian Conference on Information Systems and Technologies (CISTI), Seville, Spain, 2020, pp. 1-7. <https://doi.org/10.23919/CISTI49556.2020.9141010>.
- Gómez-Alvarez, M. C., Gasca-Hurtado, G. P., Manrique-Losada, B., & Arias, D. M. (2016). Method of pedagogic instruments design for software engineering. In 2016 11th Iberian Conference on Information Systems and Technologies (CISTI) (pp. 1-6). IEEE.
- Gasca-Hurtado, G. P., Gómez-Álvarez, M. C., Machuca-Villegas, L., & Muñoz, M. (2021). Design of a gamification strategy to intervene in social and human factors associated with software process improvement change resistance. *IET Software*, 15(6), 428-442.

- Hunicke, R., LeBlanc, M., & Zubek, R. (2004). MDA: A formal approach to game design and game research. *Memorias en Challenges in Game AI* (Vol. 4, p. 1). <http://www.aaai.org/Papers/Workshops/2004/WS-04-04/WS04-04-001.pdf>
- Ibarra, G., Vullingsh, S., Burgos, F.J. (2021). *Panorama Digital de las Micro, Pequeñas y Medianas Empresas (MiPymes) de América Latina 2021*. Santiago, GIA Consultores.
- Jovanovic, V., Andres, L., & Sherlund, B. (1999). Use of software engineering standards in teaching. In *Proceedings 4th IEEE International Software Engineering Standards Symposium and Forum (ISESS'99)*. Best Software Practices for the Internet Age. Curitiba, Brazil, 1999, pp. 122-130, <https://doi.org/10.1109/SESS.1999.766586>
- Laporte, C. Y., Muñoz, M., Gerançon, B. (2017). The education of students about ISO/IEC 29110 software engineering standards and their implementations in very small entities. 2017 IEEE Canada International Humanitarian Technology Conference (IHTC), 2017, pp. 94-98, <https://doi.org/10.1109/IHTC.2017.8058208>
- Laporte, C.Y., & Munoz, M. (2021). Not Teaching Software Engineering Standards to Future Software Engineers-Malpractice?. *Computer*, 54(05), 81-88. <https://doi.org/10.1109/MC.2021.3064438>
- Muñoz, M., (2021). Desafíos en la enseñanza de estándares internacionales en instituciones de educación superior. In the 12th International Conference of Software Process Improvement CIMPS 2022
- Muñoz, M., Mejia, J., Laporte, C.Y. (2019). Reinforcing Very Small Entities Using Agile Methodologies with the ISO/IEC 29110. In: Mejia, J., Muñoz, M., Rocha, Á., Peña, A., Pérez-Cisneros, M. (eds) *Trends and Applications in Software Engineering. CIMPS 2018, Advances in Intelligent Systems and Computing*, vol 865. Springer, Cham. https://doi.org/10.1007/978-3-030-01171-0_8
- Muñoz, M., Mejia, J., Peña, A., Laporte, C., Gasca-Hurtado, G.P., Gómez-Álvarez, M.C. (2021). An Exploratory Analysis of the Perception of the Utility of Proven Practices of the Software Basic Profile of ISO/IEC 29110 by a Set of VSEs in Mexico. In: Yilmaz, M., Clarke, P., Messnarz, R., Reiner, M. (eds) *Systems, Software and Services Process Improvement. EuroSPI 2021, Communications in Computer and Information Science*, vol 1442. Springer, Cham. https://doi.org/10.1007/978-3-030-85521-5_29
- Muñoz, M., Pena, A., Mejía, J., & Lara, G. (2016). Coverage of the university curricula for the Software Engineering industry in Mexico. *IEEE Latin America Transactions*, 14(5), 2382-2388. <https://doi.org/10.1109/TLA.2016.7530436>
- Robson, K., Plangger, K., Kietzmann, J. H., McCarthy, I., & Pitt, L. (2015). Is it all a game? Understanding the principles of gamification. *Business Horizons*, 58(4), 411-420. http://s3.amazonaws.com/academia.edu.documents/38021108/2_015_Is_it_all_a_game__BH.pdf?AWSAccessKeyId=AKIAJ56TQJRT
- Werbach, K., & Hunter, D. (2012). *For the win: How game thinking can revolutionize your business*. Wharton Digital Press.

- Wiklund, E., & Wakerius, V. (2016). The Gamification Process: A framework on gamification. <http://www.diva-portal.org/smash/get/diva2:931932/FULLTEXT01.pdf>
- Vives, L., Melendez, K., & Dávila ,A. (2022). ISO/IEC 29110 and Software Engineering Education: A Systematic Mapping Study. *Program. Comput. Softw.*, 48(8), 745–755. <https://doi.org/10.1134/S0361768822080229>
- Zichermann, G., & Cunningham, C. (2011). *Gamification by design: Implementing game mechanics in web and mobile apps*. Canada: O'Reilly Media.

Diseño de videojuegos para el análisis de habilidades personales

Adriana Peña Pérez Negrón¹, David Bonilla Carranza¹, Mirna Muñoz², Raúl Aguilar³

adriana.pena@cucei.udg.mx; jose.bcarranza@academicos.udg.mx;
mirna.munoz@cimat.mx; avera@correo.uady.mx

¹ Universidad de Guadalajara, CUCEI, Blvd. Marcelino García Barragán 1421, C.P. 44430, Guadalajara, Jal. Mexico

² CIMAT Zacatecas, C. Lasec y And. Galileo Galilei, Mn 3, L 7 Quantum Ciudad del Conocimiento, C.P. 98160 Zacatecas, Zac. Mexico

³ Facultad de Matemáticas, Universidad Autónoma de Yucatán, Inn, Anillo Periférico Norte tablaje 13615, Chuburna de Hidalgo, 97119, Mérida, México

DOI: 10.17013/risti.49.22-36

Resumen: El estudio de las habilidades personales ha ido tomando mayor relevancia en los últimos años en todos los ámbitos; probablemente debido a que son fundamentales para el trabajo en equipo, además de ser clave para el éxito en la vida profesional a nivel individual. Sin embargo, a diferencia de las habilidades técnicas, determinar si una persona posee o domina habilidades personales no es una tarea trivial. Actualmente los mecanismos más comunes de valoración son la entrevista y los instrumentos psicométricos, a los que se ha unido como alternativa el uso de videojuegos. Los videojuegos presentan la ventaja de abstraer al participante, haciéndolo olvidar que está siendo observado, generando un ambiente en el que se espera un desempeño más natural. En este contexto, existen diversas propuestas que se han tomado como base para generar un proceso para analizar habilidades personales mediante videojuegos.

Palabras-clave: análisis de estilo del jugador; juegos digitales; habilidades blandas; instrumentos de evaluación.

Video game design for personal skills analysis

Abstract: In recent years, the study of personal skills has become relevant in all domains; probably because they are essential for teamwork, as well as key for individual success in professional life. However, unlike technical skills, determining whether a person dominates certain soft skill is not a trivial task. Currently the most common mechanisms to assess them are the interview and psychometric instruments, to them the use of video games has been added as an alternative. Videogames present the advantage of abstracting the participants, making them to forget they are being observed, producing an environment in which a more natural performance is expected. In this context, various existing proposals have been used to generate a process for analyzing soft skills selecting video games.

Keywords: player style analysis; video games; soft skills; assessment instruments.

1. Introducción

Mediante el análisis del estilo del jugador, la industria del videojuego puso sobre la mesa el estudio de rasgos o características personales del usuario, en beneficio de la experiencia de juego. La experiencia de juego se refiere a la interacción del jugador con el espacio virtual, lo que en él puede hacer, incluyendo los procesos y el resultado de dichas interacciones. En un videojuego, todas las acciones y reacciones del jugador representan valiosa información que permite entender su comportamiento dentro del escenario virtual; este análisis ha sido propuesto como medio para evaluar características psicológicas de los jugadores de videojuegos (Ventura et al., 2012; Ontanon y Zhu, 2021)

El enfoque tradicional para identificar estilos de juego consiste en desarrollar un modelado del jugador (Desurvire & El-Nasr, 2013), el cual incluye detección, modelado, predicción y expresión de las características del jugador, generando patrones de comportamiento (Hullet et al., 2011). Tras la idea del estilo de jugador se encuentra el concepto de adaptabilidad, es decir, un sistema que automáticamente adapta la experiencia de juego para cada usuario (McCord et al., 2019).

De tal manera que para la adaptación se crea el modelado de usuarios cuya finalidad es la de entender sus necesidades y preferencias. El modelado del usuario se consigue a través de su perfil o mediante información obtenida de la sesión de juego a la que se le aplicada generalmente Inteligencia Artificial (IA) (Hullet et al., 2011; Mayer, 2018; Aggarwal et al., 2022). Para el modelado del jugador se puede utilizar un marco teórico que representa la categorización esperada del jugador, por ejemplo, un modelo emocional. En contraste, cuando no se usa un marco teórico, el modelo del jugador se genera sin suposiciones iniciales; en este enfoque la tendencia es identificar patrones de comportamiento. También se pueden aplicar enfoques híbridos que contienen elementos de ambos enfoques (Ontanon y Zhu, 2021).

Cuando un juego o videojuego conlleva algún otro fin, además de su objetivo inherentemente lúdico, se denomina juego serio. El ámbito de aplicación de los juegos serios es muy amplio, entre los diferentes objetivos podemos encontrar la formación, el aprendizaje, el fomento de comportamientos específicos, la aportación de conocimientos sobre problemas concretos y la determinación de perfiles psicológicos (van Lankveld et al., 2011).

En otro orden de ideas, las competencias personales o no técnicas están estrechamente relacionadas con la forma en que las personas se comportan. Estas, comprenden una combinación de capacidades, actitudes, hábitos y rasgos de personalidad que permiten a las personas tener un mejor rendimiento (García et al., 2020). A diferencia de las habilidades técnicas que se miden en función del desempeño en el campo de interés, las habilidades personales representan un reto tanto para su formalización como para su evaluación (Muzio et al., 2007).

Dentro de la literatura, a nivel individual las habilidades personales incluyen *aspectos psicológicos* como rasgos de personalidad, características, patrones mentales, juicio, prejuicios o motivaciones; *aspectos cognitivos* como la facilidad para aprender o las

preferencias de aprendizaje, el pensamiento abstracto, el intercambio de conocimientos o formación; y *aspectos de gestión* como la planificación o la estimación del esfuerzo para realizar una tarea (Mayoer, 2008). A nivel de equipo se sabe que las habilidades personales influyen en la forma de relacionarse entre sí y en el equipo en su conjunto (Milczarski et al., 2021).

Con base a lo antes mencionado, se propone la observación del comportamiento de los jugadores como medio para el análisis de las habilidades personales. Además, un videojuego es un medio en el que se pueden representar situaciones propicias para que se presenten en los jugadores la habilidad a analizar, por otro lado, en el videojuego que se pueden recolectar datos de manera no invasiva; y los juegos representan entornos virtuales en el que el jugador sea él mismo, sin la sensación de estar siendo observado.

Sin embargo, a pesar de que los videojuegos facilitan la configuración adecuada para comprender características del comportamiento individual y colectivo, su uso como instrumento de evaluación psicológica se encuentra todavía en una fase inicial. Es más, actualmente es muy común que se utilice un test psicológico para validar o corroborar el análisis de los datos obtenidos sobre comportamiento cuando el medio es un videojuego (Aggarwal et al., 2020; Ammannato y Chiesi, 2020; Guardiola y Natkin, 2015). Este artículo presenta una propuesta de un proceso que proporciona los pasos para la selección de videojuegos con la finalidad de analizar habilidades personales de acuerdo a un conjunto de necesidades establecidas, la construcción de este proceso toma como base el diseño de varios autores.

1.1. Habilidades personales en relación con el comportamiento del jugador

Las habilidades personales se clasifican a grandes rasgos en habilidades interpersonales e intrapersonales. Las habilidades intrapersonales son las que se dan hacia el interior de la persona como independencia o autoactualización, mientras que las interpersonales son habilidades sociales particularmente útiles para relacionarnos con los demás como comunicación, trabajo en equipo o resolución de conflictos (Schaap & Dippenaar, 2017; Warner, 2020).

La personalidad y las habilidades personales están estrechamente relacionadas. Como señaló García et al. (2020), el desarrollo de la personalidad es precisamente la adquisición de habilidades para la vida, aquellas habilidades necesarias para vivir con éxito, y las habilidades para la vida son habilidades personales. Por lo tanto, el desarrollo de habilidades personales es un proceso sistemático, consciente y continuo.

Para el estudio de características personales y del comportamiento, los instrumentos más habituales son los basados en auto informes como entrevistas y cuestionarios. Desafortunadamente, el auto informe está propenso a presentar sesgos debido a factores como la deseabilidad social, la tendencia a presentar una imagen favorable o esperada en lugar de la real, la memoria y la motivación (Desurvire y El-Nasr, 2013). Como alternativa, una evaluación situacional representa un enfoque no invasivo de observación y análisis; aquí es donde los videojuegos, escenarios atractivos e inmersivos, generados por computadora, representan un espacio en el que las personas pueden interactuar y tomar decisiones sin sufrir consecuencias. Estas características, asumimos alientan a los

jugadores a ser auténticos (Venture et al., 2012). Además, en un videojuego, el control de las situaciones es igual para cada uno de los participantes.

En el análisis de habilidades personales mediante videojuegos se pueden aplicar las prácticas de identificación del estilo de juego considerando que todos los elementos relacionados con la interacción del jugador en el videojuego determinan su estilo. Este análisis incluye las métricas del juego y las características estadísticas-espacio-temporales; datos que suelen asignarse a estados cognitivos como la atención, el desafío y el compromiso (Acerenza et al., 2009).

La extracción de datos de la interacción del jugador suponemos debe proporcionar información más genuina que la que se recopila mediante cuestionarios tradicionales (Hullet et al., 2011). Para van Lankveld et al. (2016) el comportamiento observado de los *inputs* del jugador en los videojuegos se asemeja al concepto de los informes hechos por un observador, y potencialmente evita los problemas de sesgo antes mencionados.

2. Trabajos relacionados

En esta sección presentamos varias propuestas encontradas en la literatura reciente en las que se diseña un proceso para crear o utilizar videojuegos, con la finalidad de estudiar el comportamiento. De cada una de ellas resaltamos principalmente las diferencias que ponen de manifiesto una nueva característica que será incluida en el proceso del uso de videojuegos como medio para el estudio de habilidades personales.

Guardiola y Stéphane (2015) presentaron una metodología de diseño de juego para generar un perfil psicológico de los jugadores. Su proceso incluye nueve pasos:

1. Identificar el modelo psicológico;
2. Tener en cuenta el público y las limitaciones;
3. Trabajar con expertos para la validación (*Experts pipeline*);
4. Establecer el concepto del juego de acuerdo con el modelo psicológico;
5. Definir los ítems psicométricos e incluirlos en el bucle de juego;
6. Racionalizar la conexión entre el diseño del juego y las dimensiones psicológicas;
7. Construir una interfaz para los resultados;
8. Evaluar la calidad de los ítems; y
9. Realizar un bucle hasta encontrar una fuerte conexión entre los perfiles generados y el cuestionario tradicional.

En este método, los autores destacan la importancia de comparar los resultados del juego con datos fiables como un instrumento psicológico tradicional, un cuestionario. Los datos deben recopilarse durante la sesión de juego y el modelo psicológico debe coincidir con la intención de creación de perfiles, finalmente el resultado debe compararse con medios confiables (i.e. el instrumento psicológico). En este marco la validación por expertos o “Experts pipeline” representa las fases de concepción y producción del video juego. En la fase de concepción se desarrolla el documento de diseño del juego (GDD por sus siglas en inglés) y se realizan evaluaciones técnicas y de expertos. A continuación, la fase de producción determina las especificaciones para llevar a cabo la producción. Posteriormente se realizan pruebas del jugador y pruebas funcionales, y por último, tiene

lugar la revisión de expertos sobre la función psicométrica. Tanto la fase de concepción como la fase de producción siguen un proceso iterativo incremental. Este proceso se aplicó en el proyecto JEU SERAI para dar orientación vocacional a estudiantes, no con la intención de sustituir a un asesor de carrera, sino para ofrecer una prueba de autoevaluación. Este marco fue tomado como base para el proceso aquí presentado.

Alloza et al. (2017) argumentan que los videojuegos comerciales (i.e. no los juegos serios) son una herramienta útil tanto para la identificación de habilidades personales como para su entrenamiento. Su proceso inicia con la identificación de videojuegos de código abierto que garanticen la posibilidad de hacer modificaciones a la aplicación, de tal manera que se puedan incluir indicadores que permitan la evaluación de habilidades personales. Para este efecto, los autores diseñaron un algoritmo de telemetría basado en una revisión literaria. A partir de los videojuegos seleccionados, se decide qué habilidad personal se puede medir. Utilizan un pre-test psicométrico aunado a reconocimiento facial para predecir los estados afectivos de los jugadores. Al final de las sesiones, los jugadores vuelven a contestar el instrumento psicométrico para comparar los resultados y establecer posibles mejoras en las habilidades personales seleccionadas. Los tres pasos siguientes describen el proceso de Allonza et al. (2017):

1. Exploración para la selección de un juego comercial adecuado.
2. Elegir las habilidades personales a mejorar.
3. Validar la mejora de las competencias personales.

El tercer paso, la validación, consiste en una prueba previa y otra posterior para hacer una comparación y establecer la mejoría en la habilidad personal seleccionada.

Consideramos del modelo de Allonza et al. (2017) su propuesta de utilizar videojuegos de código abierto como alternativa a incluir en el proceso propuesto para el análisis de habilidades personales a través de videojuegos. Esta opción debe considerarse para un manejo más eficiente de recursos, aunque no necesariamente limitada al uso de videojuegos de código abierto.

La propuesta presentada por Mayer (2018) tiene como objetivo demostrar la validez del entrenamiento y la evaluación mediante videojuegos. Ellos utilizaron un videojuego comercial, tridimensional y multijugador para el entrenamiento en equipo y la realización de pruebas psicométricas. Su proceso incluye recopilar datos haciendo encuestas previas y posteriores al juego, ambos instrumentos para evaluar las características del equipo a nivel individual a través de los registros del videojuego. Ellos consideran tanto las evaluaciones de equipo como las individuales. Mayer (2018) asegura que la ventaja de los videojuegos es que pueden generar una situación similar a la de la vida real en la que se aplique la evaluación en equipo. De este proyecto consideramos que las evaluaciones individuales en contexto de equipo y las evaluaciones de equipo son opciones a incluir en el análisis de habilidades personales.

McCord et al. (2019) realizaron un estudio para comparar los instrumentos de análisis tradicionales de personalidad con un instrumento similar en un videojuego. Su diseño es un escenario narrativo con opciones para modificar el curso de la narrativa. El juego da al jugador la impresión de que el progreso depende de la opción elegida, sin embargo, éstas solo están relacionada con alguna característica de la personalidad. Los resultados

del test de personalidad tradicional y el videojuego se usan hacer una comparación. Del estudio McCord et al. (2019) se puede destacar la importancia de evitar que el jugador elija opciones, tal vez en contra de su naturaleza, pero con la expectativa de obtener una mejor puntuación en el juego. Este es un punto de diseño que hay que cuidar mucho para validar el análisis.

Pouzevara et al. (2019) describen el proceso de desarrollo y pruebas de un videojuego para la resolución de problemas. Ellos utilizaron un diseño centrado en evidencia (ECD por sus siglas en inglés) como marco para el diseño de tareas relacionadas con los datos generados por el videojuego. ECD utiliza el análisis de dominio, el modelado de dominio y la evaluación conceptual para descomponer la habilidad en partes medibles. Para su propuesta realizaron una revisión bibliográfica y una consulta con expertos con la finalidad de definir un conjunto de habilidades que se priorizaron usando tres criterios: su importancia, el hecho de ser enseñables, y su mensurabilidad. El diseño del videojuego fue un co-desarrollo con grupos focales. El videojuego se validó inicialmente en una prueba de funcionalidad, luego con una prueba de usuario y por último con pruebas de la interfaz de usuario. Para validar el videojuego, se realizó una triangulación entre lo capturado por el juego, el test de personalidad y las autoevaluaciones. También se realizó una prueba a los participantes acerca de las habilidades necesarias para jugar el juego, posteriormente se les preguntó sobre su frecuencia en el uso de juegos digitales para establecer sus habilidades. A continuación, se aplicó el test psicológico para las habilidades personales de interés. Por último, se compararon los instrumentos psicométricos, la autoevaluación y las métricas del juego.

El estudio de Pouzevara et al. (2019) presenta varios puntos importantes tomados como base para el diseño de nuestro proceso. En primer lugar, se pone de manifiesto la importancia de utilizar un marco de diseño para las tareas u objetivos del juego. En segundo lugar, la importancia de dividir las tareas en partes medibles. Además, el videojuego requiere pruebas tanto de funcionalidad como de usuario aunado a los procesos de prueba habituales. Otro punto importante que puede inferirse de su propuesta es que las habilidades del jugador no deben interferir en el resultado del videojuego; a este respecto sugerimos utilizar juegos que no requieran habilidades físicas y en todos los casos establecer que las habilidades requeridas para jugar el videojuego no interfieran con el análisis de las habilidades personales. El último punto relevante es la consideración de una autoevaluación distinta al instrumento psicométrico que también puede apoyar la evaluación.

Zulkifly (2019) describió el proceso de adaptación de un marco teórico de personalidad para el diseño de un videojuego serio. Su método consiste en primer lugar en consultar a expertos en personalidad y diseño de videojuegos con el fin de identificar los ítems del cuestionario asociados al comportamiento que se desea medir. A continuación, se relacionan con las métricas del videojuego que mejor capturen el comportamiento identificado. En este segundo paso, el autor sugiere compilar las mejores ideas en una lista exhaustiva de métricas de juego. Posteriormente se debe establecer la mecánica de juego para implementar las métricas identificadas, debatiendo el cómo encajan mejor las mecánicas en un videojuego cohesionado y enumerando sus posibles combinaciones. A continuación, basándose en los pasos anteriores, determinar las especificaciones del

diseño del juego. Esta propuesta es interesante porque profundiza en el procedimiento que debe seguirse desde los ítems del cuestionario hasta las métricas del videojuego y sus mecánicas.

Ammannato & Chiesi (2020) realizaron un estudio utilizando la información sobre el modo en que el jugador actúa y reacciona en un videojuego competitivo para evaluar los rasgos de personalidad aplicando el aprendizaje automático para extraer datos de los registros del juego. Los autores seleccionaron un modelo de personalidad y habilidades personales (i.e., rasgos de honestidad-humildad y emocionalidad) relacionadas con un juego multiusuario asumiendo que un videojuego incluye el placer de interactuar con otros y la incertidumbre en los resultados. Para este proyecto se utilizó un videojuego MOBA (por sus siglas en inglés *Massive Online Battle Arena*) con dinámicas cooperativas y competitivas. Se aplicó previamente a los participantes un test psicométrico en formato electrónico. Los registros del videojuego se codificaron para realizar el análisis de los datos a través de modelos de aprendizaje profundo basados en las dimensiones psicométricas del instrumento de evaluación de la personalidad. En el estudio preliminar predijeron correctamente los rasgos de personalidad basándose en las acciones de los jugadores en los videojuegos. En esta propuesta se identifican dos enfoques diferentes en los juegos multiusuario, el modo competitivo y el modo cooperativo. Además, para el análisis, extraen datos no solo mediante una comparación de un videojuego psicométrico tradicional, sino también para comprender otros comportamientos ocultos de los jugadores que podrían contribuir a la revelación de la métrica psicológica.

Haizel et al., (2021), presentaron un estudio para establecer si un videojuego puede predecir la personalidad y comparar esta evaluación con métodos tradicionales. Desarrollaron un videojuego de rol, RPG por sus siglas en inglés, con libertad de elección del tipo “camina y resuelve”. Se preguntó a los participantes sobre sus hábitos con los videojuegos. Los participantes jugaron el juego antes de responder a un test de personalidad. En esta propuesta se incluye entender los hábitos de videojuego de los participantes que se incluyeron en la evaluación, esto puede ser visto como parte de las habilidades del video jugador, pero también para entender ciertas tendencias a tener en cuenta durante la evaluación psicométrica.

Peña Pérez Negrón et al. (2022) presentaron un estudio para la evaluación de estilos interactivos, es decir, comportamiento consistente en situaciones con los mismos arreglos, asociado a equipos de desarrollo de software. Su proceso consiste en seleccionar un formato de contingencia de acuerdo con el estilo interactivo a analizar. A continuación, se establecen los valores esperados propios de los miembros del equipo. A partir de los pasos anteriores, se selecciona el videojuego. Por último, describen los datos que definirán la evaluación de los participantes. Ellos utilizaron un pre-cuestionario de autoevaluación sobre las habilidades a evaluar; y un post-cuestionario para validar la percepción del jugador de su estilo interactivo en el videojuego. De esta propuesta se resalta la inclusión de valores esperados para el rendimiento del equipo.

A partir de estas propuestas se presenta a continuación en la Tabla 1, el conjunto de características destacadas que se tomarán como base para la realización de la propuesta de un proceso para el análisis de habilidades personales basado en videojuegos.

Proceso	Características
<i>Guardiola y Stéphane (2015)</i>	Comparación de los datos del videojuego con un instrumento de medición tradicional Marco base para la propuesta
<i>Alloza et al. (2017)</i>	Uso de videojuegos comerciales preferentemente de código abierto
<i>Mayer (2018)</i>	En su caso, evaluación de equipo tanto a nivel individual como a nivel grupal
<i>McCord et al. (2019)</i>	Evitar opciones que vayan en contra de la naturaleza del jugador y que sean tomadas con la finalidad de ganar el juego
<i>Pouzevara et al. (2019)</i>	• Utilización de un marco de diseño de tareas (e.g. uso de ECD) • Descomposición de las tareas en partes medibles • Realización de pruebas de funcionalidad, usuario y de interfaz, además de las pruebas habituales del desarrollo de software. • Asegurar que las habilidades del jugador no interfieran en el resultado del videojuego. • Uso de la autoevaluación además del instrumento psicométrico. Triangulación entre resultados del videojuego, el test psicométrico y la autoevaluación.
<i>Zulkifly (2019)</i>	Uso de un procedimiento para relacionar los ítems de un instrumento psicométrico con las métricas y mecánicas del videojuego.
<i>Ammannato & Chiesi (2020)</i>	• Establecimiento de un modelo competitivo y cooperativo, cuando se trate de valoración en equipo. • Análisis basado en métodos de IA para la evaluación de rasgos de personalidad mediante todas las entradas del jugador en el videojuego.
<i>Haizel et al., (2021)</i>	Considerar en el perfil del jugador y su análisis de comportamiento, sus hábitos de juego.
<i>Peña Pérez Negrón et al. (2022)</i>	Valores esperados en las pruebas psicométricas para los participantes del videojuego.

Tabla 1 – Características de diferentes autores para el proceso propuesto

3. Proceso para el uso de videojuegos en el análisis de habilidades personales

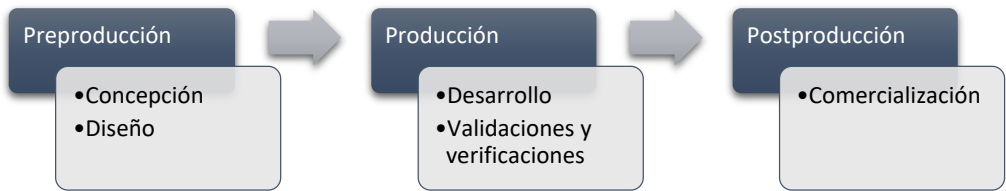


Figura 1 – Fases del ciclo de vida del desarrollo de videojuegos (Aleem et al., 2016)

El ciclo de vida de los videojuegos se considera diferente al proceso tradicional de desarrollo de software. Con base en una revisión sistemática de literatura Aleem et a. (2016) encontraron que el ciclo de vida del desarrollo de videojuegos incluye tres etapas,

ver Figura 1. En el ciclo de vida del videojuego, el proceso de desarrollo tiene lugar en la etapa de Producción, en contraste con el software tradicional que durante el proceso de producción el software se ejecuta y está listo para su uso, para los videojuegos en la etapa de Preproducción, se prueba la factibilidad de los escenarios, esta etapa está compuesta de la Concepción y el Diseño que se explican más adelante y la Etapa de Postproducción aborda las pruebas, mercadotecnia y publicidad.

3.1. Fase de Preproducción

El propósito de la fase de Preproducción es realizar la Concepción y el Diseño del juego. Las actividades de la Concepción y se describen a continuación:

1. *Identificar la habilidad personal a evaluar.* Esta actividad inicial será la referencia para establecer las características del video juego. A diferencia del modelo de Guardiola y Natkin (2015) que inicia con la identificación del modelo psicológico debido a que su intención es generar un perfil psicológico. En nuestro caso, la intención es la evaluación de habilidades personales, exista o no un instrumento psicológico para su evaluación. De tal manera que el primer paso lo constituye establecer la habilidad personal a evaluar.
2. *Establecer si la habilidad personal es intrapersonal o interpersonal.* Esta diferenciación es importante principalmente para establecer si es necesario un juego multijugador.
3. *Definir el nivel de medidas.* En esta actividad se decide si se harán medidas individuales o grupales de acuerdo a lo expuesto por Mayer (2018). En caso de tratarse de actividades grupales, se debe determinar si el juego tendrá actividades competitivas o colaborativas, o ambas.
4. *Identificar la audiencia.* En esta actividad se establece el conjunto de restricciones vinculadas al tipo de público y al objetivo útil del juego (Guardiola et al., 2005). También, es importante establecer el perfil del jugador, sus hábitos de videojuego que pudieran tener implicaciones en la evaluación de la habilidad personal (Ammannato y Chiesi, 2020).
5. *Establecer la situación o situaciones.* En esta actividad se identifican las situaciones o arreglo de contingencias en que es más probable que se presente la habilidad personal a evaluar, esto nos dará un contexto situacional en el video juego.

La fase de Pre-producción también contempla el Diseño, para ésta las tareas a realizar a continuación se describen.

1. *Determinar el marco para el diseño de las tareas.* Basado en el proceso de Pouezevara et al. (2019) esta actividad selecciona un marco de referencia para establecer las tareas que tienen relación directa con la habilidad personal a evaluar.
2. *Establecer objetivos de las tareas.* Aquí se establecen los objetivos para cada una de las tareas conforme a las métricas de la habilidad personal a evaluar. En caso de utilizar como base un instrumento psicológico conocido, se debe utilizar un procedimiento para relacionar los ítems del instrumento con métricas del juego (Zulkifly, 2019).

3. *Descomposición de tareas.* Las tareas deben descomponerse a un nivel en el que puedan ser medibles.
4. *Revisión de habilidades para realizar las tareas del videojuego.* Asegurarse que las tareas no involucren habilidades de video jugadores expertos. Además, es muy importante que las opciones del jugador no lo hagan ir contra su naturaleza para ganar el juego (McCord et al. (2019).
5. *Definición de elementos de juego.* Establecer las mecánicas, dinámicas y estéticas del video juego (Peña Pérez Negrón et al., 2021)

La realización de las actividades de Concepción y Diseño, proporcionan los insumos para la elaboración del Documento de Diseño del Juego (GDD por sus siglas en inglés). La realización de este documento es el criterio de salida de esta fase para continuar a la siguiente fase que es la Producción.

3.2. Fase de Producción

El propósito de la fase de producción es la selección o el desarrollo del videojuego, ver Figura 2.



Figura 2 – Selección de desarrollo.

Estos pasos se describen a continuación:

1. *Seleccionar tipo de juego a utilizar.* Siguiendo a Alloza et al. (2017) es importante verificar si es posible utilizar un juego comercial ya desarrollado o, en caso contrario, iniciamos el desarrollo.
2. *Revisión de ajustes del juego.* En caso de que tengamos un videojuego adecuado para hacer las prueba, debemos verificar si será necesaria su modificación para la recolección de datos o bien, si debemos buscar otros mecanismos que nos permitan capturar las interacciones del usuario que sean de interés para establecer su perfil psicológico conforme a la habilidad a observar.
3. *Desarrollo de juego.* Para llevar cabo el desarrollo del juego, en el caso en que no exista un juego comercial, se recomiendan dos métodos propuestos:

- a. Seguir cinco fases basado en la recomendación de SCRUM y su adaptación para video juegos (Acerenza et al., 2009): (1) Fase Conceptos en la que se especifican los elementos básicos para el diseño, esto es, lo que en nuestro proceso se corresponde al Pre-proceso; (2) Fase Planificación en la que se hace un cronograma de actividades funcionales y no funcionales, estableciendo prioridades, se crean los equipos conforme a los requerimientos técnicos; (3) Fase Elaboración en la que se hace la implementación; (4) Fase Beta con el objetivo de evaluar y hacer ajustes; y (5) Fase Cierre en la que se concluye el proceso y se obtiene retroalimentación para futuros proyectos
- b. Seguir las tres fases sugeridas por Hernández Rentería et al. (2019) siguiendo una metodología iterativa e incremental: (1) Diseño del concepto, en el que se definen los retos mediante un estado del arte de juegos con las mismas características; (2) Diseño del juego, en el que se hace un script ejecutivo, se identifican los verbos de acción, se hace una tormenta de ideas que derivan en una versión en papel; y (3) Desarrollo de un prototipo para establecer la jugabilidad, antes de pasar a la Implementación propiamente dicha.

En nuestro caso consideramos que el prototipo en papel es una buena opción para establecer no solo la jugabilidad sino también las economías del juego, este paso debe hacerse solo si se determina llevar a cabo el desarrollo y es conveniente desarrollar el prototipo antes de iniciar la implementación.

4. *Validaciones y verificaciones.* Implementar actividades de verificación y validación a cada una de las partes del desarrollo, siguiendo buenas prácticas de Calidad de Ingeniería del Software. Pouezevara et al. (2019) recomienda la realización de pruebas de funcionalidad, de usuario y de interfaz, además de las pruebas habituales en el desarrollo de software.
5. Establecer los datos que se recogerán del video juego. Estos datos forman la base para establecer la evaluación de la habilidad personal. Los datos del juego pueden utilizar una categorización pre establecida que nos permita clasificar a los jugadores conforme a un modelo determinado (Peña Pérez Negrón, 2022).

3.3. Fase de Postproducción

Típicamente, el propósito de esta fase es realizar la comercialización del videojuego. En nuestro caso la fase está enfocada en la validación de los resultados del análisis de habilidades personales mediante los datos recolectados, especialmente cuando se trate de una primera aproximación para medir la habilidad personal. Para la validación de los resultados del video juego, se sugiere la aplicación de alguna o varias de las siguientes técnicas, las cuales se muestran a continuación ordenadas por orden de importancia:

1. Instrumento psicológico reconocido: aplicar un instrumento psicológico reconocido, preferentemente una cualitativa que permita contrastar resultados con los obtenidos en el videojuego.
2. Juicio de uno o varios expertos: invitar a un conjunto de expertos que revisen los resultados, además estos expertos pueden conducir entrevistas de ser necesario. Los juicios de experto, sin embargo, pueden ser subjetivos y menos estandarizados que los instrumentos psicológicos.

3. Auto-evaluación del jugador: el mismo jugador puede expresar su acuerdo o desacuerdo con el perfil generado por el videojuego. Esta medida suele ser aún más subjetiva que un juicio experto, pero sigue siendo válida.
4. Recolección de datos y aplicación de técnicas IA: esta evaluación puede considerarse más objetiva que el juicio de experto o la auto-evaluación, aunque la técnica será tan buena como la cantidad y calidad de los datos que podamos proporcionarle, además de los sesgos que puedan presentar.

4. Conclusiones y trabajo futuro

Las habilidades personales representan un reto en cuanto a su evaluación. Estas habilidades actualmente presentan un factor clave para el éxito tanto en la vida personal como profesional. Existen diferentes esfuerzos e iniciativas con metodologías para su evaluación mediante el uso de video juegos, debido a las ventajas que presentan los video juegos comparados con otras alternativas para medir características personales.

En específico el uso de video juegos permite recrear las situaciones en un contexto situacional en el que se presentan dichas habilidades para la resolución de ciertas tareas no siempre relacionadas con las actividades diarias de la persona. En este artículo se presenta un proceso para el uso de video juegos para medir habilidades personales desde su concepción y hasta el análisis de los datos obtenidos. Cabe resaltar que este proceso no ha sido probado, por lo tanto, como trabajo futuro, se contempla la aplicación de dicho proceso en un estudio exploratorio para determinar el alcance real de su uso.

Referencias

- Abdul-Kadir, M. R., & Price, A. D. F. (1995). Conceptual phase of construction projects. *International Journal of Project Management*, 13(6), 387–393. [https://doi.org/10.1016/0263-7863\(96\)81776-5](https://doi.org/10.1016/0263-7863(96)81776-5)
- Acerenza, N., Coppes, A., Mesa, G., Viera, A., Fernandez, E., Laurenzo, T., & Vallespir, D. (2009). Una Metodología para Desarrollo de Videojuegos. Presented at the 38o JAIIO - Simposio Argentino de Ingeniería de Software. https://www.fing.edu.uy/sites/default/files/biblio/22811/asse_2009_16.pdf
- Aggarwal, S., Saluja, S., Gambhir, V., Gupta, S., & Satia, S. P. S. (2020). Predicting likelihood of psychological disorders in PlayerUnknown's Battlegrounds (PUBG) players from Asian countries using supervised machine learning. *Addictive Behaviors*, 101, 106132. <https://doi.org/10.1016/j.addbeh.2019.106132>
- Aleem, S., Capretz, L. F., & Ahmed, F. (2016). Game development software engineering process life cycle: a systematic review. *Journal of Software Engineering Research and Development*, 4(1), 1-30. <https://doi.org/10.1186/s40411-016-0032-7>
- Alloza, S., Escribano, F., Delgado, S., Corneanu, C., & Escalera, S. (2017). XBadges. Identifying and training soft skills with commercial video games. <https://doi.org/10.48550/ARXIV.1707.00863>

- Ammannato, G., & Chiesi, F. (2020). Playing With Networks: Using Video Games as a Psychological Assessment Tool. *European Journal of Psychological Assessment*, 36(6), 973–980. <https://doi.org/10.1027/1015-5759/a000608>
- Bonilla Carranza, D., Peña Pérez, N., A., & Contreras, M. (2021). Videogame development training approach: A Virtual Reality and open-source perspective. *JUCS - Journal of Universal Computer Science*, 27(2), 152–169. <https://doi.org/10.3897/jucs.65164>
- Dhaouadi, S., & Ben Khelifa, M. M. (2020). A multimodal Physiological-Based Stress Recognition: Deep Learning Models' Evaluation in Gamers' Monitoring Application. In 2020 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP) (pp. 1–6). <https://doi.org/10.1109/ATSIP49331.2020.9231666>
- Desurvire, H., & El-Nasr, M. S. (2013). Methods for Game User Research: Studying Player Behavior to Enhance Game Design. *IEEE Computer Graphics and Applications*, 33(4), 82–87. <https://doi.org/10.1109/MCG.2013.61>
- Ezzati, R.. (2017). Gamers' experiences in playing video games – A Theoretical Thematic Analysis. <https://doi.org/10.13140/RG.2.2.22549.37608>
- García, I., Pacheco, C., Méndez, F., & Calvo-Manzano, J. A. (2020). The effects of game-based learning in the acquisition of “soft skills” on undergraduate software engineering courses: A systematic literature review. *Computer Applications in Engineering Education*, 28(5), 1327–1354. <https://doi.org/10.1002/cae.22304>
- Guardiola, E., & Natkin, S. (2015). A Game Design Methodology for Generating a Psychological Profile of Players. In C. S. Loh, Y. Sheng, & D. Ifenthaler (Eds.), *Serious Games Analytics* (pp. 363–380). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-05834-4_16
- Haizel, P., Vernanda, G., Wawolangi, K. A., & Hanafiah, N. (2021). Personality Assessment Video Game Based on The Five-Factor Model. *Procedia Computer Science*, 179, 566–573. <https://doi.org/10.1016/j.procs.2021.01.041>
- Hernández Rentería, L. A., Ramírez Lujano, A., Contreras Davito, M., Carranza, D. B., & Peña Pérez Negrón, A. (2019). A Serious Game Proposal to Reinforce Reading Comprehension in Scholars. In *Human-Computer Interaction: 5th Iberoamerican Workshop, HCI-Collab 2019, Puebla, Mexico, June 19–21, 2019, Revised Selected Papers 5* (pp. 30–41). Springer International Publishing.
- Hullett, K., Nagappan, N., Schuh, E., & Hopson, J. (2011). Data analytics for game development (NIER track). In *Proceedings of the 33rd International Conference on Software Engineering* (pp. 940–943). <https://doi.org/10.1145/1985793.1985952>
- Mayer, I. (2018). Assessment of Teams in a Digital Game Environment. *Simulation & Gaming*, 49(6), 602–619. <https://doi.org/10.1177/1046878118770831>
- McCord, J.-L., Harman, J. L., & Purl, J. (2019). Game-like personality testing: An emerging mode of personality assessment. *Personality and Individual Differences*, 143, 95–102. <https://doi.org/10.1016/j.paid.2019.02.017>

- Mildner, P., & Floyd Mueller, F. (2016). Design of Serious Games. In R. Dörner, S. Göbel, W. Effelsberg, & J. Wiemeyer (Eds.), *Serious Games* (pp. 57–82). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-40612-1_3
- Milczarski, P., Podlaski, K., Hlobaž, A., Dowdall, S., Stawska, Z., & O'Reilly, D. (2021). Soft Skills Development in Computer Science Students via Multinational and Multidisciplinary GameDev Project. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education* (pp. 583–589). <https://doi.org/10.1145/3408877.3432522>
- Muzio, E., Fisher, D. J., Thomas, E. R., & Peters, V. (2007). Soft Skills Quantification (SSQ) Foi Project Manager Competencies. *Project Management Journal*, 38(2), 30–38. <https://doi.org/10.1177/875697280703800204>
- Ontanon, S., & Zhu, J. (2021). The Personalization Paradox: the Conflict between Accurate User Models and Personalized Adaptive Systems. In *26th International Conference on Intelligent User Interfaces* (pp. 64–66). <https://doi.org/10.1145/3397482.3450734>
- Peña Pérez Negrón, A., Bonilla Carranza, D., & Berumen Mora, J. (2021). Cutting-Edge Technology for Video Games. In *New Perspectives in Software Engineering: Proceedings of the 9th International Conference on Software Process Improvement (CIMPS 2020)* (pp. 291–304). Springer International Publishing.
- Peña Pérez Negrón, A., Muñoz, M., Bonilla Carranza, D., & Rangel, N. (2022). Towards the Evaluation of Relevant Interaction Styles for Software Developers. In J. Mejia, M. Muñoz, Á. Rocha, H. Avila-George, & G. M. Martinez-Aguilar (Eds.), *New Perspectives in Software Engineering* (Vol. 1416, pp. 137–149). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-89909-7_11
- Pouzevara, S., Powers, S., Moore, G., Strigel, C., & McKnight, K. (2019). Assessing Soft Skills In Youth Through Digital Games. Presented at the 12th annual International Conference of Education, Research and Innovation, Seville, Spain. <https://doi.org/10.21125/iceri.2019.0774>
- Saputra, R., Iqbal, B. M., & Komarudin, A. (2017). Stress Emotion Evaluation in Multiplayer Online Battle Arena (MOBA) Video Game Related to Gaming Rules Using Electroencephalogram (EEG). In *Proceedings of the 2017 4th International Conference on Biomedical and Bioinformatics Engineering* (pp. 74–77). <https://doi.org/10.1145/3168776.3168797>
- Schaap, P., Dippenaar, M. (2017). The impact of coaching on the emotional and social intelligence competencies of leaders. *South African Journal of Economic and Management Sciences*, 20(1), 1–16 <https://doi.org/10.4102/sajems.v20i1.1460>
- van Lankveld, G., Spronck, P., van den Herik, J., & Arntz, A. (2011). Games as personality profiling tools. Presented at the 2011 IEEE Conference on Computational Intelligence and Games (CIG), Seoul, Korea (South): IEEE. <https://doi.org/10.1109/CIG.2011.6032007>

- Ventura, M., Shute, V., & Kim, Y. J. (2012). Video gameplay, personality and academic performance. *Computers & Education*, 58(4), 1260-1266. <https://doi.org/10.1016/j.compedu.2011.11.022>
- Warner, LH (2020). Developing interpersonal skills of evaluators: a service-learning approach. *Am J Eval*, 41:432–451. <https://doi.org/10.1177/1098214019886064>
- Zulkifly, A. N. (2019). Personality assessment through the use of video games (Doctoral dissertation, University of Tasmania).

Automatización de los procesos del estándar ISO/IEC 29110 a través de la Adopción de DevOps

Josefina García¹, J Jesús Minero¹, Elvia Lara¹

josefina.garcia@itsn.edu.mx; jesus.minero@itsn.edu.mx; elvia.lara@itsn.edu.mx

¹ Instituto Tecnológico Superior de Nochistlán, 99900, Nochistlán Zacatecas, México.

DOI: 10.17013/risti.49.37-51

Resumen: Uno de los principales problemas a los que se enfrentan los centros de desarrollo de software certificados en el estándar ISO/IEC 29110 es como automatizar sus procesos para incrementar su productividad. DevOps aparece como una solución para este problema a través de su cultura, prácticas y herramientas con el objetivo de ayudar a las empresas a desarrollar productos y servicios de software a través de la automatización de tareas lo que permite mejorar el tiempo de entrega incrementando la competitividad de las empresas de software. El objetivo de este artículo es demostrar cómo la adopción de DevOps permite la automatización de los procesos del estándar ISO/IEC 29110 del perfil básico a través de un caso de estudio en el Centro de Desarrollo de Software del Instituto Tecnológico Superior de Nochistlán. Los resultados obtenidos muestran el 78% de los productos de trabajo de los procesos del estándar automatizados incrementando la productividad del equipo y reduciendo el tiempo de entrega del software.

Palabras-clave: DevOps; ISO/IEC 29110; automatización; productos de trabajo.

Automating ISO/IEC 29110 processes through DevOps Adoption

Abstract: One of the main problems software development centers certified in the ISO/IEC 20110 standard face is how to automate their processes to increase their productivity. DevOps appears as a solution to this issue through its culture, practices and tools with the aim of helping companies to develop software products and services through the automation of tasks, improving software delivery time and increasing the competitiveness of software development companies. The objective of this article is to demonstrate how the adoption of DevOps allows the automation of the processes of the ISO/IEC 29110 standard of the basic profile through a case study at the Software Development Center of the Higher Technological Institute of Nochistlán. The results obtained show 78% of the work products of the standard processes automated, increasing team productivity and reducing software delivery time.

Keywords: DevOps; ISO/IEC 29110; automation; work products.

1. Introducción

Las pequeñas organizaciones de desarrollo de software, constituyen más del 80% de las empresas en los países. Sus características principales son: cuentan con pocos recursos humanos y capital, la mayoría no tiene un estándar de desarrollo implementado, para mantenerse en el mercado aceptan cualquier trabajo notándose la falta de especialización y provocando que las estimaciones de tiempo y costo no sean reales. Se ha identificado que el esfuerzo necesario para la implementación de estándares más usados por las pequeñas organizaciones como CMMI DEV, ISO/IEC 12207, MPS.BR, Competisoft y MoProSoft no es atractivo por no adaptarse a sus necesidades. Es entonces cuando aparece el estándar ISO/IEC 29110 que está enfocado a las pequeñas organizaciones de hasta 25 personas, y dado su análisis es más adecuado para las pequeñas y medianas empresas de desarrollo de software porque está compuesto por solo dos áreas de proceso en su perfil básico, el proceso de Gestión del Proyecto (GP) y el proceso de Implementación de Software (IS) (Urresti, Ronald, Rodríguez, & de la Caridad, 2019).

Aun cuando el estándar ISO/IEC 29110 está siendo adoptado por las pequeñas organizaciones de desarrollo de software, existe una falta de herramientas que permitan automatizar su adopción (Mejía, Bonilla, & Villanueva Rosas, 2020). DevOps es la combinación de las palabras Desarrollo (Dev) y Operaciones (Ops), es un término que, a pesar de existir hace diez años, actualmente está tomando mayor fuerza (Guerrero, Zúñiga, Certuche Díaz, & Pardo Calvache, 2019). DevOps automatiza tareas de creación, despliegue de software, actividades de integración, pruebas, construcción y entrega de software; reduce demoras y riesgos debido a que tales actividades requieren mucho tiempo y es propenso a errores si se hace manualmente (Mohsienuddin Mohammad, 2018).

En este artículo se muestra como la adopción de DevOps permite la automatización de los productos de trabajo del estándar ISO/IEC 29110 del perfil básico a través de un caso de estudio en el Centro de Desarrollo de Software del Instituto Tecnológico Superior de Nochistlán (ITSN).

Este artículo está estructurado de la siguiente manera: la sección 2 describe el trabajo relacionado; la sección 3 describe el caso de estudio y la sección 4 muestra las conclusiones.

2. Trabajo relacionado

En esta sección se presentan trabajos relacionados que abordan la temática de la adopción de DevOps en organizaciones de desarrollo de software.

Lwakatare, Kilamo, Karvonen, & Sauvola (2019) realizaron un estudio exploratorio de cómo se implementa DevOps en pequeñas y medianas empresas de desarrollo de software. Se hizo un estudio de casos múltiples en cinco contextos diferentes de desarrollo con implementaciones de DevOps, los beneficios que se obtuvieron fueron lanzamientos rápidos y errores mínimos de implementación. Los datos se

recopilaron a través de entrevistas con 26 profesionales y observaciones realizadas en las empresas.

Guerrero, Zúñiga, Certuche y Pardo (2020) realizaron un mapeo sistemático con el objetivo de analizar el conocimiento actual sobre DevOps, los autores observaron que no hay una definición común del significado y que algunas empresas de desarrollo de software han adoptado ciertas prácticas relacionadas con el mismo, sin embargo, mencionan que es necesario desarrollar una guía o estándar que facilite el proceso de adopción de las prácticas de DevOps.

Ghantous y Gill (2017) hicieron una revisión de la literatura en la que analizaron las herramientas, beneficios y retos de las prácticas de DevOps, concluyeron que las empresas de desarrollo de software tienen un interés en la adopción de DevOps pero no tienen la comprensión de conceptos, prácticas, herramientas, beneficios y desafíos.

Bucena y Kirikova (2017) diseñaron y validaron un método de adopción de DevOps con el fin de ayudar a las pequeñas empresas a simplificar el proceso de adoptarlo, además hicieron la validación de un método con la ayuda de un equipo de Tecnologías de la Información de una compañía. Proponen un método con una lista de prácticas priorizadas y herramientas que soportan estas prácticas.

Sánchez, Martínez, Quesada y Jenkins (2020) realizaron un mapeo sistemático de literatura del periodo 2015-2019 donde se analizaron 42 artículos. Se logró identificar y clasificar un total de 20 prácticas de DevOps, 18 criterios para evaluarlas, 16 beneficios y 19 retos asociados a su adopción. Los resultados muestran que las prácticas de soporte son las que tienen mayor tendencia de uso en las organizaciones. Otro descubrimiento importante es que las organizaciones reportan el aumento en la calidad del software como el principal beneficio de DevOps. Respecto a los beneficios de adoptar DevOps, se identificó que las organizaciones han aumentado la calidad del software que producen. Sin embargo, el principal reto que enfrenta la adopción de DevOps es el factor humano.

3. Caso de estudio

El caso de estudio que se presenta a continuación fue realizado en el Centro de Desarrollo de Software (CEDESOFTE) del ITSN, el cual obtuvo su certificación en el perfil básico del estándar ISO/IEC 29110 en marzo de 2019, este perfil está compuesto por dos procesos, el proceso de GP y el proceso de IS. Su misión es generar productos de software de calidad a través del cumplimiento de estándares internacionales para el desarrollo de software para contribuir al desarrollo económico de las regiones, altos de Jalisco y sur de Zacatecas, y su visión es ser un Centro de Desarrollo de Software reconocido por la calidad de sus productos a nivel internacional.

El CEDESOFTE es representado por un Encargado, y éste asigna a un Gestor del proyecto (Miembro de la Academia de Ingeniería en Sistemas Computacionales) cuando llega una solicitud de proyecto, considerando el perfil y las necesidades del proyecto. El Gestor

del proyecto es quien asigna al Líder Técnico y equipo de trabajo para el desarrollo del proyecto.

El CEDESOFTE tiene definidos y documentados los procesos que contiene el perfil básico GP e IS. La Figura 1, muestra el flujo de trabajo de las principales actividades de los procesos.

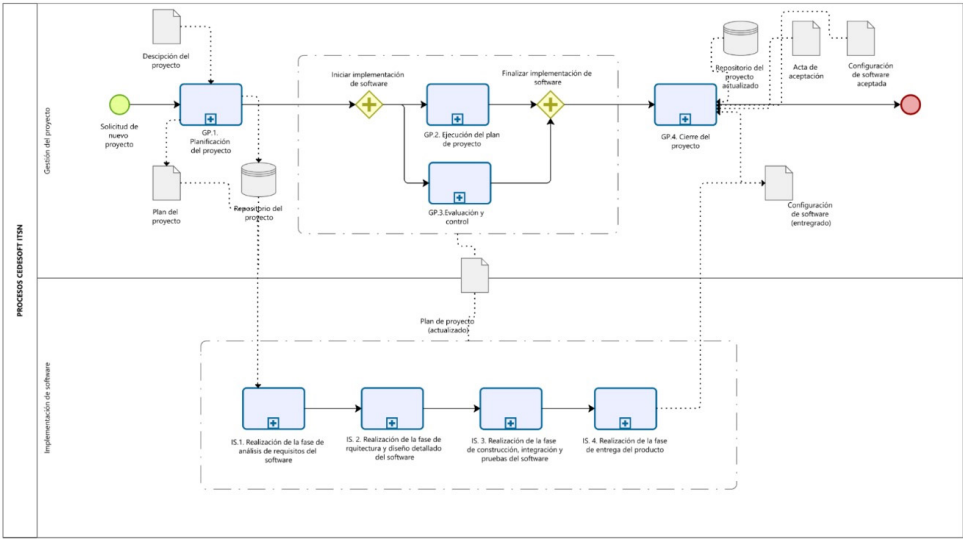


Figura 1 – Flujo de actividades de los procesos

El CEDESOFTE adoptó DevOps para la automatización de los productos de trabajo generados durante la ejecución de los procesos del estándar con el propósito de lograr una mejora continua, disminuir el tiempo de entrega y aumentar la colaboración y la productividad de los equipos.

El presente caso de estudio muestra la configuración de los procesos actuales del CEDESOFTE en Azure DevOps, después se muestra la ejecución de un proyecto en el proceso heredado y por último se presentan los resultados.

A. Configuración de los procesos actuales del CEDESOFTE en Azure DevOps.

Para configurar los procesos en Azure DevOps se siguieron los pasos que se muestran en la Figura 2.

Paso 1. Identificar los productos de trabajo de los procesos. Se inició con un análisis de los procesos del CEDESOFTE para identificar los productos de trabajo, además de las herramientas utilizadas para su creación. La Tabla 1 y Tabla 2 muestran los productos y herramientas de los procesos.

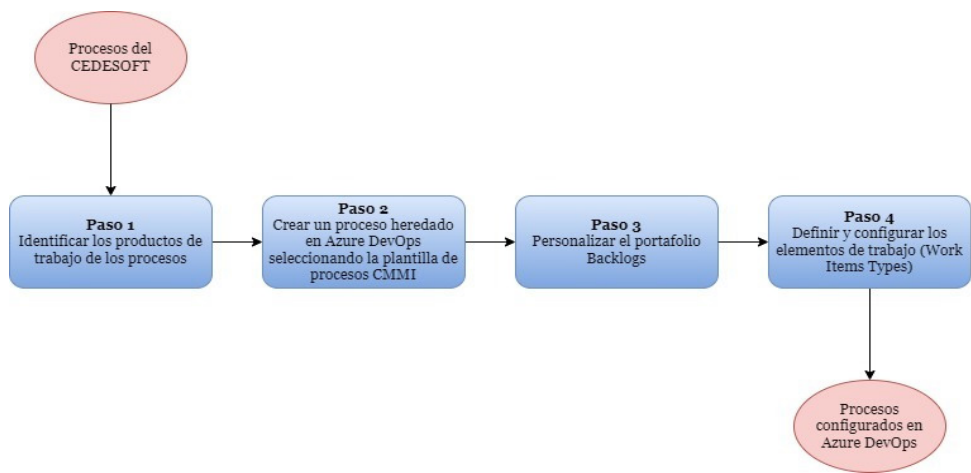


Figura 2 – Pasos para la configuración de los procesos

Producto de trabajo	Herramienta
Plan de proyecto	Archivo de Texto
Cronograma de actividades	Microsoft Project
Plan de Adquisiciones, capacitaciones y costos	Hoja de cálculo
Riesgos	Hoja de cálculo
Lista de verificación del Plan de Proyecto	Archivo de Texto
Lista de validación del Plan de Proyecto	Archivo de Texto
Repositorio del proyecto	Google Drive
Reporte de avance	Archivo de Texto
Acta de reunión	Archivo de Texto
Acciones correctivas	Archivo de Texto
Acta de aceptación	Archivo de Texto

Tabla 1 – Productos de trabajo y herramientas del proceso GP

Producto de trabajo	Herramienta
Especificación de requisitos	Archivo de Texto
Lista de verificación de requisitos	Archivo de Texto
Lista de validación de requisitos	Archivo de Texto
Matriz de trazabilidad	Hoja de cálculo
Diseño de software	Archivo de Texto
Lista De Verificación para Revisión de Diseño de Software	Archivo de Texto

Producto de trabajo	Herramienta
Casos de Prueba, Procedimientos de Prueba y Reporte de prueba	Hoja de cálculo
Lista de validación del software	Archivo de Texto
Manual de Operación	Archivo de Texto
Manual de Usuario	Archivo de Texto
Manual de Mantenimiento	Archivo de Texto
Lista de Verificación para la Revisión de los manuales de Operación, Mantenimiento y de Usuario	Archivo de Texto

Tabla 2 – Productos de trabajo y herramientas del proceso ISO

Paso 2. Crear un proceso heredado en Azure DevOps seleccionando la plantilla de procesos CMMI. Después de identificar los productos de trabajo de los procesos del CEDESOFTE, se hizo la creación de un proceso heredado en Azure DevOps seleccionando la plantilla de procesos CMMI, esta plantilla se recomienda cuando el equipo sigue métodos de proyecto más formales que requieran un marco para la mejora de procesos y un registro auditable de decisiones. Con este proceso se puede realizar un seguimiento de los requisitos, las solicitudes de cambio, los riesgos y las revisiones. Todos los proyectos que utilizan un proceso heredado obtienen las personalizaciones realizadas durante su configuración. Por otro lado, es posible configurar las herramientas ágiles (Backlogs, Sprints, Kanban boards y Taskboard) para cada equipo. La Figura 3 muestra las plantillas de procesos que ofrece Azure DevOps y la forma como se crea un proceso heredado.

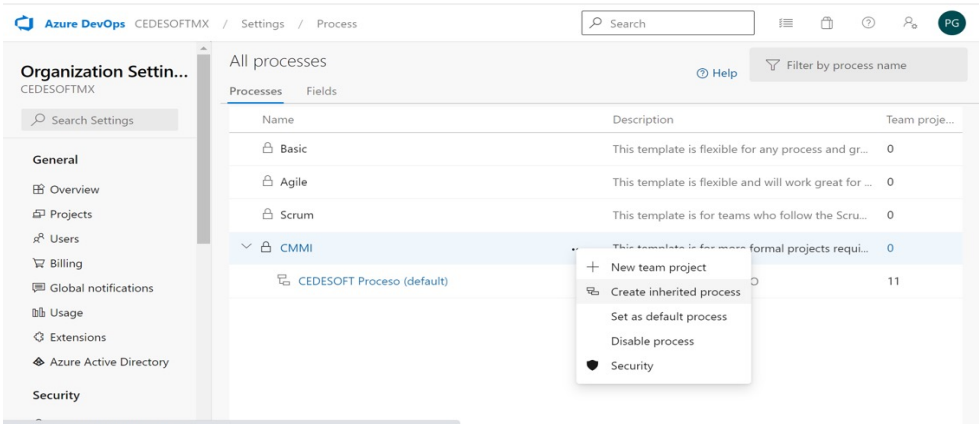


Figura 3 – Plantillas de proceso y creación de un proceso heredado

Paso 3. Personalizar el portafolio Backlogs. Este portafolio provee una forma de agrupar elementos relacionados dentro de una estructura jerárquica, se puede renombrar y editar a cualquier nivel del portafolio, como se muestra en Figura 4.

Add portfolio backlog

The following fields are automatically added to all work item types on the Portfolio backlogs: Stack Rank

Name

Activity

Work item types on this backlog level

☐ Feature construction

☐ Issue

+ New work item type

Default work item type

Save Cancel

Figura 4 – Edición de un nivel del portafolio backlogs

En la Figura 5 se muestran los niveles personalizados del backlog en el proceso heredado con sus diferentes tipos de elementos de trabajo (Work Items Types).

All processes > CEDESOFTE Proceso

Work item types Backlog levels Projects

Portfolio backlogs

Portfolio backlogs provide a way to group related items into a hierarchical structure. You can ren

+ New top level portfolio backlog

Backlog	Work item types
Proyecto	Proyecto (default)
Epic	Epic (default)
Activity	Activity
	Feature (default)

Figura 5 – Niveles personalizados del Backlog

Paso 4. Definir y configurar los elementos de trabajo. Se definieron y configuraron los siguientes elementos de trabajo: Activity Task, Backup, Change Request, Corrective Actions, Report, Review y Risk. La Figura 6 muestra algunos de los elementos de trabajo creados y configurados.

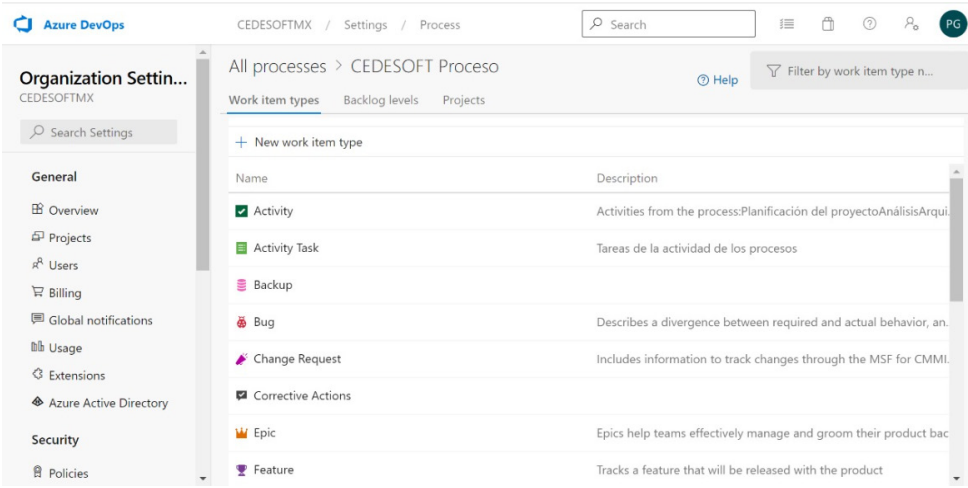


Figura 6 – Elementos de trabajo personalizados

B. Ejecución de un proyecto en el proceso heredado

La Figura 7 muestra los pasos para la ejecución de un proyecto en el proceso heredado. El inicio de un proyecto se realiza a través de una solicitud del proyecto y termina con una carta de aceptación de la entrega del software.

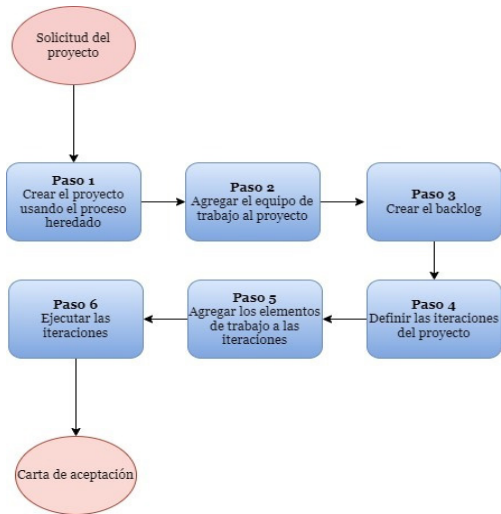
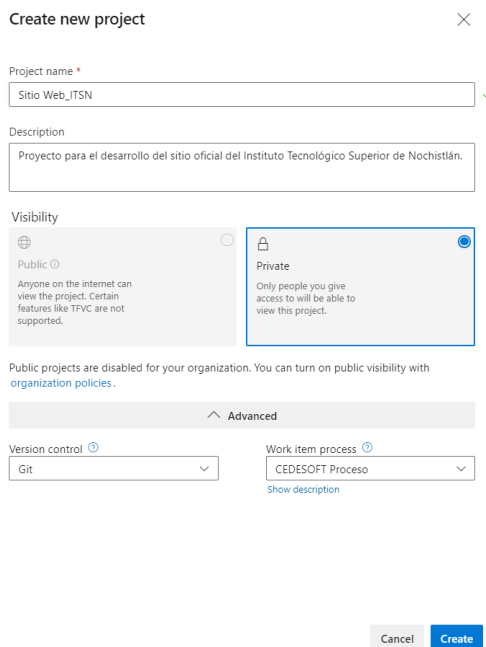


Figura 7 – Pasos para la ejecución de un proyecto en el proceso configurado

Paso 1. Crear el proyecto usando el proceso heredado. Una vez configurados los procesos del Centro de Desarrollo, se creó un proyecto en el proceso heredado con el nombre Sitio Web_ITSN. La Figura 8 muestra la creación del proyecto.



Create new project ✕

Project name *
 ✓

Description

Visibility

☐

Public ⓘ
 Anyone on the internet can view the project. Certain features like TFVC are not supported.

☒

Private
 Only people you give access to will be able to view this project.

Public projects are disabled for your organization. You can turn on public visibility with [organization policies](#).

^ Advanced

Version control ⓘ

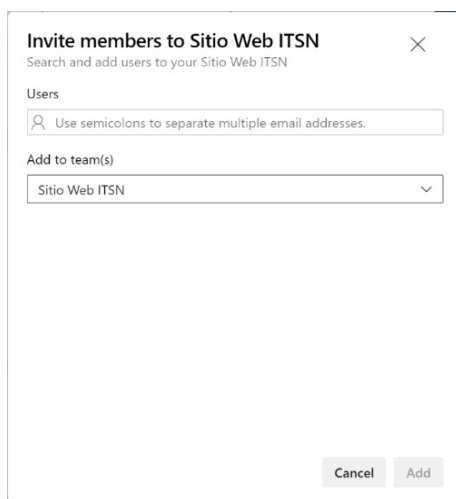
Work item process ⓘ

[Show description](#)

Cancel Create

Figura 8 – Creación del proyecto en Azure DevOps

El proyecto fue un Sitio Web para el ITSN para que los estudiantes, docentes y administrativos puedan consultar información de su interés y para que los aspirantes de nuevo ingreso conozcan la oferta educativa de la institución.



Invite members to Sitio Web ITSN ✕

Search and add users to your Sitio Web ITSN

Users

Add to team(s)

Cancel Add

Figura 9 – Ventana para invitar miembros al proyecto

Paso 2. Agregar el equipo de trabajo al proyecto. Después de crear el proyecto se agregaron los miembros del equipo de trabajo al proyecto para realizar la planificación y darle seguimiento. La Figura 9 muestra la ventana para invitar miembros al proyecto.

Paso 3. Crear el backlog. Después de agregar a los miembros del equipo de trabajo al proyecto, el siguiente paso es crear el backlog. La Figura 10 muestra parte del Backlog con sus elementos de trabajo.

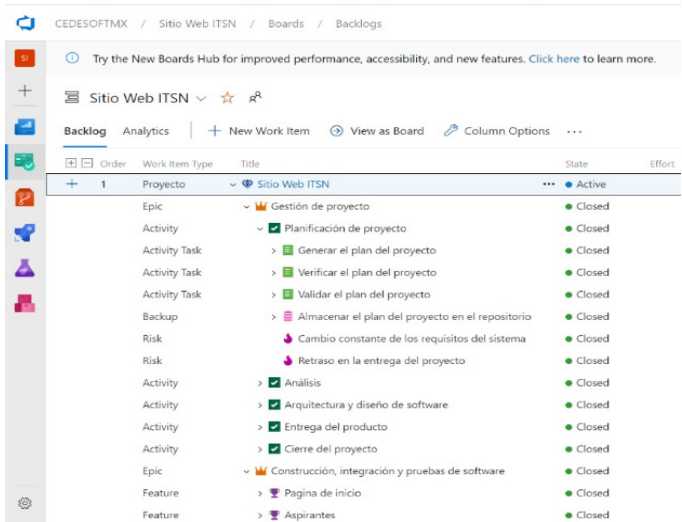


Figura 10 – Backlog del proyecto

New iteration

Iteration name

Iteration name

iteration name

Start date

End date

Location

Sitio Web ITSN\Planeación

Save and close

Cancel

Figura 11 – Creación de las iteraciones

Paso 4. Definir las iteraciones del proyecto. Una vez creada la lista de trabajo se crearon y se gestionaron las iteraciones del proyecto. La Figura 11 muestra la creación de las iteraciones del proyecto.

Para el caso de estudio se definieron seis iteraciones. La Figura 12 muestra las iteraciones del proyecto Sitio Web ITSN.

New	New child	+	-
Iterations			
Start Date			
End Date			
▼ Sitio Web ITSN			
Planeación	2/21/2022	2/25/2022	
Análisis	2/28/2022	3/4/2022	
Diseño	3/7/2022	3/18/2022	
Construcción, pruebas e integración	3/21/2022	6/10/2022	
Cierre de Proyecto	5/23/2022	5/27/2022	
Entrega	6/27/2022	6/30/2022	

Figura 12 – Iteraciones del proyecto

Paso 5. Agregar los elementos de trabajo a las iteraciones. Después de definir las iteraciones (sprints) del proyecto se agregaron los elementos de trabajo a cada una de ellas. La Figura 13 muestra algunos elementos de trabajo para el sprint de entrega con sus diferentes estados (To do, Doing, Done y Closed).

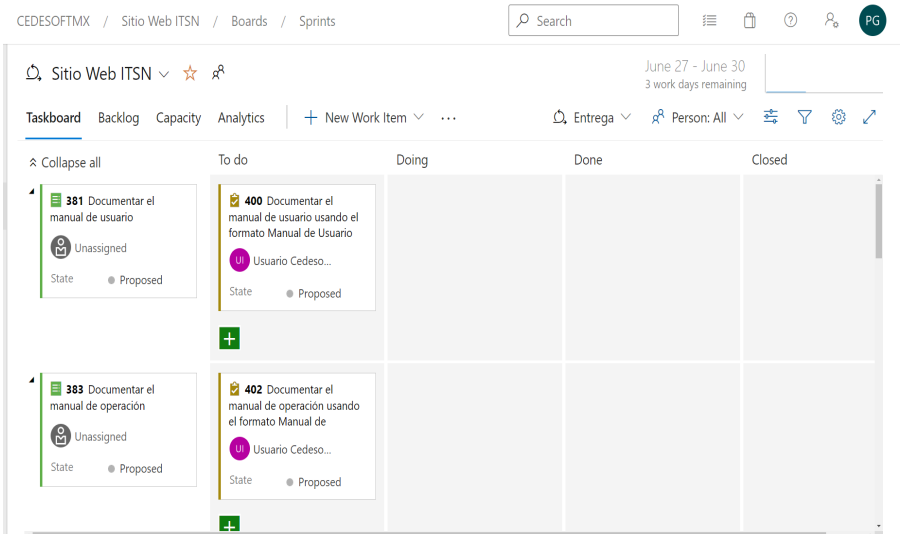


Figura 13 – Ejemplo de una iteración con sus elementos de trabajo

Paso 6. Ejecutar las iteraciones. El tablero Kanban de Azure Boards es una herramienta ágil de gestión de proyectos diseñada para ayudar a visualizar el trabajo y maximizar la eficiencia. A través de este servicio se le dio seguimiento a todos los elementos de trabajo hasta cerrarlos (closed), en la Figura 14 se muestran algunos de los elementos de trabajo cerrados en el sprint de Entrega.

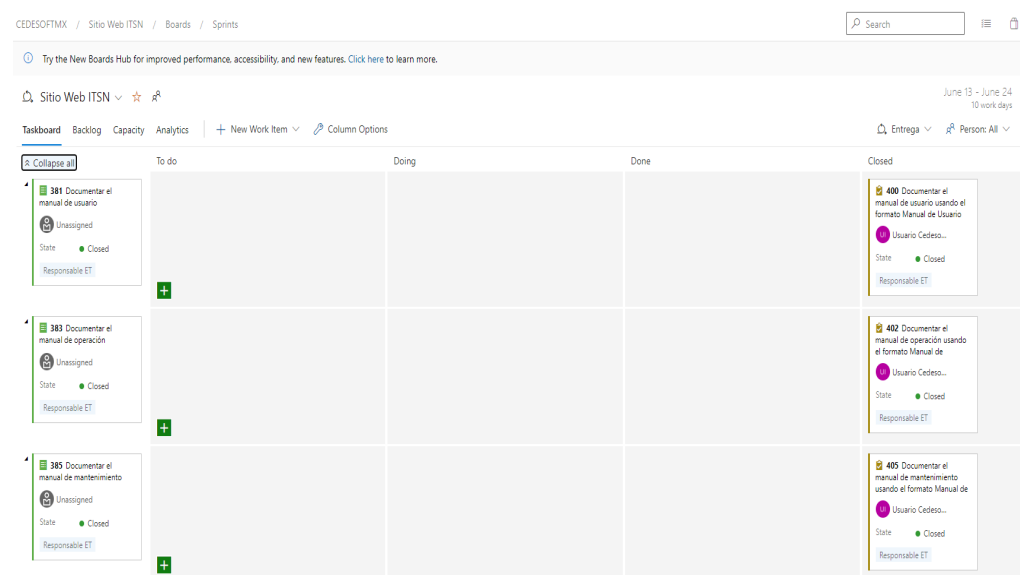


Figura 14 – Elementos de trabajo cerrados (closed)

C. Resultados de la ejecución del proyecto

En la Tabla 3 podemos observar los productos de GP y herramientas que se utilizaban antes de la adopción de DevOps y en la columna 4 se muestran los servicios que se utilizaron para la automatización de los productos.

Producto de trabajo	Herramienta	Servicio Azure DevOps
Plan de proyecto	Archivo de Texto	Azure Boards
Cronograma de actividades	Microsoft Project	Azure Boards
Plan de Adquisiciones, capacitaciones y costos	Hoja de cálculo	Azure Boards
Riesgos	Hoja de cálculo	Azure Boards
Lista de verificación del Plan de Proyecto	Archivo de Texto	Azure Boards
Lista de validación del Plan de Proyecto	Archivo de Texto	Azure Boards
Repositorio del proyecto	Drive	Azure Repos
Reporte de avance	Archivo de Texto	Azure Boards

Producto de trabajo	Herramienta	Servicio Azure DevOps
Acta de reunión	Archivo de Texto	Azure Boards
Acciones correctivas	Archivo de Texto	Azure Boards
Acta de aceptación	Archivo de Texto	-----

Tabla 3 – Productos de trabajo del proceso GP que se automatizaron

La Tabla 4 muestra los productos del proceso de IS que se lograron automatizar con los servicios de Azure DevOps.

Producto de trabajo	Herramienta	Servicio Azure DevOps
Especificación de requisitos	Archivo de Texto	Azure Boards
Lista de verificación de requisitos	Archivo de Texto	Azure Boards
Lista de validación de requisitos	Archivo de Texto	Azure Boards
Matriz de trazabilidad	Hoja de cálculo	Azure Boards
Diseño de software	Archivo de Texto	-----
Lista De Verificación para Revisión de Diseño de Software	Archivo de Texto	Azure Boards
Casos de Prueba, Procedimientos de Prueba y Reporte de prueba	Hoja de cálculo	Azure Test Plans
Lista de validación del software	Archivo de Texto	Azure Boards
Manual de Operación	Archivo de Texto	-----
Manual de Usuario	Archivo de Texto	-----
Manual de Mantenimiento	Archivo de Texto	-----
Lista de Verificación para la Revisión de los manuales de Operación, Mantenimiento y de Usuario	Archivo de Texto	Azure Boards

Tabla 4 – Productos de trabajo del proceso IS que se automatizaron

La Tabla 5 muestra que con la adopción de DevOps se logró automatizar el 90% de los productos de trabajo que se generaban de forma manual en el proceso de GP, solo un producto no se automatiza, el Acta de aceptación. En el proceso de IS se logró automatizar el 66% de ellos, los productos de trabajo Diseño del Software, Manuales de Operación, Usuario y Mantenimiento son de forma manual.

Proceso	Total de productos de trabajo del proceso	Total de productos de trabajo automatizados en DevOps	%
Gestión del proyecto	11	10	90
Implementación del software	12	8	66
Totales	23	18	78

Tabla 5 – Análisis de productos de trabajo automatizados en DevOps

Después de revisar el total de productos de trabajo de los procesos de GP e IS se logró automatizar el 78% de ellos a través de los servicios de DevOps, lo que representa un logro importante para el Centro de desarrollo de Software del ITSN que se encuentra certificado en el estándar ISO/IEC 29110.

4. Conclusiones

En este caso de estudio se muestra cómo a través de algunos de los servicios que ofrece DevOps, como son: Azure Boards, Azure Repos y Azures Test Plans, se automatizaron los productos de trabajo de los procesos GP e IS del estándar ISO/IEC 29110. Un total de 18 productos fueron automatizados lo que representa el 78%.

DevOps permitió crear y personalizar un proceso heredado a través de la selección de una de sus plantillas de proceso: Basic, Agile, Scrum y CMMI. Para este caso de estudio se seleccionó la plantilla de CMMI para la configuración de los procesos del estándar permitiendo su automatización. La plantilla de procesos de CMMI contiene elementos de trabajo que permitieron planificar, realizar un seguimiento del trabajo, gestión de solicitudes de cambio, gestión de riesgos y seguimiento a los requisitos.

Encontramos que no existen estudios que aborden la automatización de los productos de trabajo del estándar ISO/IEC 29110 con la adopción de DevOps y que ayuden a las pequeñas organizaciones de desarrollo de software a mejorar la calidad de sus productos y servicios.

Después de la adopción de DevOps se ha logrado reducir el tiempo de desarrollo del software incrementando la colaboración y la productividad de los equipos, ya que anteriormente se realizaba el llenado de los productos de trabajo utilizando diferentes herramientas como son Archivos de texto, hojas de cálculo, Microsoft Project y Drive.

Como trabajo futuro se pretende difundir la adopción de DevOps en las pequeñas organizaciones de desarrollo de software certificadas en el estándar ISO/IEC 29110 como una opción para la automatización de sus procesos.

Referencias

- Bucena, I., & Kirikova, M. (2017). Simplifying the DevOps Adoption Process. *BIR Workshops*.
- Guerrero, J., Zúñiga, K., Certuche Díaz, S., & Pardo Calvache, C. (2019). What is there about DevOps? Preliminary Findings from a Systematic Mapping. *Research Gate*.
- Guerrero, J., Zúñiga, K., Certuche Díaz, S., & Pardo Calvache, C. (2020). Un estudio de mapeo sistemático sobre DevOps. *Jornal de la ciencia de ingeniería*, 48-62.
- Ghantous, G., & Gill, A. (2017). DevOps: Concepts, Practices, Tools, Benefits and Challenges. *PACIS 2017 Proceedings*.
- Lwakatare, L., Kilamo, T., Karvonen, T., & Sauvola, T. (2019). DevOps en la práctica: un estudio de caso múltiple de cinco empresas. *Elsevier*, 114, 217-230.

- Mejía, J., Bonilla, E., & Villanueva Rosas, E. (2020). Identificación de herramientas como soporte en la implementación de la norma ISO/IEC 29110 (Perfil básico). *RISTI – Revista Ibérica de Sistemas e Tecnologias de Informação*, 15-33.
- Mohsienuddin Mohammad, S. (2018). Streamlining DevOps automation for Cloud. *IJCRT.ORG*, 251-256.
- Muñoz, M., Negrete, M., & Arcilla Cobián, M. (2021). Using a platform based on the Basic profile of ISO/IEC 29110 to reinforce DevOps environments. *Journal of Universal Computer Science*, 91-110.
- SánchezCastillo, J., Martínez, A., Quesada López, C., & Jenkins, M. (2020). Caracterización de las prácticas de DevOps en organizaciones que desarrollan software: Un mapeo. *RISTI – Revista Ibérica de Sistemas e Tecnologias de Informação*, 83-96.
- Urresti, S., Ronald, D., Rodríguez, L., & de la Caridad, G. (2019). Las PyME de desarrollo de software. Modelos de mejora de sus procesos en Latinoamérica. *Espacios*, 40(28), 9.

Propuesta de mejoras para la implementación del estándar ISO/IEC 29110 utilizando lógica matemática

Víctor Terron, Jezreel Mejía

{victor.terron; jmejia}@cimat.mx

Centro de Investigación en Matemáticas, Parque Quantum, Ciudad del Conocimiento, Av. Lasec, Andador Galileo Galilei, Manzana 3, Lote 7, CP 98160 Zacatecas, Zac., México.

* Autor de correspondencia: Jezreel Mejía

DOI: 10.17013/risti.49.52–66

Resumen: El estándar ISO/IEC 29110 está específicamente orientado a muy pequeñas organizaciones ya que se adapta al contexto y necesidades de este tipo de organizaciones. Tras haber analizado las estructuras de los procesos: actividades, roles, tareas, y productos de trabajo considerados dentro del estándar ISO/IEC 29110 en su perfil básico se identificó falta de información acerca de los estados de los productos de trabajo que son necesarios para desarrollar las tareas, lo que dificulta su correcta implementación. Para ayudar a resolver esta problemática este artículo presenta propuestas de tautologías de lógica matemática. Como resultado se presentan propuestas de mejora en el establecimiento de estados a productos de trabajo que ayudarán en el control de la versión de los productos de trabajo y su documentación asociada.

Palabras-clave: ISO/IEC 29110, tautologías, lógica matemática, mejora, evaluación, elementos explícitos e implícitos.

Improvement proposal for the implementation of ISO/IEC 29110 standard through mathematical logic

Abstract: The ISO/IEC 29110 standard is aimed explicitly at very small organizations as it adapts to the context and needs of this type of organization. After analyzing the structures of processes: activities, roles, tasks, and work products considered within the profile basic of the ISO/IEC 29110 standard, a lack of information was identified regarding the states of the work products that are necessary to carry out the tasks, which makes their proper implementation difficult. Therefore, this article presents proposals for tautologies of mathematical logic. As a result, proposals are presented to improve the establishment of states for work products that will aid in controlling the version of the work products and their associated documentation.

Keywords: ISO/IEC 29110, tautologies, mathematical logic, improvement, assessment, explicit and implicit elements.

1. Introducción

Actualmente, las entidades muy pequeñas relacionadas a la industria del desarrollo de software son muy importantes, ya que como menciona J. Mejía, et. al. (Jezreel Mejía, et. al., 2021), este tipo de organizaciones representa un 94% ($\pm 3\%$) en esta industria. Por lo tanto, éstas requieren de certificaciones en estándares que den garantía de la calidad de los productos y servicios de software generados (Mirna Muñoz, et. al., 2021). Además, como lo menciona el grupo WG24 de la ISO/IEC se deben considerar sus características y necesidades como por ejemplo condiciones financieras, habilidades, cultura organizacional y modelos de referencia (Xavier Larrucea, et. al., 2016), entre otros.

En este contexto, se ha creado un estándar orientado para pequeñas organizaciones que desarrollan software no crítico, como lo es ISO/IEC 29110 (Apinporn Methawachananont, et. al., 2020). Este estándar tiene por objetivo ayudar a entidades muy pequeñas de hasta 25 individuos que a su vez pueden ser equipos de una organización muy grande, para el desarrollo de software no crítico por medio de la incursión en dos procesos: Gestión de Proyecto (GP) e Implementación de Software (IS). Estos procesos contienen actividades, tareas y productos de trabajo (ISO/IEC WG24, 2011). Sin embargo, como lo indica M. Muñoz, et. al. (Mirna Muñoz, Patricia Montoya, 2021) se han identificado dificultades en su implementación debido a la falta de comprensión, falta de control de versiones, funcionalidades y de documentación, desarrollo de pruebas, entre otros.

Por lo tanto, este artículo propone mejoras para la implementación del estándar ISO/IEC 29110 enfocadas en la falta de control de versiones que impactan en el desarrollo de la documentación. Para lograrlo se analiza la estructura, trazabilidad de los productos de trabajo, terminología y estados de los productos de trabajo para proponer tautologías utilizando lógica matemática. Estas tautologías permitieron evaluar la persistencia de los elementos explícitos e implícitos en el estándar identificando la falta del estado de algunos productos de trabajo que son necesarios para desarrollar las tareas y actividades dentro de los procesos de GP e IS.

Después de la introducción este artículo está estructurado de la siguiente manera: en la segunda sección se describe brevemente el estándar ISO/IEC 29110; en la tercera sección se analiza el estándar; en la cuarta sección, se proponen las tautologías a los elementos del proceso y su formalización matemática; en la quinta sección se presenta la aplicación de las tautologías al estándar; en la sexta sección se presentan las conclusiones y trabajo futuro.

2. ISO/IEC 29110

El estándar ISO/IEC 29110 es una serie de normas en informes técnicos desarrollado para pequeñas organizaciones (Edgar Bonilla Rivas, Mirna Muñoz, 2020). Específicamente define los perfiles del ciclo de vida del software. Una pequeña organización (PO) (VSE por sus siglas en inglés — Very Small Entities) se define como una entidad (empresa, organización, departamento o proyecto) que tiene hasta 25 personas. La mayoría de

las empresas de software pertenecen a la categoría de PO (Jezreel Mejía, et. al., 2021) (Apinporn Methawachananont, et. al., 2020).

El estándar ISO/IEC 29110 fue desarrollado por la Organización Internacional para la Normalización y la Comisión Electrotécnica Internacional.

Este estándar aborda dos procesos, en las cuales se establecen 7 objetivos para cada una de ellas, los cuales están descritas en (ISO/IEC WG24, 2011). A continuación, se describen cada uno de los elementos de los procesos, actividades, productos de trabajo y roles.

2.1. Proceso de Gestión de Proyecto

Tiene como propósito establecer y llevar a cabo sistemáticamente tareas del proyecto de implementación, mismas que permiten lograr los objetivos del proyecto en 3 aspectos: calidad, tiempo y coste (ISO/IEC WG24, 2011).

El proceso de GP consta de 4 actividades, cada una de las actividades con tareas específicas, tal como se muestra en la Tabla I:

Identificador de la actividad	Nombre de la actividad	Número de tareas
GP 1.	Planificación del Proyecto.	15
GP 2.	Ejecución del Plan del Proyecto.	6
GP 3.	Evaluación y control del proyecto.	3
GP 4.	Cierre del proyecto.	2

Tabla I – Actividades del proceso GP.

Cada actividad genera productos de trabajo, el estándar ISO/IEC 29110 menciona 3 tipos de productos de trabajo:

- Productos de trabajo de entrada.
- Productos de trabajo internos.
- Productos de trabajo de salida.

Los productos de trabajo de entrada en el proceso de GP son indicados en la Tabla II:

Nombre	Origen
Enunciado de trabajo.	Cliente —CL.
Configuración de software.	Área de proceso de implementación de software —IS.
Solicitud de cambio.	Cliente —CL. Área de proceso de Implementación de software —IS.

Tabla II – Productos de trabajo de entrada en GP.

Los productos de trabajo internos en el proceso de GP son indicados en la Tabla III:

Nombre
Solicitud de cambio.
Acciones correctivas.
Acta de reunión.
Resultados de verificación.
Reporte de avance.
Respaldo del repositorio del proyecto.

Tabla III – Productos de trabajo internos en GP.

Los productos de trabajo de salida en el proceso de GP son indicados en la Tabla IV:

Nombre	Origen
Plan de Proyecto.	Área de proceso de implementación de Software.
Acta de Aceptación.	Alta Dirección.
Repositorio del Proyecto.	Área de proceso de implementación de Software.
Acta de reunión.	Cliente – CL.
Configuración de Software.	Cliente – CL.

Tabla IV – Productos de trabajo de salida de GP.

Así mismo dentro del proceso de GP, intervienen específicamente ciertos roles, mismos que se colocan en la Tabla V:

Nombre	Abreviatura
Cliente.	CL.
Gestor de Proyecto.	GP.
Líder Técnico.	LT.
Equipo de Trabajo.	ET.

Tabla V – Roles involucrados en el proceso de GP.

2.2. Proceso de Implementación de Software

El propósito de este proceso es la realización sistemática de las actividades de análisis, diseño, construcción, integración y pruebas para los productos conforme a los requisitos (ISO/IEC WG24, 2011)(Jezreel Mejía, Mirna Muñoz, 2019).

El proceso IS consta de 4 actividades, cada una de las actividades con tareas específicas, como se indica en la Tabla VI:

Identificador de la actividad	Nombre de la actividad	Número de tareas
IS.1.	Inicio de la Implementación del software.	2
IS.2.	Análisis de requisitos de software.	7
IS.3.	Arquitectura y diseño detallado de software.	8
IS.4.	Construcción del software.	7
IS.5.	Integración y pruebas de software.	11
IS.6.	Entrega del producto.	6

Tabla VI – Actividades del proceso de IS.

Los productos de trabajo de entrada del proceso de Implementación de software son indicados en la Tabla VII:

Nombre	Origen
Plan de Proyecto.	Área de proceso de gestión de proyecto.
Repositorio del proyecto.	Área de proceso de gestión de proyecto.

Tabla VII – Productos de trabajo de entrada de IS.

Los productos de trabajo internos del proceso de Implementación de Software son indicados en la Tabla VIII:

Nombre
Resultados de verificación.
Resultados de validación.

Tabla VIII – Productos de trabajo internos de IS.

Los productos de trabajo de salida del proceso de Implementación de Software son indicados en la Tabla IX:

Nombre	Origen
Configuración del software: - Especificación de requisitos. - Diseño de software. - Registro de trazabilidad. - Componentes de software. - Software. - Casos de prueba y procedimientos de prueba. - Reporte de pruebas. - Manual de operación. - Manual de usuario. - Manual de mantenimiento.	Gestión de Proyecto.
Solicitud de cambio.	Gestión de Proyecto.

Tabla IX – Productos de trabajo de salida de IS.

Así mismo dentro del proceso de IS, intervienen específicamente ciertos roles, mismos que se colocan en la Tabla X:

Nombre	Abreviatura
Cliente.	CL.
Gestor de Proyecto.	GP.
Analista.	AN.
Líder Técnico.	LT.
Diseñador.	DIS.
Equipo de Trabajo.	ET.
Programador.	PR.

Tabla X – Roles involucrados en el proceso de IS.

3. Análisis del estándar ISO/IEC 29110

En esta sección se presenta el análisis del estándar ISO/IEC 29110 considerando sus elementos explícitos e implícitos de los procesos de GP e IS.

3.1. Elementos explícitos e implícitos

Para desarrollar un análisis más completo de los elementos explícitos e implícitos, además del análisis del estándar ISO/IEC 29110, se analizó la guía de implementación de este estándar (ISO/IEC WG24, 2011).

Se entiende por **elemento explícito** a todo elemento que contiene una definición dentro de la guía de implementación del estándar. Por lo tanto, como se presentó en la sección anterior se puede observar que la estructura de elementos explícitos en los dos procesos del estándar se da de acuerdo con la Figura 1.



Figura 1 – Elementos explícitos del estándar ISO/IEC 29110.

Tras identificar los elementos explícitos se identificaron **elementos implícitos**, que son aquellos que no tienen definición dentro de la guía de implementación del estándar

y que solo son identificados por su nombre. Un elemento implícito identificado tras realizar el análisis fue el estado del producto de trabajo.

3.5. Elementos principales del estándar ISO/IEC 29110.

Como se puede apreciar en la Figura 2 y tras el análisis de la estructura y las tareas se identificó que dentro de la guía de implementación del estándar hay un elemento que influye en la versión de los productos de trabajo: el estado. Dicho elemento no está descrito explícitamente dentro del estándar.

Por lo tanto, elemento de estado es un elemento implícito, es decir, un elemento que forma parte de la estructura, pero que no está descrito en ninguna sección de la guía de implementación del estándar ISO/IEC 29110 y que es elemental para la comprensión e implementación apropiada del estándar. Además, tras el análisis se identificó que el elemento estado facilita la identificación de la versión de los productos de trabajo y los roles involucrados en su desarrollo.



Figura 2 – Elementos explícitos e implícitos del estándar ISO/IEC 29110.

Por lo tanto, los elementos principales del estándar incluyen elementos explícitos e implícitos que son mostrados en la Figura 2.

4. Tautología a los elementos de proceso y formalización matemática

Como menciona Kenneth Rosen las reglas de la lógica son utilizadas para distinguir entre argumentos válidos e inválidos, la lógica tiene numerosas aplicaciones en las ciencias computacionales (Rosen, 2019).

Los elementos básicos de la lógica son las proposiciones, que son oraciones declarativas que pueden ser verdaderas o falsas, pero no ambas.

Kenneth Rosen define la tautología como una proposición compuesta que es siempre verdadera, una proposición compuesta es una proposición construida utilizando proposiciones con operadores lógicos (Rosen, 2019).

Para denotar variables proposicionales o variables oracionales se utilizan letras, en la lógica proposicional hay diferentes tipos de proposiciones, tal como se indica en la Tabla XI:

Tipo	Proposiciones	Denotación
Negación.	p	$\neg p$ o $\bar{p}$
Conjunción.	p y q	$p \wedge q$
Disyunción.	p y q	$p \vee q$
XOR.	p y q	$p \oplus q$
Condicional.	p y q	$p \rightarrow q$
Bicondicionales.	p y q	$p \leftrightarrow q$
Todos (Cuantificadores).	$p(x)$	$\forall x (P(x))$
Algunos (Cuantificadores).	$p(x)$	$\exists x (P(x))$

Tabla XI – Tipos de proposiciones y cuantificadores.

Considerando las diferentes proposiciones y cuantificadores se seleccionó proposiciones de tipo condicional. Kenneth Rosen menciona que este tipo de proposiciones se expresa de diferentes formas como: si p , entonces q , si p , q , p es suficiente para q , entre otras.

4.1. Establecimiento de tautologías.

Tras el análisis de los elementos principales identificados que incluyen tanto los elementos explícitos e implícitos del estándar los cuales son: procesos, actividades, roles, tareas, productos de trabajo y estado del producto de trabajo.

A los elementos anteriores se les asignaron letras para utilizar la lógica proposicional estableciendo proposiciones para generar tautologías:

1. $P(x)$: x es un proceso.
2. $Q(y)$: y es una actividad del proceso.
3. $R(z)$: z es una tarea de la actividad del proceso.
4. $S(a)$: a es un producto de trabajo de la tarea de la actividad del proceso.
5. $T(b)$: b es un rol que participa en el producto de trabajo de la tarea de la actividad del proceso.
6. $U(c)$: c es un estado de producto de trabajo de la tarea de la actividad del proceso.
7. $V(x)$: x tiene actividades.
8. $W(x)$: x tienen estados.
9. $O(x)$: x tiene tareas.
10. $M(x)$: x genera productos de trabajo.
11. $N(x)$: x tiene roles.
12. $L(x)$: participan en.
13. $D(x)$: necesita un.

Tras el establecimiento de las proposiciones se desarrollan dos clases de tautologías:

1. Tautologías de elementos explícitos: aquellas que cumplen con los elementos explícitos véase sección 3.1.

#	Tautología	Denotación matemática
1	Todo proceso tiene actividades.	$\forall x(P(x) \rightarrow V(x))$
2	Toda actividad tiene tareas.	$\forall y(Q(y) \rightarrow O(y))$
3	Toda tarea tiene roles.	$\forall z(R(z) \rightarrow N(c))$
4	Algunas tareas generan productos de trabajo.	$\exists z(R(z) \rightarrow M(a))$
5	Algunos roles participan en la tarea.	$\exists b(T(b) \rightarrow L(z))$
6	Algunos productos de trabajo no tienen estado.	$\exists a(S(z) \rightarrow W(c))$
7	Alguna tarea que pertenezca a alguna actividad genera algún producto de trabajo.	$\exists R(z) \in \exists Q(y) \rightarrow \exists M(a)$

Tabla XII – Tautologías del estándar ISO/IEC 29110.

2. Tautologías de elementos implícitos: aquellas que cumplen con los elementos explícitos e incluyen los elementos explícitos véase sección 3.2.

#	Tautología	Denotación matemática
8	Para todo producto de trabajo es necesario un estado de producto de trabajo.	$\exists z \rightarrow D(c))$
9	Algún rol genera algún producto de trabajo si y solo si participa en alguna tarea.	$\exists b(T(b) \rightarrow M(a)) \leftrightarrow \exists x(L(b))$
10	Toda actividad y toda tarea genera todo producto de trabajo con algún un estado de producto de trabajo.	$(\forall y \wedge \forall z \rightarrow \forall a) \rightarrow \exists c$

Tabla XIII – Tautologías propuestas para el estándar ISO/IEC 29110.

5. Propuestas de mejora

Considerando las tautologías presentes en el estándar y su aplicación, se identificaron las siguientes problemáticas:

Pérdida de estados, por ejemplo, en la actividad GP.1.5. ya que recibe como producto de trabajo de entrada un enunciado de trabajo sin estado.

En total dentro del estándar se detectaron 41 incidencias relacionadas con alguna de las problemáticas mencionadas en este artículo.

6. Aplicación de tautologías al estándar ISO/IEC 29110

Una vez propuestas las tautologías al estándar ISO/IEC 29110 y que tienen por fundamento a la lógica proposicional, se detectaron aspectos de mejora en la implementación del estándar, con respecto al estado de los productos de trabajo. Tras la aplicación de todas las tautologías se identificó que las tautologías que no se cumplen dentro del estándar ISO/IEC 29110 hacen referencia a las tautologías 8 y 10.

Es decir, existen productos de trabajo a los cuales no se identificó la existencia de un estado que inciden en la creación de un producto de trabajo y que al mismo tiempo permiten establecer su versión.

La Figura 3 muestra el total de tareas relacionadas a un producto de trabajo que no indica su estado tanto en el proceso de gestión de proyecto y del proceso de implementación de software.

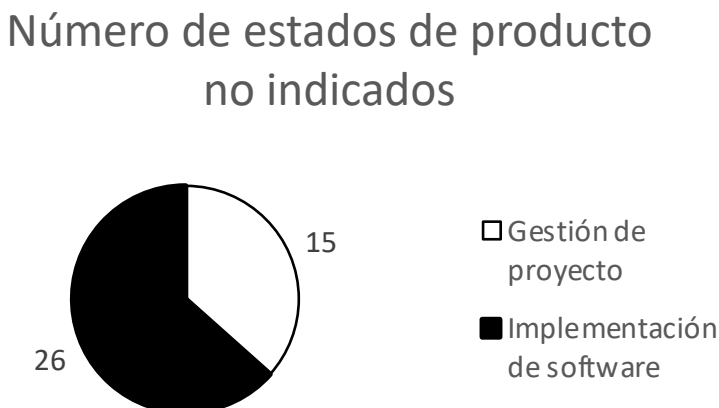


Figura 3 – Número de estados de producto no indicados.

Como se observa en la Figura 3, para el proceso de gestión de proyecto 15 de 26 tareas que hacen referencia a productos de trabajo no indican su estado.

Mientras que para el proceso de implementación de software son 27 de 41 tareas que no indican su estado.

Tras obtener estos resultados la Tabla XIV muestra las propuestas de mejora con respecto a los estados de los productos de trabajo que están relacionadas a las tareas.

Actividad	Tareas	Elemento mejorado	Propuesta de mejora
GP.1.	GP.1.5.	Producto de trabajo de entrada.	Cambiar estado del enunciado de trabajo a [Revisado].
	GP.1.12.	Producto de trabajo de entrada.	Recibe un enunciado de trabajo sin estado, se propone cambiar el estado a [Revisado].
		Producto de trabajo de salida.	Genera un plan de proyecto como producto de trabajo de salida sin estado, se propone cambiar su estado a [Revisado].
	GP.1.13.	Producto de trabajo de entrada.	Cambiar el estado del plan de proyecto a [Revisado].
GP.1.	GP.1.14.	Producto de trabajo de salida.	Cambiar el estado del acta de reunión a [Aceptado].
	GP.1.15.	Producto de trabajo de entrada.	Cambiar el estado de la estrategia de control de versiones a [Aceptado].
		Producto de trabajo de salida.	Cambiar el estado del repositorio del proyecto a [Verificado].
GP.2.	GP.2.1.	Producto de trabajo de entrada.	Cambiar el estado del plan de proyecto a [Aceptado].
		Producto de trabajo de salida.	Cambiar el estado del reporte de avance a [Iniciado].
	GP.2.2.	Producto de trabajo de entrada.	Cambiar el estado del plan del proyecto a [Aceptado].
	GP.2.3.	Producto de trabajo de entrada.	Cambiar el estado del plan de proyecto a [Actualizado].
	GP.2.4.	Producto de trabajo de entrada.	Cambiar el estado del plan de proyecto a [Actualizado].
	GP.2.5.	Producto de trabajo de entrada.	Corregir el producto de entrada a Plan de proyecto apartado de estrategia de control de versiones.
		Producto de trabajo de salida.	Cambiar el estado del respaldo del repositorio del proyecto a [Actualizado].
GP.3.	GP.3.1.	Producto de trabajo de entrada.	Cambiar el estado del plan de proyecto a [Aceptado].
IS.1.	IS.1.1.	Producto de trabajo de entrada.	Cambiar el estado del plan del proyecto a [Aceptado].
IS.2.	IS.2.2.	Producto de trabajo de entrada.	Cambiar el estado del plan de proyecto a [Revisado].
IS.2.	IS.2.3.	Producto de trabajo de entrada.	Cambiar el estado del plan del proyecto a [Revisado].
	IS.2.4.	Producto de trabajo de salida.	Cambiar el estado de la especificación de requisitos a [Validado].
	IS.2.6.	Producto de trabajo de entrada.	Cambiar el estado de la especificación de requisitos a [Validado].

Actividad	Tareas	Elemento mejorado	Propuesta de mejora
IS.3.	IS.3.1.	Producto de trabajo de entrada.	Cambiar el estado del plan del proyecto a [Revisado].
	IS.3.3.	Producto de trabajo de salida.	Cambiar el estado del diseño de software a [Preliminar].
		Producto de trabajo de salida.	Cambiar el estado del registro de trazabilidad a [Preliminar].
	IS.3.4.	Producto de trabajo de entrada.	Cambiar el estado del diseño de software a [Preliminar].
		Producto de trabajo de entrada.	Cambiar el estado del registro de trazabilidad a [Preliminar].
	IS.3.5.	Producto de trabajo de salida.	Cambiar el estado de los casos de prueba y procedimientos de prueba a [Preliminar].
IS.4.	IS.3.6.	Producto de trabajo de entrada.	Cambiar el estado de los casos de prueba y procedimientos de prueba a [Preliminar].
	IS.4.1.	Producto de trabajo de entrada.	Cambiar el estado del plan del proyecto a [Revisado].
	IS.4.4.	Producto de trabajo de entrada.	Cambiar el estado de los componentes de software a [Preliminar].
IS.5.	IS.5.1.	Producto de trabajo de entrada.	Cambiar el estado del plan de proyecto a [Aceptado].
	IS.5.3.	Producto de trabajo de salida.	Cambiar el estado de los procedimientos de prueba a [Validado].
IS.5.	IS.5.4.	Producto de trabajo de entrada.	Cambiar el estado de los casos de prueba y procedimientos de prueba a [Validado].
	IS.5.7.	Producto de trabajo de salida.	Cambiar el estado del manual de operación a [Probado].
	IS.5.9.	Producto de trabajo de salida.	Cambiar el estado del manual de usuario a [Probado].
	IS.5.11.	Producto de trabajo de salida.	Cambiar el estado de la configuración de software a [Preliminar].
IS.6.	IS.6.2.	Producto de trabajo de entrada.	Cambiar el estado de la configuración de software a [Verificado].
	IS.6.3.	Producto de trabajo de entrada.	Cambiar el estado de la configuración de software a [Verificado].
	IS.6.4.	Producto de trabajo de entrada.	Cambiar el estado de la configuración de software a [Preliminar].
	IS.6.5.	Producto de trabajo de entrada.	Cambiar el estado de la configuración de software a [Preliminar].
		Producto de trabajo de salida.	Cambiar el estado de la configuración de software a [Verificado].
	IS.6.6.	Producto de trabajo de entrada.	Cambiar el estado de la configuración de software a [Verificado].

Tabla XIV – Propuestas de mejora al estándar ISO/IEC 29110.

De acuerdo con las propuestas de estados indicados en la Tabla XIV, se pueden agrupar como se muestra en la Figura 4 y 5.

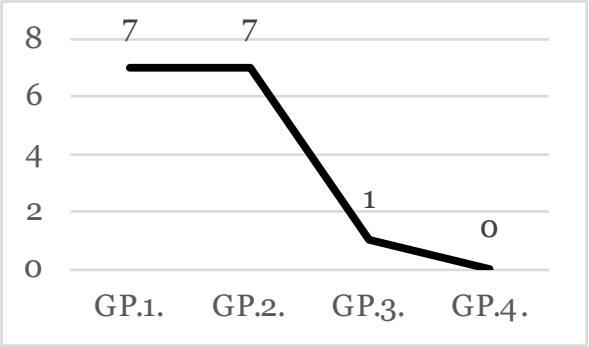


Figura 4 – Número de mejoras en GP por actividad.

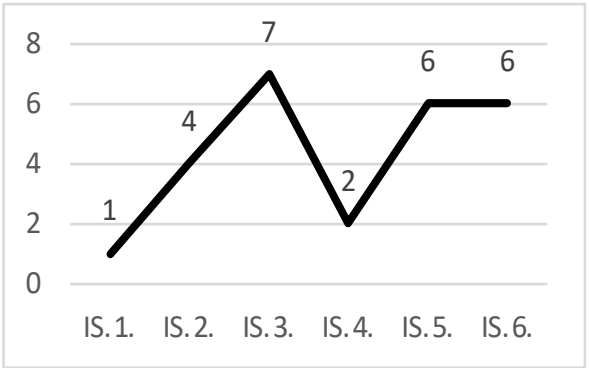


Figura 5 – Número de mejoras en IS por actividad.

Como se puede apreciar en la Figura 4, GP.1. y GP.2. presentan mayor incidencia en no indicar el estado de producto de trabajo y por tanto un mayor número de mejoras propuestas dentro de sus actividades. En la Figura 5, la actividad con un mayor número de mejoras propuestas es IS.3.

7. Conclusiones y trabajo futuro

Las empresas muy pequeñas de desarrollo de software manifiestan un interés por mejorar la calidad de sus procesos y productos, mediante la implementación del estándar ISO/IEC 29110 ya que se adecua a su contexto. Sin embargo, al implementarlo se identifican distintas problemáticas como lo es la falta de control de versiones en sus productos de trabajo, entre otros. En este contexto las propuestas de mejora realizadas

en este artículo para la implementación del estándar ISO/IEC 29110 a través del establecimiento de tautologías sustentan los hallazgos de mejora por medio de la lógica matemática.

Como resultado se identificaron 26 tareas del proceso de gestión de proyectos que hacen referencia a productos de trabajo que no indican su estado, mientras que para el proceso de implementación del software son 15 tareas. Tras este resultado la Tabla XIV propone formalizar un estado del producto de trabajo teniendo mayor incidencia en GP.1., GP.2., IS.3.

La lógica matemática a través de las proposiciones ha permitido establecer tautologías que permiten evaluar que los elementos explícitos e implícitos se cumplan en la estructura de los procesos de este estándar.

Como trabajo futuro dentro de esta investigación se propone desarrollar una herramienta web para facilitar la evaluación de los elementos de los procesos de este estándar.

Referencias

- Methawachananont, A., Buranarach, M., Amsuriya, P., Chaimongkhon, S., Krairaksa, K., Thepchai, S.. (2020). Software Process Capability Self-Assessment Support System Based on Task and Work Product Characteristics: A Case Study of ISO/IEC 29110 Standard. *IEICE Transactions on Information and Systems*, E103, 339-347. <https://doi.org/0.1587/transinf.2018EDP7303>
- Bonilla Rivas, E., Muñoz, M. (2020). Creación De Estrategias Para La Capacitación En El Estándar De Desarrollo De Software ISO/IEC 29110: Propuesta De Un Marco De Trabajo. 2020 9th International Conference On Software Process Improvement (CIMPS), (pp. 155-155). <https://doi.org/10.1109/CIMPS52057.2020.9390152>
- ISO/IEC WG24. (2011). ISO/IEC TR 29110-5-1-2: Software engineering — Lifecycle profiles for Very Small Entities (VSEs) Part 5-1-2: Management and engineerign guide: Generic profile group: Basic profile. (1 ed.). Switzerland, Geneva, Switzerland: IEEE.
- Mejía, J., Bonilla, E., Faustino, I., Jhordany, E., Villanueva, E. (2021). Propuesta de métricas para el desarrollo del software con base en el estándar ISO/IEC 29110. 2021 10th International Conference On Software Process Improvement (CIMPS), 2021 (págs. 58-65). Torreón, Coahuila, México: IEEE. <https://doi.org/10.1109/CIMPS54606.2021.9652770>
- Mejía, J., Muñoz, M. (2019). Propuesta de herramientas para la implementación de ISO/IEC 29110 (perfil básico). 2019 8th International Conference On Software Process Improvement (CIMPS) (pág. 11). Leon, México: IEEE. <https://doi.org/10.1109/CIMPS49236.2019.9082430>
- Muñoz, M., Mejía, J., Peña, A., Laporte, C., Gasca-Hurtado, G. P., Gómez-Álvarez, M.C.. (2021). An Exploratory Analysis of the perception of the utility of proven practices of the software basic profile of ISO/IEC 29110 by a set of VSEs in Mexico. Springer Nature Switzerland AG, 1-18. https://doi.org/10.1007/978-3-030-85521-5_29

- Muñoz, M., Montoya, P. (2021). Identification of issues in the implementation of the ISO/IEC29119: comparison between the state of the art and the state of practice. 2021 10th International Conference On Software Process Improvement (CIMPS) (págs. 90-97). Torreón, Coahuila, México: IEEE. doi:10.1109/CIMPS54606.2021.9652708
- Rosen, K. H. (2019). Discrete Mathematics and Its Applications (8th ed.). New York, United States of America: McGraw-Hill.
- Larrucea, X., O'Connor, R.V., Colomo-Palacios, R., Laporte, C.Y. (2016). Software Process Improvement in Very Small Organizations. IEEE Software, 33(2), 85-89. <https://doi.org/10.1109/MS.2016.42>

Aplicación móvil para monitoreo del consumo de materia seca en rumiantes usando la tecnología bluetooth

Jesús Sierra Martínez¹, Juan Carlos Elizondo Leal¹, Daniel López Aguirre¹,
Yadira Quiñonez², José Hugo Barrón Zambrano¹, Alan Díaz Manríquez¹,
Vicente Paul Saldívar Alonso¹, José Ramon Martínez Angulo¹

chuy_984@outlook.com; jcaelizondo@docentes.uat.edu.mx;
dlaguirre@docentes.uat.edu.mx; yadiraqui@uas.edu.mx;
hbarron@docentes.uat.edu.mx; amanriquez@docentes.uat.edu.mx;
vpsaldiv@docentes.uat.edu.mx; jrangulo@docentes.uat.edu.mx

¹ Facultad de Ingeniería y Ciencias, Universidad Autónoma de Tamaulipas, Victoria, 87120, México.

² Facultad de Informática Mazatlán, Universidad Autónoma de Sinaloa, Mazatlán, 82000, México.

DOI: 10.17013/risti.49.66–82

Resumen: La automatización de sistemas inteligentes controlados remotamente para recolectar, analizar y procesar datos es una de las aplicaciones más comunes en diferentes campos de aplicación. Usualmente, para analizar la dieta del animal se requiere mucho tiempo y capital humano; los experimentos consisten en colocar al animal en una jaula para recolectar datos de consumo. En este sentido, incorporar tecnología para analizar el consumo de materia seca en rumiantes es fundamental para incrementar el desempeño productivo del animal. Este documento se enfoca en desarrollar un sistema de monitoreo para la recolección de datos, tales como la temperatura y la humedad, así como el peso de los alimentos a lo largo del experimento nutricional, utilizando una aplicación móvil desarrollada en Android Studio que envía y recibe la información vía tecnología Bluetooth a un dispositivo electrónico basado en el microcontrolador Arduino Nano.

Palabras-clave: Automatización; Microcontroladores; Producción Animal; Ingesta de Materia Seca.

Mobile application for monitoring dry matter intake in ruminants using Bluetooth technology

Abstract: Automating remotely controlled intelligent systems to collect, analyze, and process data is one of the most common applications in different fields of application. Usually, analyzing the animal's diet requires much time and human capital; the experiments consist of placing the animal in a cage to collect consumption data. In this sense, incorporating technology to analyze the consumption of dry matter in ruminants is essential to increase the productive performance of the animal. This document focuses on developing a monitoring

system for data collection, such as temperature and humidity, as well as the weight of food throughout the nutritional experiment, using a mobile application developed in Android Studio that sends and receives the information via Bluetooth technology to an electronic device based on the Arduino Nano microcontroller.

Keywords: Automation; Microcontrollers; Animal Production; Dry Matter Intake.

1. Introducción

Hoy en día la tecnología es un referente importante en cualquier área de aplicación, la ciencia en convivencia con la tecnología ha permitido la creación de herramientas y dispositivos que simplifican las actividades diarias. En la actualidad, la automatización y el control de sistemas inteligentes se han convertido en un tema de gran interés para los investigadores de la comunidad científica en diferentes campos de aplicación; debido a que es posible crear aplicaciones que generan nuevas oportunidades tanto a la economía como a la sociedad. En la última década, mejorar las actividades diarias y crear nuevos productos y servicios ha llamado la atención de la sociedad en general, la industria, la academia y la investigación.

Cuando se trabaja con tecnología del Internet de las Cosas (IoT) se describe un escenario donde cualquier objeto físico puede convertirse en un terminal conectado a internet en un entorno doméstico para controlar y monitorear diferentes cosas en casa de forma remota desde cualquier lugar y en cualquier momento a través de una conexión a internet (Vashi, 2017). La conexión de los dispositivos se caracteriza principalmente por ser controlados remotamente desde cualquier lugar, en cualquier momento, por cualquier cosa y cualquier persona (Chin, 2019), para poder recopilar, analizar y procesar datos en plataformas informáticas (McCann, 2018). Algunas de las aplicaciones más comunes están relacionadas con el hogar inteligente (Erzİ, 2020, Rokonuzzaman, 2022), las ciudades inteligentes (Kirimtat, 2020, Shabber, 2021), la atención médica (Elkahlout, 2020, Hussain, 2020), transporte inteligente (Quiñonez, 2019, Cheng, 2021), la agricultura (Kassim, 2020, Abdul Hafeez, 2022), entre otros.

En este sentido, otra aplicación de interés se centra en la investigación del estado de salud de los rumiantes y de la Ingesta de Materia Seca (IMS), debido a que es el principal sustento de los ganaderos, entonces, es indispensable que los rumiantes tengan un buen estado de salud y que la IMS sea el adecuado para asegurar la producción de alimentos. La continua y creciente población que asume nuestra sociedad nos lleva a la necesidad de requerir más alimentos. Los productos animales han sido componentes de la alimentación humana. Se prevé que la demanda de estos productos (por ejemplo, leche y carne de vacuno) en las regiones tropicales aumente debido al crecimiento demográfico y al cambio climático. Es por ello, que se debe de aumentar el rendimiento de la producción animal, el cuál, debe lograrse mediante la mejora de la gestión y las tecnologías de producción (Ribeiro, 2020).

Los rumiantes son esenciales para la sostenibilidad en la producción agrícola debido a su capacidad de transformar los carbohidratos y proteínas no digeribles en los vegetales en alimentos disponibles para el consumo humano, como la leche y la carne. Además, estos animales, que ocupan un papel importante en el ciclo de nutrientes, devuelven hasta el 90% de los nutrientes digeridos al suelo a través de sus excrementos, orina y

restos (Liggesmeyer, 2009). La demanda por estos alimentos continuará creciendo y se espera un aumento adicional del 70% para el año 2050 para alimentar a una población mundial de 9 mil millones de personas, de las cuales el 70% vivirá en centros urbanos (Riaz, 2014). Por lo tanto, es necesario mejorar la tecnología de gestión y producción para aumentar el rendimiento de la producción animal (Silanikove, 1992).

La tecnología de los sistemas de sensores se ha utilizado para estimar conductas como el comportamiento de descanso (Ledgerwood, 2010, Rodríguez-Jimenez, 2018), la rumia (Zehner, 2017, Grinter, 2019, Zambelis, 2019) y el comportamiento de alimentación (Bikker, 2014, Büchel, 2014, Mattachini, 2016). Sin embargo, la mayoría de los sistemas de sensores sólo estiman el comportamiento de alimentación como un porcentaje de tiempo durante 24 horas [10]. El registro de las dinámicas individuales de IMS recogidas en tiempo real a través de los sistemas de sensores es una herramienta valiosa para los ensayos de investigación en la producción animal.

Hay sistemas empotrados disponibles en la actualidad que estiman la ingesta de alimentos de los rumiantes, pero estos se basan principalmente en los movimientos del animal en entornos abiertos y no proporcionan una medición directa de lo que el animal está consumiendo (Silanikove, 1994). Normalmente, un sistema empotrado se controla con un microcontrolador (μC) que maneja una o varias funciones específicas en el sistema (Silanikove, 2000). En este sentido, en este documento se propone el desarrollo de un sistema de monitoreo de recogida de datos, como la temperatura y la humedad, así como el peso del alimento a lo largo del experimento nutricional, utilizando una aplicación móvil desarrollada en Android Studio.

2. Análisis de la problemática y enfoque propuesto

En este apartado se resume la problemática detectada con el consumo de ingesta de materia seca en rumiantes y se describe la solución propuesta.

2.1. Descripción del problema.

La IMS es un parámetro fundamental en la investigación de la nutrición animal porque estima el suministro global de nutrientes, especialmente los evaluados durante los experimentos de investigación. La IMS en el ganado lechero es esencial para correlacionar el estado de salud durante periodos de estrés, como el período de transición en las vacas lecheras, además es primordial para calcular la eficiencia en las vacas lecheras (es decir, kg de leche/kg de MS consumida) (Carpinelli, 2019). La IMS de los animales puede verse afectado por las características físicas y químicas (por ejemplo, el contenido de materia seca y fibra) de los ingredientes de la dieta y sus interacciones (Allen, 2000). Estos factores y sus interacciones hacen que sea un reto para cualquier sistema de sensores estimar el IMS real en animales en un formato de dieta completamente mezclada. Por el contrario, los factores mencionados y sus interacciones son menos problemáticos cuando la IMS se mide por la desaparición de la dieta completamente mezclada (es decir, MS ofrecida - MS rechazada) en los experimentos de investigación a través de un sistema de compuertas o un sistema automático de alimentación que puede registrar continuamente la ingesta (Carpinelli, 2019, Schirmann, 2011).

Actualmente, en los laboratorios de experimentación ganadera se utilizan jaulas en las que se realizan experimentos para el análisis de la dieta del animal. Los experimentos que se realizan consisten en colocar al animal en una jaula durante un tiempo para recopilar datos sobre su consumo; cada jaula tiene un comedero y un bebedero. La recogida de datos consiste en pesar el alimento antes de dispensarlo al animal y, pasado un tiempo, pesar el material rechazado, siendo el consumo la diferencia entre estas dos medidas. Sin embargo, para llevar a cabo la recolección de los datos se requiere de gran cantidad capital humano. En este sentido, se propone el desarrollo de un sistema automatizado a través de una aplicación móvil capaz de capturar datos con mayor frecuencia y así determinar la dinámica de consumo y considerar nuevas variables, como la temperatura y la humedad. De esta manera, es posible recopilar una gran cantidad de datos para su posterior análisis.

2.2. Solución propuesta.

Los sistemas automatizados son soluciones innovadoras que suelen utilizar software y hardware para proporcionar altos niveles de fiabilidad de la información sobre la intervención humana. Como ya se mencionó anteriormente el análisis realizado, a continuación, se presenta una propuesta de solución con el diseño de un sistema que permita monitorear y almacenar información sobre el peso del alimento consumido por animales de producción, así como la temperatura y humedad ambiental del entorno del animal con su respectiva fecha y hora de captura sin la intervención humana; a través de una aplicación móvil capaz de configurar el inicio y el fin del experimento, así mismo se presentan diversas funciones de configuración del sistema de monitoreo.

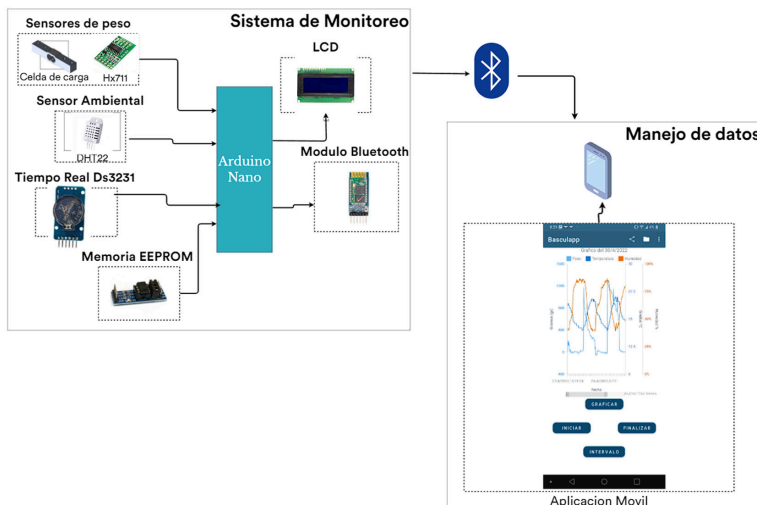


Figura 1 – Diagrama de arquitectura del sistema y aplicación móvil

A continuación, se describe una propuesta de solución que cubre las necesidades antes descritas. Como se puede observar en Figura 1 se presenta el diagrama de arquitectura

del sistema que consta de dos fases: el sistema de monitoreo (Hardware) y el sistema de manejo de datos (Aplicación Móvil). Por un lado, el sistema de monitoreo está construido con los siguientes componentes: Arduino Nano, sensores ambientales, memoria EEPROM, reloj en tiempo real, celda de carga, transmisor y un módulo Bluetooth para establecer comunicación con la aplicación móvil. Por otro lado, la aplicación móvil ha sido desarrollada en el sistema operativo Android para permitir la configuración y manipulación de los datos obtenidos por el sistema de monitoreo.

3. Herramientas de implementación

Actualmente, existen diferentes plataformas de bajo costo permiten el desarrollo de dispositivos tecnológicos; algunas de estas plataformas son Arduino (Arduino, 2023), Raspberry Pi (Raspberry, 2023), OpenPicus (OpenPicus, 2023), Cubieboard (Cubieboard, 2023), Udoo (Udoo, 2023), entre otras. Se realizó un análisis exhaustivo de cada una de las características de las plataformas; los más utilizados y conocidos son Arduino, OpenPicus y Raspberry Pi. Este trabajo se centra en la plataforma Arduino, específicamente en la plataforma de Arduino Nano por la diferencia de tamaño con relación a las otras placas de Arduino.

La placa Arduino Nano se basa en el microcontrolador ATMEGA328 de 8 bits con una memoria de SRAM de 2 KB, flash de 32 KB, EEPROM de 1 KB, la frecuencia de velocidad del reloj es de 16 MHz; tiene 14 pines de entrada/salida que utilizan de 0-5 voltios y cinco pines (3, 5, 6, 9, 11) se pueden configurar como salidas para manejar señales de modulación de ancho de pulso (PWM), cuenta con ocho pines de entrada/salida analógica para medir voltaje analógico en el rango de 0-5 voltios, además, cuenta con dos pines de interrupciones externas (2, 3) para activar la interrupción en un valor bajo o en un cambio de valor y dos pines seriales (0-Rx, 1-Rx) utilizados para recibir y transmitir datos en serie TTL, finalmente, mencionar el pin para restablecer el microcontrolador.

Los protocolos de comunicación inalámbricos dividen la información en paquetes pequeños y la envuelven antes de transmitirla. Hay muchos tipos diferentes de protocolos inalámbricos con sus propias ventajas y desventajas. Algunos ejemplos de protocolos de comunicación inalámbrica incluyen Bluetooth, GSM, ZIGBEE y WIFI (Fernandez, 2013). En este trabajo, en el sistema de monitoreo se utilizó el módulo de Bluetooth HC05 como protocolo de comunicación entre el dispositivo y la aplicación móvil. Este módulo diseñado especialmente para dispositivos de bajo consumo, que requieren corto alcance de emisión y basados en transceptores de bajo costo. Su uso es adecuado cuando puede haber dos o más dispositivos en un área reducida sin grandes necesidades de ancho de banda. Bluetooth tiene la ventaja de simplificar el descubrimiento y configuración de los dispositivos, ya que éstos pueden indicar a otros los servicios que ofrecen, lo que redundante en la accesibilidad de estos sin un control explícito de direcciones de red, permisos y otros aspectos típicos de redes tradicionales. El Bluetooth es utilizado de diversas maneras, pero principalmente es empleado para establecer conexión con un dispositivo móvil, ya sea para pasar datos al dispositivo externo o viceversa [25].

El diseño y construcción del dispositivo del sistema de monitoreo se utilizaron diferentes sensores tales como: sensores de peso (Celda de carga y módulo Hx711), sensor ambiental (DHT22) y un reloj de tiempo real (Ds3231). La celda de carga convierte la carga que se

le aplica sobre ella en señales eléctricas. Estas señales son variaciones de resistencia por lo que es necesario un módulo que pueda convertirlas en datos que pueda procesar un Arduino, por ello se utilizó el módulo transmisor HX711 que funciona como convertidor Analógico - Digital. El módulo DHT22 es un sensor de temperatura y humedad que cuenta con cuatro pines (VCC, Data, NC, GND) que se conectan directamente al microcontrolador de Arduino y se utiliza la librería DHT.h para obtener la temperatura y humedad relativa del ambiente. El módulo DS3231 es un reloj de tiempo real con alta precisión y bajo costo que cuenta con un oscilador de cristal con compensación de temperatura y un cristal integrado. Tiene una entrada de batería que mantiene la hora exacta incluso cuando se corta la alimentación principal. Almacena información sobre segundos, minutos, horas, día, fecha, mes y año. La información se transfiere a través del protocolo I2C mediante un bus bidireccional en serie. Finalmente, se utilizó una pantalla LCD de 20 x 4 adaptada para que funcione con Arduino y se muestren los resultados del sistema de monitoreo.

4. Esquema general del desarrollo de la propuesta

Para el desarrollo del sistema de monitoreo, se utilizó el modelo propuesto por Liggesmeyer y Trapp (2019), con el objetivo de desarrollar mejoras parciales después de cada modificación e implementar los aspectos más importantes del desarrollo. La ventaja de utilizar este modelo es que no modifica el flujo del ciclo de vida y aumenta la probabilidad de éxito porque ofrece a los clientes un prototipo donde se puede probar el sistema y les permite verificar los requisitos específicos y tener más claro el desarrollo. Las etapas del modelo utilizado en el proceso de desarrollo del sistema se describen a continuación:

- **Requerimientos:** se realizó una entrevista a expertos en nutrición animal, donde se obtuvo las ideas principales del funcionamiento del sistema. Después, se desarrolló un prototipo inicial que sirvió para realizar pruebas de funcionamiento y a partir de esto trabajar en mejoras en el sistema de toma de datos y el software encargado de trabajar los datos capturados
- **Diseño Funcional:** una vez definidos los requisitos del sistema, se crearon los diagramas UML para la documentación posterior.
- **Arquitectura del sistema:** se desarrolló una arquitectura para el funcionamiento del sistema de captura de datos, para definirla se utilizó un diagrama de arquitectura, mostrando los componentes utilizados para el desarrollo.
- **Arquitectura de software:** se utilizaron diagramas de flujo para definir el algoritmo a realizar para el software, permitiendo la documentación con la parte lógica del desarrollo.
- **Diseño del software:** se diseñó el protocolo de comunicación capaz de comunicar el sistema con el software, para atender las solicitudes realizadas por los usuarios.
- **Implementación:** se desarrolló el sistema utilizando como base el microcontrolador Arduino Nano, la implementación de los componentes se realizó utilizando el entorno de desarrollo proporcionado por Arduino.
- **Pruebas unitarias:** al codificar cada requerimiento, se realizaron pruebas separadas de cada componente con el fin de comprobar el correcto funcionamiento del sistema, la técnica de prueba utilizada fue la prueba de

componentes.

- Integración y pruebas de software: se integraron los componentes y realizamos las pruebas de integración con la finalidad de encontrar problemas de funcionamiento y comunicación entre los componentes de software, la técnica utilizada fue la prueba de extremo a extremo.
- Integración y pruebas del sistema: se integraron los componentes del hardware y se realizaron pruebas de integración con el fin de encontrar fallas en el mismo, además, se realizaron pruebas en la comunicación entre el software y el hardware, tratándolo como un solo conjunto, la técnica utilizada fue el ejemplo de prueba.
- Pruebas funcionales: se realizaron las pruebas teniendo en cuenta los requisitos solicitados por los expertos en nutrición animal, con el fin de verificar que el sistema cumple con lo solicitado.
- Pruebas de validación: las pruebas de validación se realizaron por expertos en nutrición animal para validar el sistema de monitoreo propuesto, en esta etapa se adquirieron nuevos requisitos para posibles mejoras del sistema.

5. Resultados experimentales

En este apartado se resume las diferentes pruebas realizadas y resultados obtenidos con el sistema de monitoreo tanto en el diseño de conexiones en protoboard como en el diseño de PCB y el desarrollo de la aplicación móvil.

5.1. Diseño en el protoboard del sistema de monitoreo

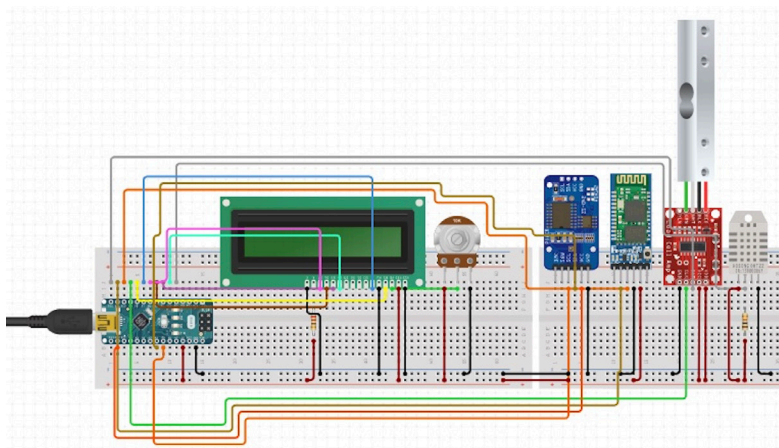


Figura 2 – Diagrama de conexiones utilizando todos los componentes: la placa Arduino Nano, el módulo Bluetooth y los sensores de peso (Celda de carga y módulo Hx711), sensor ambiental (DHT22) y sensor de tiempo real (DS3231)

Primero, se definió la arquitectura del sistema con base en el microcontrolador de Arduino Nano, el módulo de Bluetooth HC05 como protocolo de comunicación y los

diferentes sensores utilizados para cumplir con el objetivo propuesto. La Figura 2 muestra el diagrama de conexiones con todos los componentes utilizados.

5.2. Diseño PCB del sistema de monitoreo.

Una vez comprobado el correcto funcionamiento del diseño de conexiones elaborado en el protoboard con cada uno de los componentes utilizados, se realizó el diseño PCB para generar placas con un tamaño reducido y asegurar una conectividad idónea entre los componentes. En este sentido, se utilizó el software EasyEDA para la creación de un diseño PCB. Primero, se obtuvo el esquemático con todos los componentes utilizados y se realizaron las conexiones de acuerdo con el diseño realizado en el protoboard (Fig. 2). En la Figura 3 se muestra el resultado del sistema desarrollado.

En la Figura 4a se presenta el resultado final del diseño del PCB generado, minimizando al máximo los espacios para obtener una placa de tamaño reducido. La Figura 4b se muestra el resultado del diseño del PCB impreso, este sistema se incrustó en una caja de prototipos de 15 x 6 x 9,9 cm.

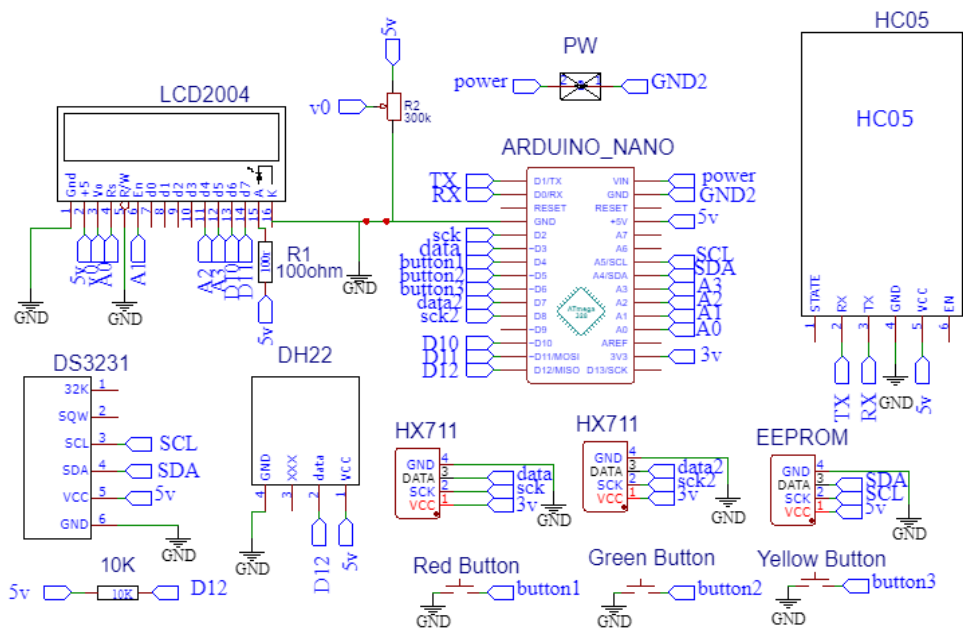


Figura 3 – Diseño esquemático

En la Figura 4a se presenta el resultado final del diseño del PCB generado, minimizando al máximo los espacios para obtener una placa de tamaño reducido. La Figura 4b se muestra el resultado del diseño del PCB impreso, este sistema se incrustó en una caja de prototipos de 15 x 6 x 9,9 cm.

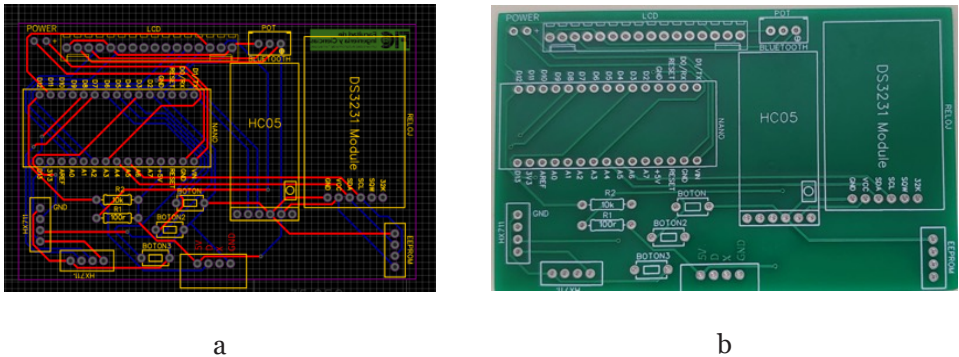


Figura 4 – (a) Diseño del PCB generado, (b) Diseño del PCB impreso.

5.3. Desarrollo de la aplicación móvil

Las primeras pruebas con el sistema de monitoreo se realizaron de manera individual con cada uno de los componentes para comprobar el correcto funcionamiento del sistema. Después, se realizó el diseño de una aplicación móvil en Android Studio para usuarios finales. En la Figura 5 se muestra una secuencia de imágenes con el diseño final de la aplicación, se puede observar la pantalla de inicio con el nombre y los logos de la aplicación y los logos de la Universidad Autónoma de Tamaulipas y de la Facultad de Ingeniería y Ciencias. También se presentan las opciones que se pueden realizar con la aplicación.

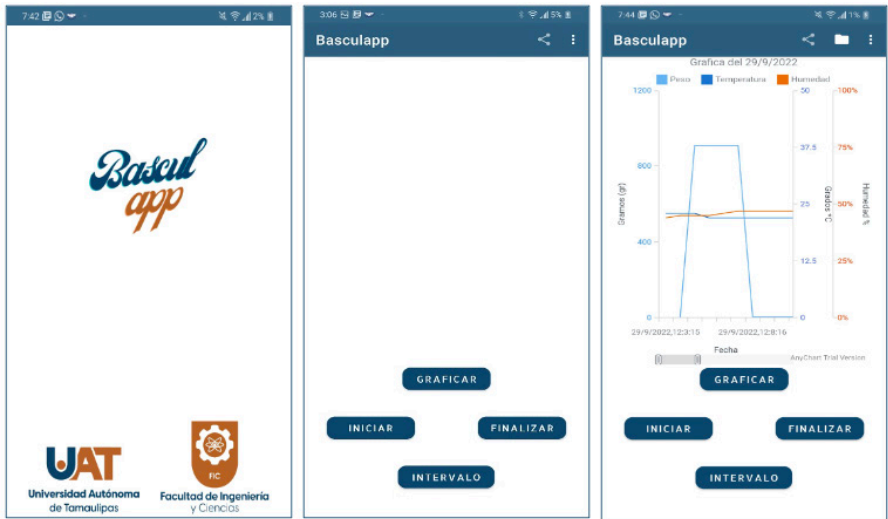


Figura 5 – Diseño final de la aplicación móvil.

En la Figura 6 se muestra las opciones para iniciar un experimento, mediante el ingreso de parámetros como la fecha y hora a través de un calendario para elegir el día de inicio del experimento y definir la finalización del experimento forma programada o de manera inmediata.

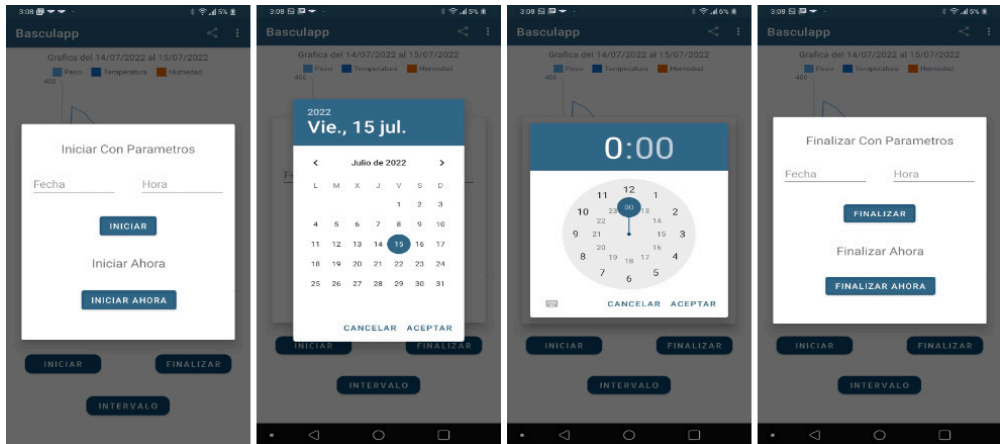


Figura 6 – Pantallas para iniciar y finalizar experimentos.

En la Figura 7 se presenta una secuencia de imágenes que muestran las diferentes opciones con las que cuenta el menú de la aplicación, por ejemplo, la opción para conectarse a un dispositivo que llevará a cabo una actividad, la cual permite visualizar la relación de dispositivos bluetooth disponibles para la conexión. La opción de descargar permite descargar un archivo CSV con los datos capturados por el sistema durante el experimento, además, se puede compartir el archivo CSV a través de correo electrónico o red social de elección. También se pueden realizar cambios de hora del sistema, modificar datos, hacer la tara, consultar diferentes configuraciones y por último desconectarse del dispositivo.

Se diseñó un diagrama de secuencia que describe la forma de comunicación entre el usuario, la aplicación y el dispositivo a través de arduino (Fig. 8). Como se puede observar en la Figura 8, el usuario inicia eligiendo una opción ya sea iniciar, finalizar, intervalo o graficar donde configura los parámetros de fecha y hora del experimento de inicio o de finalización, esta información es enviada al dispositivo utilizando del módulo de bluetooth que se encuentra conectado en el Arduino Nano. De acuerdo con los parámetros recibidos, se ejecuta el algoritmo para calcular la cantidad alimento que animal ha ingerido. Finalmente, se envía la información al usuario y genera un reporte en un archivo CSV con la información solicitada que se puede graficar o compartir por correo electrónico.

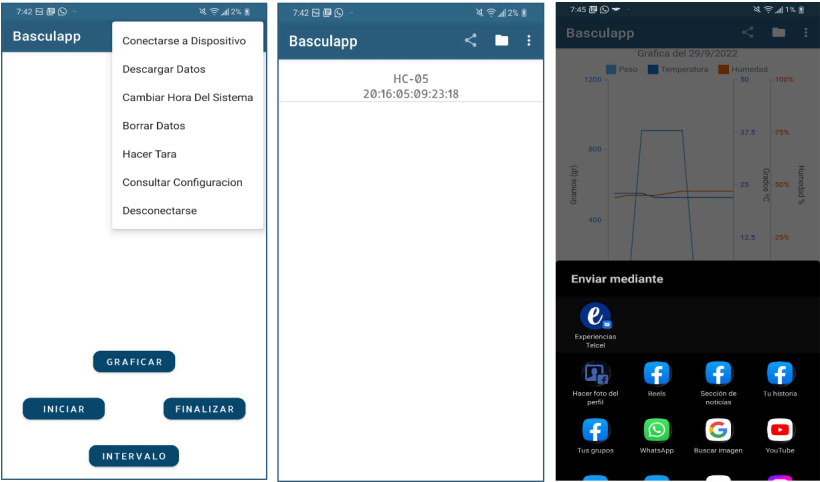


Figura 7 – Menú de opciones para realizar diferentes configuraciones en la aplicación.

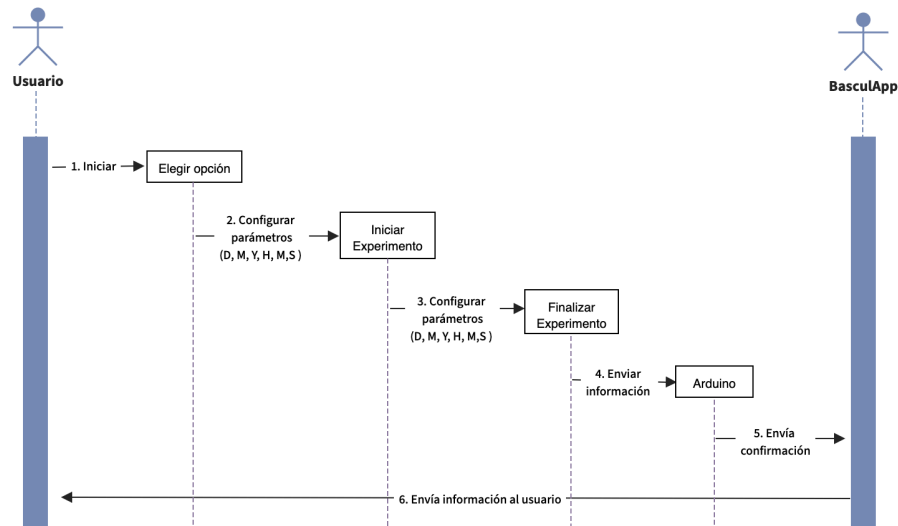


Figura 8 – Diagrama de secuencias de la aplicación móvil.

A continuación, en la Figura 9 se presentan las pruebas de funcionamiento con el prototipo final del sistema instalado en la jaula de metabolismo diseñada para pequeños rumiantes. La prueba se realizó con dos ovejas con un peso vivo promedio de 24.5 kg, que fueron alimentadas con un alimento comercial (Uni-Ovino Engorda, Alimentos Unión Tepexpan®; 15% de proteína) para ovejas en crecimiento. Los datos de consumo de MS se recogieron con un intervalo de muestreo de 15 minutos durante tres días para el animal No. 1 y cuatro días para el animal No. 2.

Finalmente, en la Figura 10 se muestran los resultados obtenidos de la dinámica de consumo de la temperatura y la humedad. De acuerdo con los resultados, se puede observar un pico cuando se ofrece un alimento diario (aproximadamente 1200 g/día). También, se registró un mayor consumo de MS durante las primeras horas del día, y el consumo de materia seca disminuyó con temperaturas cercanas a los 34°. Además, la gráfica muestra una menor cantidad de alimento al final del día, esto se interpreta como que el animal no cumplió con su consumo voluntario de MS. Los porcentajes de humedad registrados estuvieron en el rango de 40 a 85 % durante el período evaluado. Con los datos recolectados el experto en nutrición animal será capaz de ajustar la dieta que mejor se adapte a las características de las ovejas.



Figura 9 – Pruebas de funcionamiento del sistema de monitoreo y la aplicación móvil.

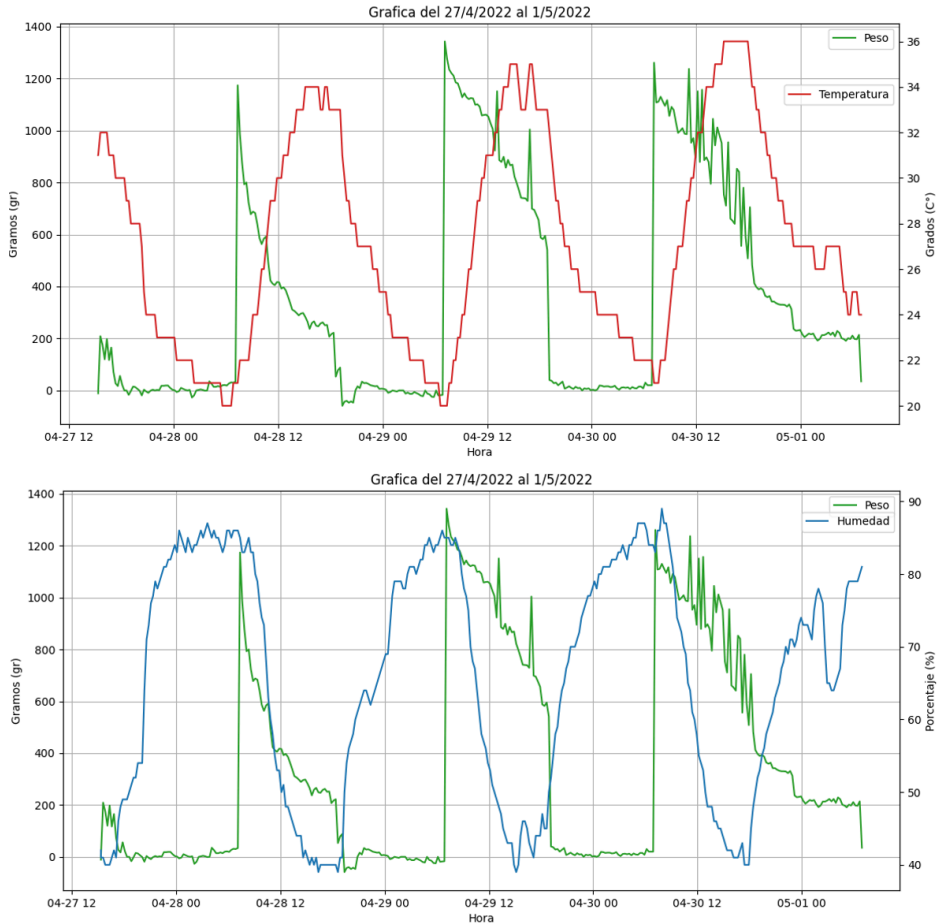


Figura 10 – Dinámica del consumo con relación a la temperatura y la humedad.

6. Conclusiones

En la actualidad, la automatización de procesos está teniendo un crecimiento considerable en diferentes campos de aplicación, debido al uso de sensores y el desarrollo de sistemas es posible reducir la intervención humana, además, de reducir el factor del error obteniendo datos más precisos. En este trabajo se desarrolló un sistema de monitoreo de consumo de materia seca, que consta de un dispositivo electrónico basado en el microcontrolador de Arduino Nano, y utilizando sensores ambientales, una memoria EEPROM, un reloj en tiempo real, una celda de carga, un transmisor y un módulo Bluetooth para establecer comunicación con la aplicación móvil. La aplicación móvil permite la captura de datos como peso de alimento vertido al animal, temperatura y humedad relativa, fecha y hora de captura de datos.

De acuerdo con los experimentos realizados, se obtuvo que el sistema de monitoreo desarrollado es de gran ayuda para los expertos en nutrición animal, recogiendo, a lo largo del tiempo, información esencial, como el peso del alimento y la humedad y temperatura del ambiente donde se alimenta al animal. En general, permite obtener la dinámica de consumo de materia seca, dato fundamental a la hora de analizar las dietas proporcionadas por expertos en nutrición animal.

En futuros trabajos sería interesante generar un filtro para evitar variaciones bruscas generadas, ya sea por el clima o por el movimiento del animal. También, implementar otro método de comunicación entre el dispositivo y la aplicación móvil, utilizando el módulo del Sim900 en lugar del módulo de Bluetooth. Además, se necesitan pruebas más rigurosas donde, por ejemplo, se realicen experimentos en entornos controlados para evaluar la efectividad de los datos obtenidos.

Referencias

- Abdul Hafeez, P., Singh, G., Singh, J., Prabha, C., & Verma, A. (2022). IoT in Agriculture and Healthcare: Applications and Challenges. In *Proceedings of the 3rd International Conference on Smart Electronics and Communication*, Trichy, India (pp. 446-450). IEEE. <https://doi.org/10.1109/ICOSEC54921.2022.9952061>.
- Allen, M.S. (2000). Effects of Diet on Short-Term Regulation of Feed Intake by Lactating Dairy Cattle. *Journal of Dairy Science*, 83, 1598–1624.
- Arduino (2023), <https://www.arduino.cc/>
- Bikker, J.P., van Laar, H., Rump, P., Doorenbos, J., van Meurs, K., Griffioen, G.M., & Dijkstra, J. (2014). Technical note: Evaluation of an ear-attached movement sensor to record cow feeding behavior and activity. *Journal of Dairy Science*, 97, 2974–2979.
- Büchel, S., & Sundrum, A. (2014). Technical note: Evaluation of a new system for measuring feeding behavior of dairy cows. *Computers and Electronics in Agriculture*, 108, 12–16.
- Carpinelli, N.A., Rosa, F., Grazziotin, R.C.B., & Osorio, J.S. (2019). Technical note: A novel approach to estimate dry matter intake of lactating dairy cows through multiple on-cow accelerometers. *Journal of Dairy Science*, 102, 11483–11490.
- Cheng, D., Li, C. & Qiu, N. (2021). The application prospects of NB-IoT in intelligent transportation. In *Proceedings of the 4th International Conference on Advanced Electronic Materials, Computers and Software Engineering*, Changsha, China (pp. 1176-1179). IEEE. <https://doi.org/10.1109/AEMCSE51986.2021.00240>.
- Chin, J., Callaghan, V., & Allouch, S. B. (2019). The Internet-of-Things: Reflections on the Past, Present and Future from a User-centered and Smart Environment Perspective, *Journal of Ambient Intelligence and Smart Environments*, 11(1), 45-69.
- Cubieboard (2023), <http://cubieboard.org/>

- Elkahlout, M., Abu-Saqer, M. M., Aldaour, A. F., Issa, A. & Debeljak, M. (2020). IoT-Based Healthcare and Monitoring Systems for the Elderly: A Literature Survey Study. In *Proceedings of the International Conference on Assistive and Rehabilitation Technologies*, Gaza, Palestine (pp. 92-96). IEEE. <https://doi.org/10.1109/iCareTech49914.2020.00025>.
- Erzİ, H. M. & Aydin, A. A. (2020). IoT Based Mobile Smart Home Surveillance Application. In *Proceedings of the 4th International Symposium on Multidisciplinary Studies and Innovative Technologies*, Istanbul, Turkey (pp. 1-5). IEEE. <https://doi.org/10.1109/ISMSIT50672.2020.9255303>.
- Fernández, A. (2013). *Phyton 3 al descubierto* (2nd Ed.). Alfaomega Grupo Editor.
- Grinter, L.N., Campler, M.R., & Costa, J.H.C. (2019). Technical note: Validation of a behavior-monitoring collar's precision and accuracy to measure rumination, feeding, and resting time of lactating dairy cows. *Journal of Dairy Science*, 102, 3487–3494.
- Hussain, F. & Hammad, M. (2020). Smart healthcare facilities via IOT in the healthcare industry. In *Proceedings of the 3rd Smart Cities Symposium*, Online Conference (pp. 386-391). IEEE. <https://doi.org/10.1049/icp.2021.0945>.
- Kassim, M. R. M. (2020). IoT Applications in Smart Agriculture: Issues and Challenges. In *Proceedings of the Conference on Open Systems*, Kota Kinabalu, Malaysia (pp. 19-24). IEEE. <https://doi.org/10.1109/ICOS50156.2020.9293672>.
- Kirimtat, A., Krejcar, O., Kertesz, A. & Tasgetiren, M. F. (2020). Future Trends and Current State of Smart City Concepts: A Survey. *IEEE Access*, 8, 86448-86467.
- Ledgerwood, D.N., Winckler, C., & Tucker, C.B. (2010). Evaluation of data loggers, sampling intervals, and editing techniques for measuring the lying behavior of dairy cattle. *Journal of Dairy Science*, 93, 5129–5139.
- Liggesmeyer, P., & Trapp, M. (2009). Trends in Embedded Software Engineering. *IEEE Software*. 26, 19–25.
- Mattachini, G., Riva, E., Perazzolo, F., Naldi, E., & Provolo, G. (2016). Monitoring feeding behaviour of dairy cows using accelerometers. *J Agricult Engineer*, 47, 54.
- McCann, J., Quinn, L., McGrath, S., & O'Connell, E. (2018). Towards the Distributed Edge – An IoT Review. In *Proceedings of the International Conference on Sensing Technology*, Limerick, Ireland (pp. 263-268). IEEE. <https://doi.org/10.1109/ICSensT.2018.8603663>.
- OpenPicus (2023), <http://www.openpicus.com/>
- Quiñonez, Y., Lizarraga, C., Peraza, J., & Zatarain, O. (2019). Sistema inteligente para el monitoreo automatizado del transporte público en tiempo real. *RISTI - Revista Ibérica de Sistemas e Tecnologias de Informação*, (31), 94-105.
- Raspberry Pi, (2023), <https://www.raspberrypi.org/>

- Riaz, M.Q., Südekum, K.-H., Clauss, M., & Jayanegara, A. (2014). Voluntary feed intake and digestibility of four domestic ruminant species as influenced by dietary constituents: A meta-analysis. *Livestock Science*, 162, 76–85.
- Ribeiro, D.M., Salama, A.A.K., Vitor, A.C.M., Argüello, A., Moncau, C.T., Santos, E.M., Caja, G., de Oliveira, J.S., Balieiro, J.C.C., Hernández-Castellano, L.E., Zachut, M., Poleti, M.D., Castro, N., Alves, S.P., & Almeida, A.M. (2020). The application of omics in ruminant production: a review in the tropical and sub-tropical animal production context. *Journal of Proteomics*, 227, 103905.
- Rokonuzzaman, M., Akash, M. I., Khatun Mishu, M., Tan, W. S., Hannan, M. A. & Amin, N. (2022). IoT-based Distribution and Control System for Smart Home Applications. In *Proceedings of the 12th Symposium on Computer Applications & Industrial Electronics*, Penang, Malaysia (pp. 95-98). IEEE. <https://doi.org/10.1109/ISCAIE54458.2022.9794497>.
- Rodriguez-Jimenez, S., Haerr, K.J., Trevisi, E., Loor, J.J., Cardoso, F.C., & Osorio, J.S. (2018). Prepartal standing behavior as a parameter for early detection of postpartal subclinical ketosis associated with inflammation and liver function biomarkers in peripartal dairy cows. *Journal of Dairy Science*, 101, 8224–8235.
- Schirmann, K., Chapinal, N., Weary, D.M., Heuwieser, W., & von Keyserlingk, M.A.G. (2011). Short-term effects of regrouping on behavior of prepartum dairy cows. *Journal of Dairy Science*, 94, 2312–2319.
- Shabber, S. M., Bansal, M., Devi, P. M. & Jain, P. (2021). iHAS: An Intelligent Home Automation Based System for Smart City. In *Proceedings of the International Symposium on Smart Electronic Systems*, Jaipur, India (pp. 48-52). IEEE. <https://doi.org/10.1109/iSES52644.2021.00023>.
- Silanikove, N. (1992). Effects of water scarcity and hot environment on appetite and digestion in ruminants: a review. *Livestock Production Science*, 30, 175–194.
- Silanikove, N. (1994). The struggle to maintain hydration and osmoregulation in animals experiencing severe dehydration and rapid rehydration: the story of ruminants. *Exp Physiol*, 79, 281–300.
- Silanikove, N. (2000). Effects of heat stress on the welfare of extensively managed domestic ruminants. *Livestock Production Science*, 67, 1–18.
- Udoo, (2023), <https://www.udoo.org/>
- Vashi, S., Ram, J., Modi, J., Verma, S., & Prakash, C. (2017). Internet of Things (IoT): A vision, architectural elements, and security issues. In *Proceedings of the International Conference on IoT in Social, Mobile, Analytics and Cloud*, Palladam (pp. 492-496). IEEE. <https://doi.org/10.1109/I-SMAC.2017.8058399>.
- Zambelis, A., Wolfe, T., & Vasseur, E. (2019). Technical note: Validation of an ear-tag accelerometer to identify feeding and activity behaviors of tiestall-housed dairy cattle. *Journal of Dairy Science*, 102, 4536–4540.

Zehner, N., Umstätter, C., Niederhauser, J.J., & Schick, M. (2017). System specification and validation of a noseband pressure sensor for measurement of ruminating and eating behavior in stable-fed cows. *Computers and Electronics in Agriculture*, 136, 31–41.

Determinación del mejor algoritmo de detección de matrículas en ambientes controlados y no controlados

Elias Ccoto Hualpa¹, Abel Angel Sullon Macalupu¹, Jorge Eddy Otazu Luque¹, Jorge Sánchez-Garces²

elias.ch@upeu.edu.pe; angeli@upeu.edu.pe; jorgeol@upeu.edu.pe; jasg@upeu.edu.pe.

¹ Escuela Profesional de Ingeniería de Sistemas, Universidad Peruana Unión, Carretera Salida ArequipaKm 6 Chullunquiani, Juliaca, Puno, Perú.

² Facultad de Ingeniería y Gestión Universidad Nacional Tecnológica de Lima Sur, Lima, Peru.

DOI: 10.17013/risti.49.83–99

Resumen: La seguridad es una prioridad en la gestión y parte de esto es el control y monitoreo para detectar situaciones que agredan el bien público o privado. En este sentido el reconocimiento de matrículas de auto se suma a estos sistemas de monitoreo y control. Este trabajo de investigación aplicó una serie de algoritmos de inteligencia artificial para automatizar dicha detección. En este sentido se utilizó funciones de procesamiento de imágenes con el framework OpenCV considerando que las fuentes de información pudieron tener distintos escenarios, siendo que el ambiente de detección es abierto y expuesto a las condiciones meteorológicas de la zona. Las tomas de fotos se realizaron en un ambiente ubicado al sur peruano, cuyas condiciones fueron de lluvia, día soleado, día que cayó el granizo. La base de datos de las imágenes entonces se dividió en dos categorías; ambientes controlados donde se consideró una misma distancia, un solo ángulo, pero no necesariamente un mismo clima; y los ambientes no controlados con diferentes ángulos, diferentes distancias y climas. Al procesamiento de imágenes también se utilizó la transformación morfológica, suavizado gaussiano y umbral gaussiano. Con las imágenes segmentadas y extraídos los dígitos de la matrícula; se comparó 3 algoritmos K-NN, SVM y Tesseract. en cada algoritmo se utilizó sus hiperparámetros para el respectivo reconocimiento de caracteres en las imágenes, se obtuvo los siguientes resultados con imágenes tomadas con distintos ángulos y en distintas luminosidades (ambiente no controlado) donde el mejor Overall accuracy con 86 % y el segundo grupo fueron imágenes tomadas con un ángulo similar y distancias similares (ambiente controlado), obtuvo un Overall accuracy de 95.5 %.

Palabras-clave: KNN; SVM; Tesseract; OpenCV; Machine Learning y hiperparámetros.

Determination of the best license plate detection algorithm in controlled and uncontrolled environments.

Abstract: Security is a priority in management and part of this is control and monitoring to detect situations that attack the public or private good. In this sense, the recognition of car license plates is added to these monitoring and control systems. This research work applied a series of artificial intelligence algorithms to automate such detection. In this sense, image processing functions were used with the OpenCV framework, considering that the information sources could have different scenarios, since the detection environment is open and exposed to the weather conditions of the area. The photos were taken in an environment located in southern Peru, whose conditions were rainy, sunny, and the day the hail fell. The image database was then divided into two categories; controlled environments where the same distance, a single angle, but not necessarily the same climate was considered; and uncontrolled environments with different angles, different distances and climates. Morphological transformation, Gaussian smoothing and Gaussian thresholding were also used for image processing. With the segmented images and the number plate digits extracted; 3 algorithms K-NN, SVM and Tesseract were compared. In each algorithm, its hyperparameters were used for the respective recognition of characters in the images, the following results were obtained with images taken with different angles and in different luminosities (uncontrolled environment) where the best Overall accuracy with 86% and the second group were images taken with a similar angle and similar distances (controlled environment), obtained an Overall accuracy of 95.5%.

Keywords: Tesseract; OpenCV; Machine Learning and hyperparameters.

1. Introducción

En la actualidad hay un aumento en vehículos teniendo un aumento proporcional; según (SUNARP, 2020) en Perú hasta el año 2019 hay más de 701 mil vehículos registrados teniendo un aumento del 6.4 %, comparado al 2018, por lo tanto, las infracciones vehiculares y el robo de vehículos incrementaron. Según (INEI, 2021) informa que, hubo más 20000 denuncias de vehículos robados en el Perú durante el 2019 teniendo un promedio de crecimiento de 5 % anual; tan solo en la capital de Perú se registró 17000 denuncias por robo, siendo el departamento del país con más denuncias. Agravando el problema con la clonación de placas según (SEMARNAT, 2006).

El desarrollo de modelos computacionales permiten resolver el problema planteado, automatizando el proceso de detección para identificar los objetos involucrados en posibles actos de delincuencia, junto a esto se evalúan estos modelos para que la detección esta una excelente precisión, pero previo al uso de estos modelos de detección es necesario contar con algoritmos que permiten el procesamiento de las imágenes siendo que estas pueden ser tomadas en distintos escenarios , climas, distancias, ángulos. Se realizó una investigación de los trabajos actuales sobre este tipo de soluciones.

El autor (Agarwal et al., 2018) utilizó el procesamiento morfológico(escala de grises), y el autor (Divya et al., 2021) utilizó tesseract esto para el reconocimiento de manuscritos, el autor (Varma et al., 2020) realizó sus investigaciones para el reconocimiento de matrículas con imágenes capturadas en carreteras de Irak esto en ambientes no controlados usando trasformaciones morfológicas, el autor (Silva & Jung, 2020) para mejorar el reconocimiento de matrículas en ambientes con un mínimo de luminosidad utilizó redes neuronales.

Los autores citados mencionaron el uso de algoritmos de aprendizaje profundo, supervisados y no supervisados, pero no diferenciaron escenarios como controlados y no controlados, que para la implementación en ambientes abiertos y públicos es necesario validar dicho aspecto.

Por lo tanto, la motivación fue hallar un modelo de detección de placas de vehículos para ambientes controlados adecuada y un modelo para detección de placas para ambientes no controlados.

Para esta comparación se usó 2 tipos de algoritmos; siendo el primero el Tesseract con su algoritmo LSTM(Long short-term memory) y los algoritmos supervisados de Support vector machine y K-nn(K-Nearest Neighbors); ambos tuvieron con OpenCv el procesamiento de imágenes con los algoritmos de transformación morfológica, suavizado gaussiano y umbral gaussiano. Con estos algoritmos la precisión con que se detecta los dígitos de la matrícula del auto es mucho mejor se en ambientes controlados o no controlados; logrando otorgar una solución para el monitoreo y acceso de los vehículos a espacios públicos o privado; tales como universidades, supermercado y zonas de aparcamiento con el objetivo de reducir las infracciones y robos.

2. Materiales

- Cámara para la captura de imágenes en ambiente controlados: Cámara YCC365, HD, 1280*1024
- Cámara para la captura de imágenes en ambiente no controlados: Smart phone LG Q60, cámara principal 16MP + 2MP + 5MP y cámara frontal 13MP

3. Muestra

Para hallar un modelo adecuado para cada ambiente ya mencionado en la introducción, siendo el objetivo del estudio se consideró los valores de cada hiperparámetro de los algoritmos para la detección y reconocimiento de caracteres. El proceso se describe en la tabla 01.

Algoritmo	hiperparámetros	Valores
KNN	K.	1,2,3,4,5,6,7,8,9,10
	C.	1, 10, 100, 1000, 0.1, 1
	Gamma.	0.001, 0.0001, 1, 0.1, 0.01, 1
SVM	Kernel.	cv2.ml.SVM_LINEAR, cv2.ml.SVM_RBF, cv2.ml.SVM_CHI2, cv2.ml.SVM_SIGMOID, cv2.ml.SVM_POLY
	Type.	cv2.ml.SVM_EPS_SVR, cv2.ml.SVM_C_SVC, cv2.ml.SVM_NU_SVC, cv2.ml.SVM_ONE_CLASS, cv2.ml.SVM_NU_SVR
	Degree.	1, 2, 3
	P.	0, 0.1
	Psm.	1,2,3,4,5,6,7,8,9,10,11,12,13
LSMT	Psm.	1,2,3,4,5,6,7,8,9,10,11,12,13

Tabla 1 – Valores de los hiperparámetros de los algoritmos

Una vez obtenido los valores de los hiperparámetros; se tomó la primera muestra, donde se utilizó 80 imágenes en ambientes no controlados estas tomadas con distintos ángulos, distintas distancias. El objetivo fue hallar el mejor hiperparámetro de cada algoritmo descrito en la tabla 02, aplicando los valores mencionados de los hiperparámetros a cada una de las imágenes de las 80.

Algoritmo	hiperparámetros	Cantidad hiperparámetros	Muestra Total
KNN.	K.	10	$80 \cdot 10 = 800$
SVM.	C, Gamma, Kernel, Type, degree y P.	144	$80 \cdot 144 = 11520$
LSMT.	Psm.	13	$80 \cdot 13 = 1040$

Tabla 2 – Muestras por cada algoritmo de la primera ejecución del modelo

Para las muestras en controlados se utilizó 100 imágenes con distancias similares (3 a 5 metros). La cámara que se utilizó fue colocada en control de ingreso de una institución educativa peruana. El objetivo fue hallar los mejores hiperparámetros de cada algoritmo en la totalidad de las muestras descrito en la tabla 03.

Algoritmo	hiperparámetros	Cantidad hiperparámetros	Muestra Total
KNN	K	10	$100 \cdot 10 = 1000$
SVM	C, Gamma, Kernel, Type, degree y P	144	$100 \cdot 144 = 14400$
LSMT	psm	13	$100 \cdot 13 = 1300$

Tabla 3 – Muestras por algoritmo para la segunda ejecución del modelo.

4. Metodología

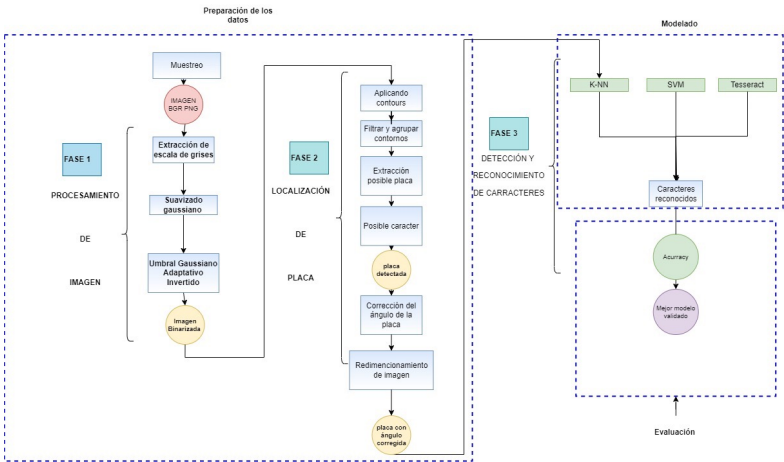


Figura 1 – Método de la investigación

En la figura 01 se muestra la metodología que se utilizó en esta investigación y donde tiene 3 fases principales utilizadas de la metodología CRISP-DM, se utilizó esta metodología por la flexibilidad que tiene de poder adaptarla al contexto situacional del problema que se abordó. Estas 3 fueron la preparación de los datos y consistió en el procesamiento de imágenes, la localización de placas; la segunda fue el modelado que consistió en la codificación y comparación de tres algoritmos para el reconocimiento de las placas y por último la evaluación que pudo encontrar el mejor hiperparámetro en cada muestra del ambiente controlado y no controlado para obtener el mejor modelo de predicción.

4.1. Nivel de investigación

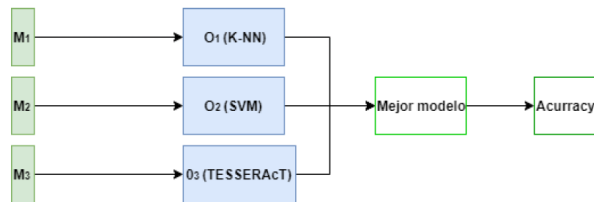


Figura 2 – Nivel de Investigación

Donde:

M = Muestra

O = Observación

La investigación es cuantitativa de nivel descriptiva comparativo porque se realizaron análisis comparativos entre algoritmos computacionales determinando el algoritmo prevalente en función al mejor accuracy así obteniendo el mejor modelo computacional.

4.2. Preparación de los datos: Tratamiento de imagen



Figura 3 – Vehículo a detectar

Para el tratamiento de imágenes se usa herramientas de OpenCv(The Open Computer Vision). Según (Arévalo et al., 2016), la librería OpenCV permite el procesamiento de imágenes, así como su tratamiento así pudiendo desarrollar aplicaciones de visión por computador. La entrada fue una imagen RGB, en formato PNG.

- I. *Extracción de escala de grises*: en esta fase a la imagen entrante se le aplica la función `imgGrayscale`, así se obtiene cada píxel tiene un valor entre 0 y 255, siendo que los valores más cercanos a cero son los más oscuros y los valores más próximos a 255 son los más claros.
- II. *Suavizado gaussiano*: Utiliza la función gaussiana lineal. La principal función del filtrado gaussiano es reducir los detalles y ruidos en las imágenes. La utilización del filtro gaussiano a una imagen tiene la ventaja de evitar errores en la detección de la placa.
- III. *Umbral Gaussiano Adaptativo Invertido*: El objetivo de aplicar la umbralización en una imagen es crear una imagen binaria a partir de la imagen en escala de grises. Este proceso se conoce como binarización de imágenes. Se toma una ventana de tamaño predefinido y se encuentra una suma ponderada de píxeles vecinos para realizar el umbral adaptativo.

```
cv2.adaptiveThreshold(image, 255,
```

```
cv2.ADAPTIVE_THRESH_GAUSSIAN_C,
```

```
cv2.THRESH_BINARY_INV,
```

```
blockSize, WEIGHT)
```

Aquí:

- `ADAPTIVE_THRESH_GAUSSIAN_C` El umbral será en función de Gaussian;
- `BINARY_INV` donde se binariza y se invierte la binarización;
- `BLOCK_SIZE` es el tamaño de la ventana de umbral.
- `WEIGHT` se usa para calcular la suma ponderada de los valores en la zona.

4.3.Preparación de los datos: Localización de placa

Luego del último etapa de hacer unos arreglos a las imágenes, Umbral Gaussiano Adaptativo Invertido, dio como resultados una imagen binarizada, con valores de cero y 255. La imagen será utilizada como entrada para la etapa de detección y reconocimiento.

- I. *Aplicando contours*: Un contorno es un vínculo de puntos de igual intensidad a lo largo del límite, (Phangtriastu et al., 2017). En OpenCV, encontrar contornos es como encontrar un objeto blanco del negro fondo, por lo tanto, durante la

etapa de Umbral Gaussiano Adaptativo, se tuvo que aplicar la operación de Inversión. La función `cv2.drawContours()` permitió aplicar contornos a una imagen binarizada.



Figura 4 – Imagen binarizada

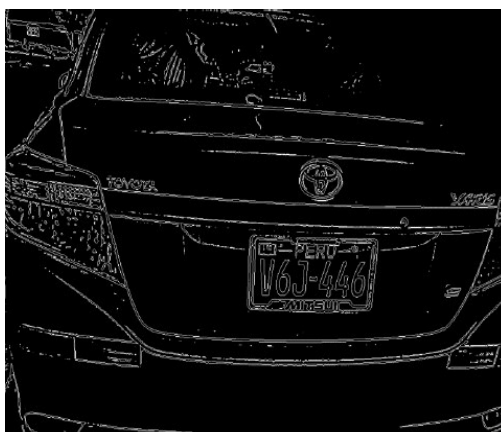


Figura 5 – Imagen Contorno

- II. *Filtrar y agrupar contornos*: Para zonas pequeñas, especialmente en los bordes afilados y valores atípicos de ruido, se aplican contornos. Así, el ojo humano puede calcular fácilmente que tales contornos son innecesarios, pero esto debe ser incorporado en el algoritmo (Abedin et al., 2018). Inicialmente, se aplicaron cuadros delimitadores a cada contorno. Luego, para cada contorno, se consideraron los siguientes factores, como el área mínima del contorno, el ancho y la altura mínimos del contorno, las relaciones de aspecto mínima y máxima posibles.

```
cv2.findContours(imgThreshCopy cv2.
```

```
RETR_LIST cv2.CHAIN_APPROX_SIMPLE)
```

Esto resultó en el filtrado de la mayoría de los contornos innecesarios, acercándonos a nuestro objetivo: detectar una placa. La segunda etapa del filtrado consta en comparar cada contorno con cualquier otro contorno en parámetros como la distancia entre los contornos. La tercera etapa es agrupar los contornos, un grupo de contornos satisface todas estas condiciones, se agrupan en uno. Es posible que se puedan obtener dos o más de tales grupos.

- III. Extracción de posible placa: para esta fase se enfocó en la ubicación de las posibles placas utilizando la función `cv2.boundingRect(self.contour)`. Esta permite ubicar las placas dentro de rectángulos, donde fueron posicionadas las imágenes.

`cv2.boundingRect(self.contour)` la función detecta los rectángulos dentro de nuestra imagen, la función `contour` devolverá: coordenada X y Y, tanto la altura y el ancho y altura encerrando la placa en un rectángulo así como se ve en la figura 2.

```
IntboundingRectX = intX.
```

```
IntboundingRectY = intY.
```

```
IntboundingRectWidwt = intWidwt.
```

```
IntboundingRectHeight = intHeight.
```



Figura 6 – Rectángulo detectado

Obtenida todas las coordenadas y las medidas se pasan a la siguiente fase que es las extraer la posible placa delimitando y recortada del resto de la imagen como se puede ver en la figura 3.



Figura 7 – Placa extraída

- IV. *Posible caracter*: En esta fase se presenta otra dificultad que la identificación de la placa correcta de un n resultados de la fase anterior. En esta fase se ubicó la placa correcta, para se identifiquen los posibles caracteres dentro de la placa, la posible placa será la que contenga más caracteres para esto se emplea la función.

```
Cv2.findContours(imgThreshCopy, Cv2.RETR_LIST,
cv2.CHAIN_APPROX_SIMPLE)
```

La función `findContours` identifica y muestra los contornos de la placa, para esto se usan parámetros: la primera `Cv2.RETR_LIST` ayuda a solo obtener los contornos que no sean dependientes de otros; el otro parámetro que se usa es `cv2.CHAIN_APPROX_SIMPLE` que mostrará las coordenadas finales de cada contorno, logrando segmentar y hacer el conteo de los caracteres, así como se muestra en la figura 4; una vez realizado se obtienen las posibles placas para la identificación de los caracteres.



Figura 8 – Caracteres detectados

- V. *Corrección del ángulo de la placa*: El propósito es corregir los ángulos estandarizando las posibles placas para su fácil detección. Se mantiene la relación de longitudes entre los puntos que residen en una línea recta. Sin embargo, los ángulos dentro de las líneas y las longitudes dentro de los puntos no se conservan. En OpenCV, la corrección del ángulo de la placa se puede lograr mediante la función `getRotationMatrix2D`.

`Cv2.getRotationMatrix2D(tupleCenter)`

`floatCorrectionAnglenDeg 1.0)`

- VI. *Redimensionamiento de la imagen*: en estas fases se cambiaron los tamaños de la placa extraída de 20x30 en las variables `RISIZED_CHAR_IMAGE_WIDTH` y `RISIZED_CHAR_IMAGE_HEIGHT`.

Esta función garantizó que todas las imágenes sean uniformes para formato de entrada para el modelo de aprendizaje, este proceso es necesario ya que las imágenes tomadas en ambientes no controlados fueron en distintos ángulos y distancias así perdiendo su uniformidad.

4.4. Funciones de código detección de los caracteres

Modelado: Detección y reconocimiento de caracteres

Se utilizó subprocesos en donde se usaron los algoritmos de OpenCv y los de Tesseract para el reconocimiento de caracteres de la placa.

K-NN: Según (Kumar Sahoo, 2020) es el algoritmo más sencillo para la técnica de aprendizaje automático es k-NN. La clasificación de los caracteres se realiza mediante el cálculo de la distancia Euclidiana.

La ventaja de un valor de k mayor es que se suprime el ruido, pero a la vez los límites de los caracteres serán menos distinguibles. La búsqueda aleatoria toma como entrada el modelo y una lista de hiperparámetros que se prueban durante el entrenamiento. En esta fase, se probaron valores positivos diferentes para los parámetros K, Hallando el valor para óptimo K para el modelo.

SVM: Support Vector Machine es un método de aprendizaje supervisado para el análisis de clasificación y regresión (Tabrizi & Cavus, 2016). En SVM, los datos se representan en un espacio n-dimensional donde se puede predecir si un nuevo ejemplo de entrenamiento cae en la misma categoría o en otra diferente. El objetivo principal de SVM es encontrar un hiperplano en el espacio n-dimensional que pueda clasificar los puntos de datos. (Singh et al., 2021) menciona que se podría elegir varios hiperplanos potenciales para distinguir las dos clases de puntos de datos. Pero el hiperplano ideal es el que maximiza el margen entre los puntos de datos de ambas clases. Los hiperplanos son límites para la toma de decisiones que ayudan a distinguir los puntos de datos. Los puntos de datos que caen a ambos lados del hiperplano pueden atribuirse a varias clases. La dimensión del hiperplano depende del número de características. El hiperplano se puede encontrar usando la siguiente ecuación.

Para obtener la mejor predicción para SVM se utilizaron distintas combinaciones de hiperparámetros siendo: Kernel, gamma, C, type1, degree y P obteniendo 144 de posibles combinaciones el valor de estos hiperparámetros

#	Hiperparámetros	Valores
1	C	1, 10, 100, 1000 y 0.1
2	Gamma	0.001, 0.0001, 0.1, 0.01 y 1
3	Kernel	Lineal, Polynomial, Sigmoid, RBF y CHI2
4	Type	EPS_SVR, C_SVC, NU_SVC, ONE_CLASS y UN_SVR
5	Degree	1, 2 y 3
6	P	0 y 0.1

Tabla 4 – Valores de los hiperparámetros de svm

TESSERAC: Tesseract es un motor OCR y es un código abierto más usado sobre todo en imágenes que contienen textos (Divya et al., 2021).

Se usó el parámetro PSM (page segmentation modes, en inglés) de la librería Pytesseract que hace referencia a la forma de segmentación en este caso de la placa. Cada número refiere a una forma de segmentación.

Estos parámetros se utilizaron para hallar el modelo óptimo mediante el accuracy para la detección y reconocimiento de placas.

5. Resultados

En esta parte se presentaron los resultados experimentales de la localización de matrículas y, además, se presentó una comparación de los algoritmos SVM, KNN y el algoritmo de Tesseract. Para cada algoritmo se utilizó las mismas funciones de procesamiento de imágenes con OpenCV; donde se obtuvo imágenes segmentadas y binarizadas, Estas imágenes fueron tomadas como input para cada algoritmo, así como se muestra en la figura 7.

5.1. Primera Ejecución del Modelo

Para esta primera ejecución se utilizó ochenta imágenes con ángulos y distancias distintas

KNN: Se usó el hiperparámetro k con 10 valores que se puede ver en la tabla 1, obteniendo el valor del hiperparámetro con el valor de 4 con el mejor valor Overall Accuracy de 0,841 tal como se muestra en la tabla 5.

#	K	Accuracy
1.	K4.	0,841
2.	K3.	0,818
3.	K1.	0,807
4.	K5.	0,798
5.	K2.	0,782
6.	K7.	0,763
7.	K6.	0,760

#	K	Accuracy
8.	K8.	0,751
9.	K9.	0,747
10.	K10.	0,733

Tabla 5 – Accuracy de los hiperparámetros de knn.

Hecha las pruebas el k-vecinos resulto tener un valor de 4 como el más óptimo, este favor fue el más óptimo para el reconocimiento de caracteres, esto tiene sentido ya que cuando K es más cercano al valor de 1, la predicción puede ser más exacta, pero a la vez es más probable que pueda confundir caracteres similares como por ejemplo la “o” con la “O” ya no tiene más K-vecinos para validar su predicción; al contrario, cuando K es más alejado de 1 tiende a errar la predicción ya que tiene demasiados K-vecinos que confunde al algoritmo, esto fue demostrado en los resultados detallados en la tabla 2; por lo cual el valor adecuado para K fue 4 por alejado al 10 y cercano a 1.

SVM: Se usaron 6 hiperparametros con valores diferentes detallado en la tabla 1, obteniendo 144 combinaciones de hiperparametros, estas fueron probadas con las 100 imágenes y resultó un total de 11520 muestras para la evaluación del modelo SVM (ver tabla 3). La mejor combinación de hiperparámetro tuvo los siguientes valores: C = 1, Gamma = 1, Kernel = Lineal, Type1 = C_SVC, Degree = 1 y P = 0 teniendo un Overall accuracy de 0,864 así como se detalla ver tabla 5.

#	C	gamma	kernel	type1	degree	P	Accuracy
1	1	1	Lineal	C_SVC	1	0	0,864
2	1	0,001	Lineal	C_SVC	3	0	0,844
3	100	1	Lineal	C_SVC	1	0	0,842
4	1	0,001	Lineal	C_SVC	1	0	0,842
5	1	1	Lineal	C_SVC	2	0	0,841
6	100	1	Lineal	C_SVC	2	0	0,840
7	100	1	Lineal	C_SVC	3	0	0,840
8	10	0,0001	Lineal	C_SVC	1	0	0,840
9	10	0,0001	Lineal	C_SVC	2	0	0,840

Tabla 6 – Accuracy de los hiperparámetros de svm.

Resultó con un Kernel lineal, siendo este fue el más rápido y sencillo a la hora del mapeo de muestras; type1 fue C_SVC, el hiperparámetro trabaja adecuadamente con un Kernel Lineal y depende de los valores del hiperparámetro C; si C es cercano a cero no se consideran puntos cercanos del núcleo en el mapeo así como se muestra en la figura 7 y cuanto más alejado de cero los puntos más cercanos serán considerados en el mapeo; El hiperparámetro C dió como valor 1 siendo el óptimo para nuestro modelo ya que este no cuenta con muchas penalizaciones en el mapeo siendo más preciso en los resultados.

TESSERACT: se considero el parametro (psm), teniendo valores del 1 al 13, donde se consiguió 1040 muestras para su evaluación. Resultó tener parametro psm con un valor 9 con un Overall accuracy de 0.479 siendo el más adecuado para este framework así como se detala en la tabla 7.

#	psm	Accuracy
1	9	0,479
2	12	0,471
3	7	0,470
4	5	0,467
5	6	0,466
6	8	0,431
7	13	0,376
8	11	0,371

Tabla 7 – Accuracy del hiperparámetros de tesseract.

Este valor 9 de PSM considera a la imagen extraída como una sola palabra en un círculo, siendo este el parámetro adecuado en el reconocimiento de caracteres, pero no alcanzó un Overall accuracy alto para ser óptimo para el modelo, esto porque Tesseract fue desarrollado para detección de textos donde no haya ruidos o manchas en la imagen y en fondos uniformes.

Segunda Ejecución del Modelo

Las Imágenes utilizadas para esta segunda ejecución se tomó 100 imágenes con distancias y ángulos similares resultando un nuevo Overall accuracy para cada uno del algoritmo se detalla en la tabla 8.

Algoritmo	Muestras	Accuracy
SVM.	100	0,955
K-NN.	100	0,863
Tesseract.	100	0,292

Tabla 8 – Overall accuracy de los mejores hiperparámetros

Se muestra en tabla 8 los resultados que variaron específicamente en uno de los algoritmos que fue SVM donde se logró obtener un 95.5% de precisión en el reconocimiento de caracteres obteniendo una mejora a comparación con las muestras de la primera ejecución, SVM tuvo una mejora considerable en la precisión al momento de discernir caracteres confusos o similares (o, O, I, 1, M, V, L,), Pero KNN tuvo más dificultad con en estos caracteres, aun así experimentó un aumento en la precisión al 86.3% en cambio, en Tesseract hubo disminución en su Overall accuracy, esto debido a que tuvo problemas con los caracteres especiales (!, ", \$, %, , /,(,),=,?, !); tal como ya se explicó este último

algoritmo es mayormente usado para manuscritos o textos en superficies homogéneas. Se describe mayores detalles en la tabla 9.

Modelo	Hiperparámetro	Acc	Observaciones
SVM,	C=1, Gamma=1, Kernel=lineal, Type1=SVM_CSVC; degree =1y P = 0	95.5%	El algoritmo obtuvo un resultado adecuado con caracteres similares y confusos como son: o, O, L, I,1, J entre otros. Obtenido un Overall accuracy de 95.5% y siendo el adecuado para usar en el modelo.
KNN,	K=4	86.3%	El algoritmo de KNN experimento dificultades a la hora de discernir caracteres similares como por ejemplo la o con la O obteniendo un accuracy de 86.3% Overall esto depende de la forma que se captura la imagen que ya que influye los ruidos de las imágenes.
Tesseract	psm=9	29.2%	Tesseract experimento muchos problemas con los ruidos en la imagen, esto porque cualquier mancha en la imagen es considerada como caracteres especiales como son "\$, +, *, -", entre otros por lo cual se obtuvo un Overall accuracy de 29.2%

Tabla 9 – comparación de algoritmos.

6. Discusiones

La motivación del trabajo fue la identificación las placas de los vehículos en ambientes controlados y para ambientes no controlados; en ambos casos se utilizaron funcionalidades de procesamiento de imágenes que pudieron mejorar el reconocimiento de los caracteres eliminando los ruidos.

Según los resultados mostrados se pudo obtener el mejor modelo para el reconocimiento de placas, este modelo fue con el algoritmo de SVM, con sus mejores hiperparámetros, siendo el Kernel lineal como el más adecuado, según (Tabrizi & Cavus, 2016) el Kernel lineal es uno de los más simples, pero, el más rápido al momento del mapeo de muestras; también menciona que el kernel lineal es utilizado cuando las muestras son linealmente separables siendo el caso en las muestras utilizadas.

Así mismo los resultados mostraron que el hiperparámetro C tuvo un valor de 1, el uso de este valor del hiperparámetro fue comprobado por (Chougule & Shah, 2019) obteniendo un acierto de 98% en la clasificación y detección de huellas dactilares; por otra parte uno de los factores que influyeron según (Salau et al., 2021) es la forma en la que las imágenes fueron capturadas, también (Khare et al., 2019) indica que otro factor es el clima al momento de capturar las imágenes, esto es validado por (Silva & Jung, 2020) donde su investigación fue realizada en climas variados y pudo obtener un porcentaje alto al momento de reconocer placas de vehículos. Para tener un mejor overall accuracy, se tuvo que realizar procesamiento de imágenes por ejemplo el autor (Agarwal et al., 2018) logra un 93% realizando procesamiento morfológico, del mismo modo (Farhat et al., 2018) que obtuvo un 95% y (Varma et al., 2020) tuvo un 98%.

7. Conclusiones

El desarrollo de modelos computacionales permite resolver uno de los problemas bandera de la sociedad actual que es la inseguridad pública, entonces nace la pregunta ¿es posible conseguir una buena precisión para que el modelo pueda detectar de forma óptima los dígitos de una matrícula de auto en condiciones controladas y no controladas? Para el problema planteado se determinaron una serie de elementos importantes para lograr un modelo de detección computacional con una métrica de Overall accuracy (OA) de excelente.

Primero se dividió el trabajo en tres fases (procesamiento de imágenes, localización de la placa de auto en toda la imagen, detección de los dígitos de la matrícula); fue importante ubicar el momento en que se aplicaron las distintas técnicas; esto lo ofreció el primer elemento metodológico; el segundo fueron las técnicas algorítmicas para procesar la imagen y estas fueron transformación morfológica, suavizado gaussiano y umbral gaussiano; consiguiendo una imagen visiblemente clara y con el menor ruido a pesar de las condiciones adversas climatológicas o distancias y ángulos diferentes.; el tercer elemento fue la aplicación del muestreo de las imágenes tomadas, a pesar que fueron 80 imágenes en la primera categoría de ambiente controlado y 100 imágenes en la categoría de ambiente no controlado; al aplicar el muestreo; la muestra llegó a tener 1000 combinaciones para KNN, 14400 para SVM y 1300 para Tesseract, haciendo las corridas con estas combinaciones de los hiperparámetros a cada imagen, con esto aumentó la posibilidad de tener un mejor OA y conseguir una detección excelente.

Se llegó a la conclusión que el mejor algoritmo fue el de SVM, cuyo objetivo principal es encontrar un hiperplano en el espacio n-dimensional que pueda clasificar los puntos de datos; logrando agrupar los puntos y determinando el dígito a ser predecido. La tabla 7 describe una tasa de reconocimiento con promedio de aproximadamente el 86.4% con imágenes tomadas de distintos ángulos y distancias, pero con las muestras con un ángulo recto y distancia similares hubo un aumento considerable obteniendo un Overall accuracy del 95.5 (tabla 9)

Referencias

- Abedin, M. Z., Nath, A. C., Dhar, P., Deb, K., & Hossain, M. S. (2018). License plate recognition system based on contour properties and deep learning model. 5th IEEE Region 10 Humanitarian Technology Conference 2017, R10-HTC 2017, 2018-Janua, 590–593. <https://doi.org/10.1109/R10-HTC.2017.8289029>
- Agarwal, P., Chopra, K., Kashif, M., & Kumari, V. (2018). Implementing ALPR for detection of traffic violations: A step towards sustainability. *Procedia Computer Science*, 132, 738–743. <https://doi.org/10.1016/j.procs.2018.05.085>
- Arévalo, V., González, J., & Ambrosio, G. (2016). La librería de visión artificial OPENCV. 1–6. https://www.researchgate.net/publication/236668252_La_Libreria_a_de_Vision_Artificial_OpenCV_Aplicacion_a_la_Docencia_e_Investigacion_in_spanish

- Chougule, A., & Shah, M. (2019). Local Binary Pattern with Hyperparameter Tuned Support Vector Machine for Fingerprint Classification. 2019 International Conference on Intelligent Computing and Control Systems (ICCS), Iccics, 1084–1087. <https://doi.org/10.1109/ICCS45141.2019.9065509>
- Divya, P., Varma, M., Ratna Mouli, U., Srinivas, Garima, Nikhil, & Vishistha. (2021). Web based optical character recognition application using flask and tesseract. Materials Today: Proceedings <https://doi.org/10.1016/j.matpr.2020.10.850>
- Farhat, A., Hommos, O., Al-Zawqari, A., Al-Qahtani, A., Bensaali, F., Amira, A., & Zhai, X. (2018). Optical character recognition on heterogeneous SoC for HD automatic number plate recognition system. Eurasip Journal on Image and Video Processing, 2018(1). <https://doi.org/10.1186/s13640-018-0298-2>
- INEI. (2021). PERU Instituto Nacional de Estadística e Informática INEI. <https://www.inei.gob.pe/estadisticas/indice-tematico/vehicle-theft/>
- Khare, V., Shivakumara, P., Chan, C. S., Lu, T., Meng, L. K., Woon, H. H., & Blumenstein, M. (2019). A novel character segmentation-reconstruction approach for license plate recognition. Expert Systems with Applications, 131, 219–239. <https://doi.org/10.1016/j.eswa.2019.04.030>
- Kumar Sahoo, A. (2020). Automatic recognition of Indian vehicles license plates using machine learning approaches. Materials Today: Proceedings. <https://doi.org/10.1016/j.matpr.2020.09.046>
- Phangtriasu, M. R., Harefa, J., & Tanoto, D. F. (2017). Comparison between Neural Network and Support Vector Machine in Optical Character Recognition. Procedia Computer Science, 116, 351–357. <https://doi.org/10.1016/j.procs.2017.10.061>
- Salau, A. O., Yesufu, T. K., & Ogundare, B. S. (2021). Vehicle plate number localization using a modified GrabCut algorithm. Journal of King Saud University - Computer and Information Sciences, 33(4), 399–407. <https://doi.org/10.1016/j.jksuci.2019.01.011>
- SEMARNAT. (2006). Cuidado con los plaguicidas.
- Silva, S. M., & Jung, C. R. (2020). Real-time license plate detection and recognition using deep convolutional neural networks. Journal of Visual Communication and Image Representation, 71. <https://doi.org/10.1016/j.jvcir.2020.102773>
- Singh, K. R., Neethu, K. P., Madhurekaa, K., Harita, A., & Mohan, P. (2021). Parallel SVM model for Forest Fire Prediction. Soft Computing Letters, 100014. <https://doi.org/10.1016/j.socl.2021.100014>
- SUNARP. (2020). Sunarp: número de autos que circulan en el país acumula una década de crecimiento continuo. <https://www.sunarp.gob.pe/PRENSA/inicio/post/2020/01/08/sunarp-numero-de-autos-que-circulan-en-el-pais-acumula-una-decada-de-crecimiento-continuo>

- Tabrizi, S. S., & Cavus, N. (2016). A Hybrid KNN-SVM Model for Iranian License Plate Recognition. *Procedia Computer Science*, 102, 588–594. <https://doi.org/10.1016/j.procs.2016.09.447>
- Varma, P. R. K., Ganta, S., Hari Krishna, B., & Svsrk, P. (2020). A Novel Method for Indian Vehicle Registration Number Plate Detection and Recognition using Image Processing Techniques. *Procedia Computer Science*, 167, 2623–2633. <https://doi.org/10.1016/j.procs.2020.03.324>

Análisis automático de micrografías SEM mediante aprendizaje profundo

Iker Toscano¹, Diego Alejandro Morales Bravo¹, Miguel De-la-Torre²,
Brenda Acevedo Juárez³, Gabriel Alberto García Mireles³

**iker.toscano2886@alumnos.udg.mx; diego.morales2887@alumnos.udg.mx;
miguel.dgomora@acadeicos.udg.mx; brenda.acevedo@academicos.udg.mx;
gabriel.garcia@unison.mx**

¹ Maestría en Ingeniería de Software, Centro Universitario de Los Valles, Ameca, México.

² Departamento de Ciencias Computacionales e Ingenierías, División de Estudios Científicos y Tecnológicos, Universidad de Guadalajara, México.

³ Departamento de Ciencias Naturales y Exactas, División de Estudios Científicos y Tecnológicos, Universidad de Guadalajara, México.

⁴ Departamento de Matemáticas, Universidad de Sonora, México.

DOI: 10.17013/risti.49.100–114

Resumen: El microscopio electrónico de barrido (SEM) es comúnmente utilizado para analizar nanopartículas de distintos materiales y mejorar métodos de fabricación, sistemas de purificación, mejoras en la industria médica, entre otras. En este artículo, se presenta un mapeo sistemático de la literatura referente al uso de técnicas de aprendizaje profundo (AP) para la detección y clasificación de nanopartículas contenidas en micrografías SEM. Los resultados reflejan que las redes neuronales convolucionales (CNN) son las técnicas más utilizadas para analizar micrografías, obteniendo una alta precisión en los proyectos realizados en las publicaciones revisadas. Como prueba de concepto, se presentan ejemplos del uso de los enfoques más comunes en micrografías SEM de CaCO_3 , incluyendo el uso de herramientas como OpenAI. Los resultados revelan ventajas y retos que se presentan al utilizar técnicas de aprendizaje profundo en el análisis de micrografías SEM.

Palabras-clave: aprendizaje profundo; micrografías SEM; CaCO_3 ; nanopartículas; redes neuronales convolucionales.

Automated analysis of SEM micrographs using deep learning

Abstract: The scanning electron microscope (SEM) is commonly used to analyze nanoparticles of different materials and improve manufacturing methods, purification systems, and improve the medical industry, among others. In this article, a systematic mapping of the literature regarding the use of deep learning (AP) techniques for the detection and classification of nanoparticles contained in SEM micrographs is presented. The results reflect those variants of convolutional neural networks (CNN) that are the most widely used techniques to analyze micrographs, obtaining high precision in the projects carried out in the reviewed publications.

As proof of concept, examples of the use of the most common approaches in SEM micrographs of CaCO_3 , including tools like OpenAI, are presented. The results reveal advantages and challenges that arise when using deep learning techniques in the analysis of SEM micrographs.

Keywords: deep learning; SEM micrographs; CaCO_3 ; nanoparticles; convolutional neural networks.

1. Introdução

Las micrografías, y en particular las obtenidas con el microscopio electrónico de barrido mejor conocido como SEM (scanning electron microscope) son cada vez más utilizadas para determinar de forma fiable la caracterización de la morfología como el tamaño, la distribución de tamaños y forma de los materiales, a niveles micro y nanométricos. Esta popularización del SEM se debe en parte a las innovaciones en el campo de la microscopía, las cuales han favorecido que el SEM esté cada vez más al alcance de la creciente comunidad científica, de un número mayor de universidades y una mayor cantidad de empresas dedicadas al desarrollo de materiales para el desarrollo de nuevos productos y procesos. Este crecimiento en el uso de micrografías SEM también está relacionado con las múltiples áreas de aplicación que surgen del diseño de nuevos materiales. Así, a manera de ejemplo, en Marcellos, Soares, Naiff, & Pires, 2020 se analizan micrografías SEM para detectar cristales en la vaterita (CaCO_3), y en Samuel M, Bose, & KC, 2009 se describe el proceso por el cual se analizan las nanopartículas de óxido de Zinc (ZnO_2), ambas tienen múltiples aplicaciones en diferentes áreas como cosméticos, textil, catalizadores, cerámicos, sensores de gas, aditivos de caucho, pigmentos, biomateriales, y otras. Además, las últimas tienen aplicaciones en películas finas para celdas solares, optoelectrónica, transmisores y transductores, como se evidencia en Löwy, 2011, Kaya, y otros, 2019, Rühle, Krumrey, & Hodoroaba, 2021, Sharma, Jaiswal, & Duffy, 2019, y Ijaz, Gilani, Nazir, & Bukhari, 2020.

Para el análisis de las micrografías, el proceso consiste en visualizar un conjunto de partículas (aprox. 250) correspondientes a una muestra estadística representativa, midiéndolas una a una para identificar la forma y construir un histograma de la distribución de tamaños y formas. La extracción de estos datos usualmente es realizada por un experto de forma manual, o con programas informáticos comerciales o de código abierto pocas veces especializados, y se lleva a cabo de forma semiautomática. Esto, lleva a largas jornadas de trabajo, además de que este tipo de software actúa como una caja negra y puede tener un impacto negativo en el resultado de la medición al generar errores (Dyba, Dingsoyr, & Hanssen, 2007). Esta situación genera motivación para trabajar en la mejora del software.

Algunas de las tareas más comunes en el análisis automático de micrografías SEM incluyen la detección de determinadas formas de micro y nanopartículas (como esferas o cubos, etc.); la segmentación de las partículas una vez detectadas; la clasificación en alguna serie de categorías que depende de la aplicación; y la medición de determinadas características como el área, diámetro, longitud, etc. Además, dada la necesidad de grandes cantidades de datos de entrenamiento al utilizar algunos algoritmos de aprendizaje automático, también se considera importante la tarea de aumentación de datos, que permite generar micrografías sintéticas con características similares a las reales.

Actualmente, diferentes métodos son utilizados para segmentar las nanopartículas. Sin embargo, estos métodos se encuentran en etapa de investigación y desarrollo aun existiendo retos que afrontar. En algunas ocasiones existe la opción de utilizar una dispersión de luz dinámica (Dynamic Light Scattering, DLS) que se utiliza para obtener información como el volumen y la densidad (Rühle, Krumrey, & Hodoroaba, 2021). Esta técnica de análisis DLS no siempre se práctica, y no es posible observar o contabilizar las nanopartículas como en el caso de las micrografías. Recientemente se han desarrollado sistemas para obtener información como tamaño y conteo a partir de micrografías SEM de diferentes materiales, con la ayuda de técnicas de aprendizaje profundo como se demuestra en el estudio (Nartova, y otros, 2022).

Dada la necesidad de conocer a mayor profundidad los procesos existentes sobre el desarrollo de software utilizando aprendizaje profundo para la caracterización de la morfología de los materiales en micrografías SEM, este artículo tiene como objetivo describir los métodos, técnicas y tendencias empleadas en la automatización del proceso de análisis de micrografías SEM mediante un mapeo sistemático (Systematic Mapping Study, SMS) así como algunas pruebas de concepto utilizando desarrollos existentes en micrografías de los materiales como CaCO_3 . El documento se estructura describiendo la metodología de las investigaciones relacionadas al análisis de microscopía SEM utilizando aprendizaje profundo, mediante un mapeo sistemático y la prueba de concepto de las técnicas más relevantes encontradas en la literatura, incluyendo la metodología estándar de evaluación (Sección 2). En seguida, se presentan los resultados y se discute sobre las técnicas más actuales para el análisis de micrografías SEM y aplicación al de las técnicas en las micrografías de CaCO_3 (Sección 3). Finalmente se exponen las conclusiones, y se especifican las líneas de trabajo por abordar en referencia a los hallazgos encontrados (Sección 4).

2. Metodología

2.1. Enfoques de aprendizaje profundo para el análisis de microscopías SEM

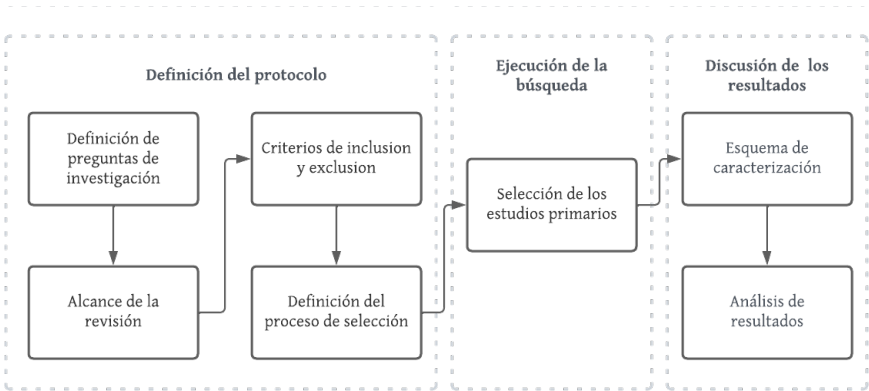


Figura 1 – Diagrama de flujo del proceso de SMS

Un SMS fue realizado con el fin de evaluar el estado actual de los métodos, técnicas y tendencia en el uso de aprendizaje profundo para el análisis de micrografías SEM, las cuales contienen micro y nanopartículas. El SMS permite obtener un panorama amplio del campo de investigación, puntos de atención y tendencias de trabajos realizados hasta el momento (Petersen, Feldt, Mujtaba, & Mattsson, 2008; Carrizo & Rojas, 2018; Dyba, Dingsoyr, & Hanssen, 2007). La metodología empleada en este trabajo se describe en (Carrizo & Rojas, 2018), este consiste en tres etapas de ejecución: definición de protocolo, ejecución de la búsqueda y discusión de resultados (ver Figura 1).

Definición del protocolo

La selección de artículos primarios pretende dar respuesta a las preguntas de investigación, para este caso en particular se consideraron las diferencias entre las técnicas y los métodos que consiguen clasificar y segmentar las nanopartículas contenidas en micrografías SEM, y la precisión de sus resultados; es así como se generaron las siguientes preguntas de investigación:

- ¿Cuál es el objetivo de los documentos que describen los métodos de segmentación automática de nanopartículas en micrografías SEM con ayuda de aprendizaje profundo?
- ¿Cuáles son las técnicas descritas en los documentos seleccionados para la segmentación de nanopartículas en micrografías SEM?
- ¿Cuál es la tendencia en el uso de aprendizaje profundo para la segmentación de nanopartículas en micrografías SEM?
- ¿Cuáles han sido los resultados del análisis desde el enfoque del uso de aprendizaje profundo en la segmentación de nanopartículas en micrografías SEM?

El alcance de SMS se refiere a la revisión de artículos primarios publicados en las bases de datos, y con los criterios de búsqueda que estuvieran escritos en idioma inglés, y durante el periodo 2017- 2022. En este estudio se emplearon las cuatro bases de datos más relevantes en el área de ciencias de la ingeniería y el desarrollo de software según Dyba, Dingsoyr, & Hanssen, 2007: Web of Science (WOS), Scopus, SpringerLink y ScienceDirect. Las palabras clave para conducir la búsqueda para la SMS fueron: “SEM”, “*scanning electron microscope*”, “*deep learning*”, “*CNN*”, “*segmentation*”, “*microparticles*” y “*nanoparticles*”, con las cuales se construyó la cadena de búsqueda general: (*SEM OR scanning electron microscope*) AND (*deep learning OR CNN*) AND *segmentation AND (microparticles OR nanoparticles)*. Las cadenas de búsqueda para las diferentes bases de datos se construyeron en base a la cadena de búsqueda general, y fueron adaptadas a los requerimientos de cada una de las bases de datos.

Considerando las preguntas de investigación fueron generados criterios de inclusión y exclusión que se muestran en la tabla 1, los cuales se emplearon en la ejecución de la búsqueda.

Criterios de inclusión	Criterios de exclusión
• Tener relación con el uso de técnicas de aprendizaje profundo para la segmentación de nanopartículas y evalúan las metodologías, técnicas y herramientas utilizadas. • Detallan el uso de técnicas de aprendizaje profundo para la detección de nanopartículas en micrografías SEM. • Planteen alternativas que hagan más eficiente la recolección de datos para el entrenamiento de redes neuronales. • Muestren otras metodologías sobre técnicas de aprendizaje profundo para la detección de nanopartículas en micrografías SEM.	• Artículos sin resultados que representen una mejora Artículos que no presenten evidencia o una metodología de desarrollo clara. • Se descartaron artículos que no contengan una justificación clara del desarrollo del proyecto o sistema.

Tabla 1 – Criterios de inclusión y exclusión para la selección de artículos primarios

Al definir el proceso de selección se generaron cuatro filtros posteriores la aplicación de los criterios de inclusión y exclusión, asegurando la calidad de la información recopilada:

1. **Primer filtro:** revisión del título.
2. **Segundo filtro:** revisión del resumen o *abstract*.
3. **Tercer filtro:** artículos sin resultados que representen una mejora mencionada en *abstract*.
4. **Cuarto filtro:** las publicaciones restantes fueron leídas en su totalidad y se realizó un análisis minucioso de su contenido.

Ejecución de la búsqueda

Una vez que se realizó la consulta en las bases de datos, se aplicaron los criterios de inclusión y exclusión, y los 4 filtros se seleccionaron los estudios primarios que son relevantes para responder a las preguntas de investigación.

Discusión de los resultados

Un esquema de caracterización fue creado para poder enfocar y simplificar los estudios primarios obtenidos. Los aspectos generales sobre las técnicas de aprendizaje profundo para la segmentación y detección de nanopartículas en las imágenes provenientes de microscopios SEM fueron analizados. Después se identificó el origen del conjunto de imágenes utilizadas para entrenar los algoritmos de aprendizaje profundo. Además del grado de validación, que incluye si la propuesta dentro del estudio fue validada o solo es una propuesta para utilizar técnicas de aprendizaje profundo. Finalmente, se considera el porcentaje de exactitud que presentan los artículos en sus resultados.

2.2. Prueba de concepto: segmentación de micro y nanopartículas en micrografías de carbonato de calcio

Tareas automatizadas mediante técnicas de aprendizaje profundo

Una de las dificultades con los algoritmos de aprendizaje automático es la cantidad de datos de entrenamiento. En la actualidad, no hay acceso a bases de datos suficientemente grandes como para entrenar modelos en nanotecnología, biología o medicina, por lo que se han desarrollado técnicas de aumentación de datos que aseguran que los algoritmos de aprendizaje máquina aprendan, reduciendo el sobreentrenamiento. En general, se toma una imagen base y se aplican transformaciones como giros, recortes, acercamientos y adición de ruido blanco, entre otras. Entre las técnicas de aumentación de datos que se prueban en este trabajo, se incluyen las encontradas en los trabajos primarios (Nartova, y otros, 2022), así como las disponibles en la plataforma de OpenAI.

Por otra parte, las redes neuronales utilizadas para la detección de objetos utilizan una red troncal pre-entrenada para extraer características de la imagen, y luego utilizan una red de detección para realizar la predicción de la clase y la ubicación del objeto. En cuanto a la segmentación de objetos, se entrena con una representación de múltiples objetos, ya sea superpuestos o solapados. Este tipo de escenarios se encuentra en diversas aplicaciones biológicas, como el análisis morfológico de objetos moleculares o celulares, o la composición de nanoestructuras para observar objetos individuales teniendo en cuenta su tamaño y forma. Sin embargo, una limitación de estas aplicaciones es que se necesita conocer la distribución de tamaños de las nanoestructuras.

Una alternativa para la detección de nanopartículas se presenta en Monchot, y otros, 2021. En este trabajo se adapta una Mask R-CNN para la detección de nanopartículas de ZnO_2 en micrografías SEM. Esta fue creada por facebook Mask R-CNN es un modelo que mejora el detector de objetos Faster R-CNN al agregar una tercera rama que produce máscaras de objeto. Esto significa que no solo se identifica la clase del objeto y su posición, sino también su forma exacta. Por lo tanto, Mask R-CNN requiere una disposición espacial mucho más fina de los objetos en las imágenes para producir las máscaras con precisión.

Un método para segmentar objetos con superposición parcial es el denominado ACED. Consiste en binarizar la imagen para obtener la silueta, eliminando el fondo. En el primer paso se extrae el contorno de los límites que pueden inferir los objetos superpuestos. Como una opción se utiliza el detector de bordes Canny para extraer los bordes.

Por otro lado, un ejemplo de sistema para identificar, medir y contar nanopartículas es el sistema LIST, que ejecuta tres etapas: (1) reconocimiento de la morfología, (2) medición del tamaño y (3) reconocimiento de la barra de escala y texto para la conversión de la escala. En la primera etapa se determina la morfología de las nanopartículas en la imagen SEM de entrada. A continuación, se realiza la medición del tamaño en función del tipo de morfología de entrada (Kim, Han, & Han, 2020). Finalmente, se realiza el reconocimiento de la barra de escala y texto para extraer la información de la escala y convertir el tamaño estimado en píxeles en la unidad adecuada (por ejemplo, μm o nm).

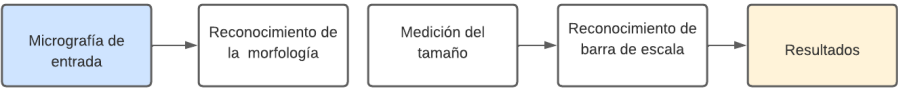


Figura 2 – Flujo de procesamiento de imágenes del software LIST

3. Resultados y discusión

3.1. Mapeo sistemático: técnicas para el análisis de micrografías SEM

Los resultados del proceso de SMS se presentan en la tabla 2. De la búsqueda en la base de datos se recuperaron 161 artículos, aplicando los criterios de búsqueda, de inclusión y de exclusión y posterior a aplicar los cuatro filtros, los artículos primarios que seleccionados se restringió a 16.

Base de datos	Filtro				
	Resultados	Primero	Segundo	Tercer	Cuarto
Scopus	37	26	25	10	7
WOS	6	4	1	1	1
Springer link	68	5	5	5	5
Science Direct	50	9	6	3	3
Total	161	44	37	18	16

Tabla 2 – Resultados de la búsqueda después de aplicar los filtros

El resultado del análisis de los 16 artículos que se encontraron, desde la cantidad de estudios de la información sobre la temática, y por qué es de interés abordar esta problemática tan actual se obtuvo la Tabla 3.

Aspectos generales	Cantidad de trabajos
Utilizan redes neuronales convolucionales para la segmentación	15
Utilizan otras técnicas para la segmentación de nanopartículas	1
Entradas para el entrenamiento	
Utilizan redes GAN para generar datos de entrenamiento	4
Modelan nanopartículas en 3D	2
Analizan micrografías no sintéticas	10
Grado de validación	
Muestran resultados validados	16
Resultados	
Presentan resultados con una precisión mayor a 90%	5

Aspectos generales	Cantidad de trabajos
Presentan resultados con una precisión entre 80% y 90%	3
Presentan resultados con una precisión entre 70% y 80%	1
No presentan de precisión obtenida	7

Tabla 3 – Clasificación de los artículos primarios para su análisis.

Adicionalmente se realizó la búsqueda en las bases de datos recomendadas en el área: EBSCO, IEEE y ACM, sin encontrar resultados relevantes, se encontró una fuente relevante que se agregó al análisis: Rühle, Krumrey, & Hodoroaba, 2021.

Como resultados en términos de precisión, en los estudios evaluados tenemos que tres artículos presentan una precisión por encima de un 90%: Kaya, y otros, 2019; Cabrera, y otros, 2022; Furat, y otros, 2022). Salvo algunas excepciones tenemos resultados por debajo del 80%: Dyba, Dingsoyr, & Hanssen, 2007; Riewald, Kurzhals, Bianchini, & Heino, 2022; Kavuran, 2021. Sin embargo, es importante mencionar que se encontraron importantes sugerencias, tal como el uso de CNN más robustas o datasets con una complejidad menor en las nanopartículas mostradas en cuanto a su morfología.

El análisis de los artículos primarios permitió encontrar las observaciones eficaces y métodos que presentan una alternativa automatizada en la segmentación y clasificación de nanopartículas. La mayoría de los documentos reportaron como primordial el segmentar y analizar las nanopartículas, así como caracterizar por tamaño y forma para posteriores análisis en diferentes industrias lo cual responde a la pregunta de investigación (Soosen Samuel, Bose, & George, 2009). Las técnicas que son descritas para la detección y conteo de nanopartículas corresponden al uso del aprendizaje profundo en arquitecturas CNN, las cuales mencionan ser las más eficaces, lo cual responde a la pregunta de investigación número (Löwy, 2011).

Las tendencias que se observaron para realizar estos estudios es el uso de CNN para la segmentación y detección de nanopartículas y el uso de redes GAN para clasificar por morfologías las nanopartículas, respondiendo la pregunta de investigación (Kaya, y otros, 2019). Se puede deducir que el enfoque de la inteligencia artificial para la segmentación de micro y nanopartículas en micrografías SEM es el uso de aprendizaje profundo utilizando CNN. Para esto tenemos arquitecturas como Marsk-Net, U-Net, DenseNet, que facilitan el usar estas técnicas para desarrollar metodologías que permiten tener resultados precisos. Para la implementación comúnmente se utilizan librerías como Keras y Tensor Flow, que ponen a disposición éstas y otras herramientas de aprendizaje máquina.

3.2. Análisis de los estudios primarios

Entre los artículos primarios podemos encontrar una fuerte tendencia sobre el objetivo de encontrar una alternativa al análisis de micro y nanopartículas obtenidas de microscopios empleando tecnología de aprendizaje profundo para la detección y segmentación de nanopartículas. Sobre las técnicas de segmentación se encontró que se enfocan en su mayoría en el uso de redes neuronales convolucionales (CNN). Siendo

esta técnica la preferida debido a las capas que involucra para procesar información y características no lineales. Un diseño típico de una CNN aprende patrones inherentes a través de filtros aplicables en diferentes resoluciones y las compara con las imágenes de entrada. También se encontró que librerías como Keras permiten visualizar patrones de manera sencilla (Baskaran, y otros, 2021; Khairunissa & Lee, 2022; Riewald, Kurzhals, Bianchini, & Heino, 2022; Kavuran, 2021). Sin embargo, si las muestras a analizar no tienden a tener una mayor complejidad, como nanopartículas superpuestas unas a otras, desenfoque, ó un número muy grande de antipartículas dentro de la micrografía, se puede hacer uso de alternativas como *Google inception network*. Tal es el caso, del estudio de Geshe, Chaisoontornyotin, Elkhatib, & Goual, 2020, en donde se realizó el reconocimiento de nanopartículas y su tamaño.

Como entradas para el entrenamiento de las técnicas de redes neuronales en los artículos primarios, se encontró que un *dataset* de imágenes es necesario, y para generarlo se encontró una variedad de alternativas empleadas. Una de las técnicas a destacar es el uso de redes generativas antagónicas Generative Adversarial Networks (GAN). Estas redes hacen posible el generar nuevas micrografías a partir de unas cuantas muestras como en los siguientes artículos: Kaya, y otros, 2019; Fu, y otros, 2021; Kim, Han, & Han, 2020; Coquelin, Fischer, Feltin, Devoille, & Felhi, 2019. Sin embargo, también es común el uso de micrografías tomadas de manera tradicional de departamentos especializados en universidades (Kavuran, 2021).

3.3. Pruebas de concepto

Las micrografías que se emplearon para la prueba de concepto fueron generadas en la síntesis de micro y nanopartículas de CaCO_3 en el trabajo de Hernández-Zepeda, 2022. La importancia de la obtención de estas micro y nanopartículas es porque pueden ser empleadas para la remoción de metales pesados, portador de fármacos o aditivo industrial, depende de diferentes variables como: su morfología, área superficial específica, fase cristalina y pureza. Para lograr un producto de síntesis adecuado dentro de una síntesis química, se deben controlar los disolventes y las temperaturas. La variación de estos parámetros ha demostrado tener efectos significativos en la forma, el tamaño, la uniformidad, la estabilidad y otras propiedades de las partículas (Khan, Saeed, & Khan, 2019).

En Hernández-Zepeda para la obtención de nanopartículas de CaCO_3 se emplearon conchas de *Crassostrea corteziensis*, que son residuos que se generan del cultivo de ostiones en Boca de Camichin, Nayarit. Las conchas de esta especie están formadas de calcita que se utilizó como precursor para obtener micro y nanopartículas de CaCO_3 . Mediante el método de sonoquímico y como disolventes se empleó ácido cítrico y bromuro de cetiltrimetilamonio. Una serie de 12 experimentos fueron realizados, obteniendo partículas de tamaños entre 70 nm a 250 nm considerando de diferente morfología consideradas micro y nano partículas y se obtuvieron 33 micrografías.

Aumentación de datos

En la Figura 3 se muestran ejemplos de aumentación de datos a partir de una micrografía de CaCO_3 , las operaciones aplicadas fueron técnicas de procesamiento de imágenes utilizando la librería Pillow de Python. En una sola imagen se pueden llevar varias

combinaciones de estas transformaciones. Es importante señalar que se debe de limitar el grado del giro para no generar datos que se alejen demasiado de una imagen real y puedan afectar negativamente el entrenamiento del modelo. Técnicas que se analizarán más adelante presentan en alguno de sus pasos para la aumentación de datos cambios sobre la imagen como los que se mostraron.

Los autores Coquelin, Fischer, Feltin, Devoille, & Felhi, 2019 introducen el uso de redes generativas adversariales. Las redes generativas adversariales consisten en dos redes neuronales con arquitecturas independientes que son entrenadas al mismo tiempo. El modelo GAN (generative adversarial networks) consiste una red generadora y una discriminadora. La red discriminadora tiene como tarea principal la clasificación de una imagen como reales o falsas. La red generadora busca poder crear una imagen lo suficientemente parecida a las que se encuentran en la base de datos con la que se entrenan para poder engañar a la discriminadora. El modelo entrena ambas redes y las enfrenta en una dura competición por mejorarse a sí mismas. Al final, el discriminador será capaz de identificar la mínima diferencia entre lo real y lo generado, y la red generativa será capaz de crear imágenes que el discriminador no pueda distinguir.

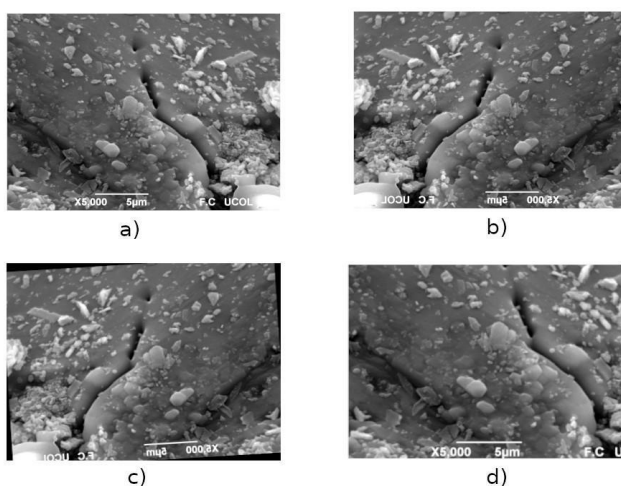


Figura 3 – a) Micrografía de CaCO_3 obtenida por microscopio SEM. b) Micrografía con transformación horizontal. c) Micrografía con giro de 1 grado. d) Micrografía con recorte.

Para mostrar la viabilidad del uso de redes GAN como técnica de aumentación de datos se hizo una implementación de una red GAN utilizando como base de datos para el entrenamiento 200 imágenes de microscopias SEM de CaCO_3 generadas en la plataforma OpenAI. Una muestra de las imágenes resultantes se puede observar en la Figura 4.

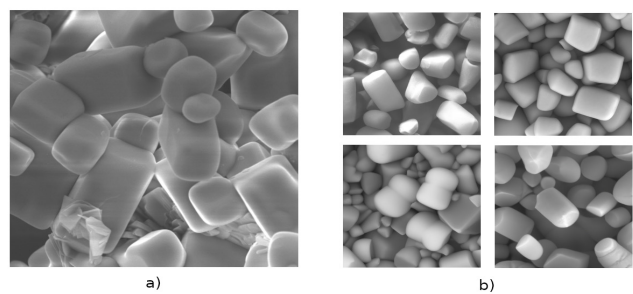


Figura 4 – a) Micrografía de base para generación de variantes.
b) Variaciones de la imagen original

La arquitectura de la red discriminadora tiene de entrada una imagen de dimensiones 256 x 256 x 1 y de salida una sola neurona que indica si es real o no la imagen. El segundo modelo discriminador funciona con una capa de entrada de 128 neuronas y genera de salida una imagen del mismo tamaño que la entrada de la red discriminadora. Múltiples entrenamientos fueron realizados variando los valores del ritmo de aprendizaje de ambas redes, alguno de los resultados de estos entrenamientos se muestra en la siguiente Figura 5.

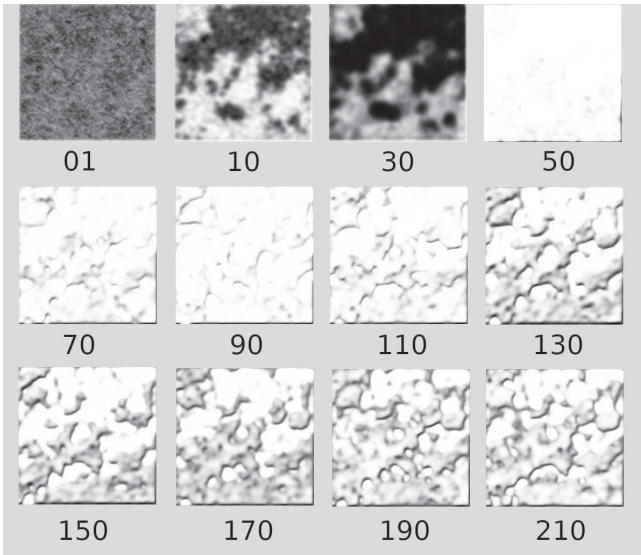


Figura 5 – Se muestran las imágenes de salida de la red generadora. El número de épocas de entrenamiento es el que se muestra debajo de cada una de las imágenes.

Aunque múltiples pruebas de entrenamiento se realizaron, no se pudo conseguir un mejor resultado que el que se muestra en la Figura 5. La dificultad de implementación de redes neuronales GAN cae en el hecho de que el entrenamiento de ambas redes sucede al mismo tiempo. El caso más común que ocurrió en esta implementación fue el problema de colapso de la medición de error por parte de la red generadora. Este colapso de error hace que la red generadora deje de aprender, ya que encuentra que las imágenes que esté generando ya son suficientes para burlar al discriminador proporcionando siempre la misma imagen sin importar el tiempo de entrenamiento. Por este motivo de dificultad, hoy en día, es por el cual la implementación de redes neuronales GAN son desplazadas por otras técnicas de aumentación de datos. Como ejemplo de esto, es el uso de plataformas de OpenIA que ya cuentan con generadoras de imágenes, por lo que es más fácil el generar bases de datos para entrenamiento.

Análisis de partículas

El proceso de canalización en la medición de nanopartículas implica la clasificación de la forma de las partículas y la obtención de una imagen binaria correcta para medir con precisión el tamaño. Se utilizan algoritmos de segmentación y procesamiento de imágenes para lograr este proceso en específico Inception-V3 entrenada con ImageNet y DenseNet. Se aplica una equalización de histograma y una umbralización adaptativa de Otsu para obtener la imagen binaria correcta. Luego, se extraen los centros de núcleo y se examina la existencia de regiones de cáscara en la imagen binaria. La Figura 6 muestra algunas micrografías de carbonato de calcio procesadas con el software LIST. En ellas se pueden observar algunas partículas detectadas, y el resultado de binarizar la imagen.

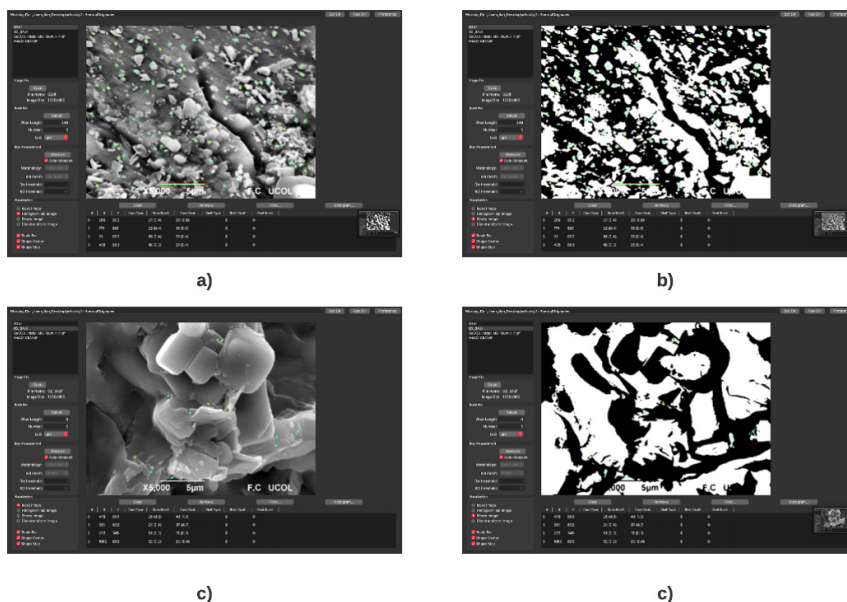


Figura 6 – Micrografías de CaCO_3 después de ser procesadas por LIST

5. Conclusiones

En este artículo se presentan los resultados de una revisión de literatura respecto de la clasificación de nanopartículas en micrografías SEM utilizando técnicas de aprendizaje profundo. Además, se exponen los principales hallazgos de una prueba de concepto referente al uso de dichas técnicas en micrografías de CaCO_3 .

La primera parte del artículo presenta la metodología utilizada para realizar la revisión literatura a través de un mapeo sistemático, siguiendo las recomendaciones de Petersen et al. (2008). La búsqueda de artículos se enfocó en la búsqueda automática en cuatro bases de datos (Web of Science, ScienceDirect, SpringerLink, y Scopus) utilizando como términos clave “scanning electron microscope”, “nanoparticles” y “deep learning”.

Como resultado del mapeo sistemático, se seleccionaron 16 estudios primarios, los cuales describen técnicas de aprendizaje basadas en redes neuronales profundas para realizar tareas como detectar, segmentar y clasificar nanopartículas contenidas en las micrografías. La clasificación de los estudios primarios mostró que las técnicas de segmentación usan con mayor frecuencia redes neuronales convolucionales. Además, para la generación de imágenes de entrenamiento, la mayoría de los estudios indican que las imágenes usadas fueron generadas por la misma universidad o instituto. En el caso de que dichas micrografías no puedan tomarse, existe la posibilidad de usar las redes neuronales generativas adversarias (GAN) para generar imágenes sintéticas.

En la segunda parte de este trabajo, se realizó un estudio exploratorio con las redes GAN con la finalidad de comprobar su viabilidad como técnica de aumentación de datos. El entrenamiento de la red se realizó con 200 imágenes de microscopía de carbonato de calcio. Sin embargo, los resultados de este estudio exploratorio no fueron satisfactorios debido al problema del colapso de la medición de error, lo que provocó que la red dejó de aprender.

Con base en los resultados del mapeo sistemático, no se puede pasar por alto la relevancia de las redes GAN para la aumentación de datos, ya que permite generar suficientes micrografías para entrenamiento. Similarmente, el software LIST permite el análisis de micrografías SEM, incluyendo la medición y el conteo, aunque presenta algunas limitantes para detectar estructuras complejas. Es en este punto donde se puede visualizar un área de oportunidad para la investigación y desarrollo de algoritmos que mejoren la tecnología actual.

Referencias

- Baskaran, A., Kautz, E., Chowdhary, A., Ma, W., Yener, B., & Lewis, D. (2021). Adoption of Image-Driven Machine Learning for Microstructure Characterization. *JOM*, 73(11), 3639-3657.
- Cabrera, C., Cervantes, D., Muñoz, F., Hirata, G., Juárez, P., & Flores, D.-L. (2022). Classifying nanostructured and heterogeneous materials from transmission electron. *Neural Computing and Applications*, 1-13.
- Carrizo, D., & Rojas, J. (2018). Metodologías, técnicas y herramientas en ingeniería de requisitos: un (Vol. 26). SciELO Agencia Nacional de Investigación y Desarrollo (ANID).

- Coquelin, L., Fischer, N., Feltin, N., Devoille, L., & Felhi, G. (2019). Towards the use of deep generative models for the characterization in size of. *Materials Research Express*, 6(8), 85001.
- Dyba, T., Dingsoyr, T., & Hanssen, G. (2007). Applying Systematic Reviews to Diverse Study Types: An Experience Report. *First International Symposium on Empirical Software Engineering*.
- Fu, Q., Zhang, Y., Wang, P., Pi, J., Qiu, X., Guo, Z., Xu, J. (2021). Rapid identification of the resistance of urinary tract pathogenic bacteria using (Vol. 413). Springer Science and Business Media LLC.
- Furat, O., Finegan, D., Yang, Z., Kirstein, T., Smith, K., & Schmidt, V. (2022). Super-resolving microscopy images of Li-ion electrodes for fine-feature (Vol. 8). Springer Science and Business Media LLC.
- Gesho, M., Chaisoontornyotin, W., Elkhatib, O., & Goual, L. (2020). Auto-segmentation technique for SEM images using machine learning: *Ultramicroscopy*, 217, 113074.
- Hernández-Zepeda, C. (2022). Síntesis de micro y nanopartículas de CaCO_3 a partir de conchas de ostión de *Crassostrea corteziensis* por el método sonoquímico. Ameca, Jalisco: Universidad de Guadalajara.
- Ijaz, I., Gilani, E., Nazir, A., & Bukhari, A. -G. (2020). Detail review on chemical, physical and green synthesis, classification, characterizations and applications of nanoparticles (Vol. 13). Taylor and Francis Ltd. <https://www.tandfonline.com/doi/abs/10.1080/17518253.2020.1802517>
- Kavuran, G. (2021). SEM-Net: Deep features selections with Binary Particle Swarm Optimization (Vol. 27). Elsevier BV.
- Kaya, E., Kaya, O., Alkan, G., Gürmen, S., Stopic, S., & Friedrich, B. (2019). New Proposal for Size and Size-Distribution Evaluation of Nanoparticles (Vol. 13). MDPI AG.
- Khairunissa, M., & Lee, H. (2022). Effective quantum mechanicseembedded nanoparticle occlusion (Vol. 24). Springer Science and Business Media LLC.
- Khan, I., Saeed, K., & Khan, I. (2019). Nanoparticles: Properties, applications and toxicities (Vol. 12). Elsevier.
- Kim, H., Han, J., & Han, T.-J. (2020). Machine vision-driven automatic recognition of particle size and morphology in (Vol. 12). Royal Society of Chemistry (RSC).
- Löwy, I. (2011). Historiography of biomedicine: “bio,” “medicine,” and in between (Vol. 102). The University of Chicago Press.
- Marcellos, C., Soares, E., Naiff, D., & Pires, G. (2020). Analyze of scanning electron microscope images using deep learning for calcium carbonate polymorph detection. <https://www.nidf.ufrrj.br/wp-content/uploads/2021/03/Analyze-of-scanning-.pdf>
- Monchot, P., Coquelin, L., Guerroudj, K., Feltin, N., Delvallée, A., Crouzier, L., & Fischer, N. (2021). Deep Learning Based Instance Segmentation of Titanium Dioxide Particles in the. *Nanomaterials*, 11(4), 968.

- Nartova, A. V., Mashukov, M. Y., Astakhov, R. R., Kudinov, V. Y., Matveev, A. V., & Okunev, A. G. (2022). Particle Recognition on Transmission Electron Microscopy Images Using Computer Vision and Deep Learning for Catalytic Applications. *Catalysts*, 1-13.
- Petersen, K., Feldt, R., Mujtaba, S., & Mattsson, M. (2008). Systematic Mapping Studies in Software Engineering (Vol. 17).
- Riewald, F., Kurzhals, P., Bianchini, M., & Heino. (2022). The LiNiO₂/sub Cathode Active. *Journal of The Electrochemical Society*, 169(2), 20529. <https://doi.org/10.1149/1945-7111/ac4bf3>
- Rühle, B., Krumrey, J., & Hodoroba, V.-D. (2021). Workflow towards automated segmentation of agglomerated, non-spherical. *Scientific Reports*, 11(1).
- Samuel M, S., Bose, L., & KC, G. (2009). Optical Properties of ZnO₂ Nanoparticles. *Academic Review*, 57-65.
- Sharma, S., Jaiswal, S., & Duffy, B. (2019). Nanostructured materials for food applications: spectroscopy, microscopy and physical properties. *Bioengineering*. <https://www.mdpi.com/430090>
- Soosen Samuel, M., Bose, L., & George, K. (2009). Optical properties of ZnO nanoparticles (Vol. 16).

Características de las Bases de Datos para el Cómputo en la Nube: Un Estudio Secundario

Raúl A. Aguilar Vera¹, Naomi García Sánchez¹, Julio C. Díaz Mendoza¹,
Brenda L. Flores Ríos²

avera@correo.uady.mx; A17016302@alumnos.uady.mx; julio.diaz@correo.uady.mx;
brenda.flores@uabc.edu.mx

¹ Facultad de Matemáticas, Universidad Autónoma de Yucatán, C.P. 97000, Mérida, Yucatán, México.

² Instituto de Ingeniería, Universidad Autónoma de Baja California, Mexicali, Baja California, México.

DOI: 10.17013/risti.49.115–130

Resumen: El crecimiento exponencial de las Bases de Datos a principios del presente siglo generó la necesidad de mecanismos innovadores para almacenamiento y gestión de información. Uno de estos paradigmas emergentes es el cómputo en la nube, un modelo que brinda acceso a un conjunto de recursos informáticos a través de la red de una manera sencilla. El presente estudio tiene como objetivo elaborar un estado del arte sobre aspectos vinculados con los esquemas de persistencia utilizados para el cómputo en la nube. Para ello se especificaron 5 preguntas de investigación que sirvieron para guiar una Revisión Sistemática de Literatura con la que se analizaron 72 Estudios Primarios. Entre los hallazgos se encontraron ocho técnicas específicas para el modelado de Bases de Datos, siendo la más citada el cifrado de datos; así mismo, se pudo identificar que Amazon es el proveedor corporativo más popular para dicho tipo de sistemas de gestión.

Palabras-clave: Cómputo en la Nube; Estudio Secundario; Modelado de Bases de Datos.

Characteristics of Databases for Cloud Computing: A Secondary Study

Abstract: The exponential growth of Databases at the beginning of this century generated the need for innovative mechanisms for information storage and management. One of these emerging paradigms is cloud computing, a model that provides access to a group of computing resources through the network in a simple way. The aim of this study is to elaborate a state of the art on aspects related to Databases used for cloud computing. For this, 5 research questions were specified that served to guide a systematic review of the literature with which 72 Primary Studies were analyzed; among the findings, eight specific techniques for Database modeling were found, the most cited being data encryption; likewise, it was possible to identify that Amazon is the most popular corporate provider for this type of management systems.

Keywords: Cloud Computing; Database Modeling; Secondary Study.

1. Introducción

A poco más de medio siglo de su concepción como disciplina ingenieril, la Ingeniería de Software (IS) ha acumulado un cuerpo de conocimientos relacionado con los procesos de desarrollo y gestión del software (Bourque & Fairley, 2014); sin embargo, el desarrollo acelerado y la evolución de los sistemas de información, ha dado pie en el nuevo milenio, a un conjunto de paradigmas emergentes (Aguilar, Oktaba y Juárez, 2019). Entre estos paradigmas, uno de singular importancia para la IS, es un modelo innovador conocido como Cómputo en la Nube (CN), el cual permite habilitar el acceso ubicuo bajo demanda a un grupo compartido de recursos informáticos configurables (Mell & Grance, 2011). La comprensión del constructo CN resulta importante para poder describir de forma apropiada la gestión de la información al interior de dicho ambiente; su éxito radica, en parte, por la disminución en tamaño para aquellas empresas que lo han incorporado, lo cual resulta un factor crítico en el beneficio económico de las mismas. El modelo del CN ha dado paso a la evolución de centros de datos que brindan a los clientes la infraestructura física requerida para alojar sus sistemas de información, incluidas las fuentes de alimentación redundantes, capacidades de comunicación de gran ancho de banda, monitoreo del entorno, así como servicios de seguridad (Zhao et al., 2014) permitiendo de esta manera la administración de sus procesos desde la nube (Peralta et al., 2016).

El presente estudio es una Revisión Sistemática de Literatura (RSL) con la cual se pretende sintetizar la investigación disponible sobre el paradigma del CN, en particular, sobre aspectos vinculados con las técnicas de modelado, tipos de datos, aspectos de calidad; así como proveedores y herramientas más comunes y/o populares para dichos sistemas.

El documento se encuentra organizado de la siguiente manera: la sección 2 presenta antecedentes y trabajos previos relacionados con el CN; la sección 3 describe la metodología utilizada en el presente estudio secundario. Las secciones 4 y 5 describen la planificación y ejecución del estudio, de acuerdo con la metodología descrita. En la sección 6 se describen los resultados del estudio de acuerdo con las Preguntas de Investigación planteadas. Finalmente, las conclusiones y trabajos futuros se presentan en la sección 7.

2. Antecedentes

Al inicio del presente siglo, los sistemas de información comenzaron a gestionar grandes cantidades de información para las aplicaciones de escritorio, con la expansión de los sistemas al incorporarse al internet, así como con la aparición de una gran diversidad de tipos de datos que empezaron a ser utilizados, la necesidad de mayor capacidad de persistencia en las Bases de Datos (BD) se incrementó de manera exponencial, por lo que el almacenamiento de la información pasó de medios locales, a medios gestionados mediante el CN. Para 2012 en internet ya existían 2.27 miles de millones de usuarios que requerían almacenamiento en algún dispositivo, y para el año 2020, de acuerdo con el informe de *Data Never Sleeps*, se generaban al día más de 2,500 millones de bytes, en particular, en plataformas como *Netflix*, *YouTube*, *Facebook* y otras aplicaciones de uso cotidiano (DOMO, 2021) en las que la variedad de tipos de datos ya no resultaba

compatible con los tradicionales esquemas de persistencia vinculados al modelo relacional.

El CN generalmente se rige por cinco características fundamentales: (1) ser un servicio automático por demanda, (2) ofrecer amplio acceso a la red, (3) disponer de un conjunto de recursos en común, (4) brindar elasticidad rápida, y (5) ser un servicio con medida. Con las características antes citadas, podemos decir que la era del CN inicia formalmente en 2006 con la aparición del CN elástico, un servicio de almacenamiento simple y novedoso introducido por la Web de Amazon (AWS). No obstante, independientemente de cuál haya sido la primera empresa en dar pauta al desarrollo de dicha tecnología, la principal razón por la que se destacó, es debido a que tanto el procesamiento de la información, como su almacenamiento, puede ser más eficiente si se cuenta con un ambiente más grande y accesible como lo es internet (Marinescu, 2018).

En el caso de la comunidad de Ciencia y la Tecnología, el paradigma del CN ha venido demostrando mayor interés de investigación; de hecho, en el contexto de la IS ha dado como resultado sistemas y aplicaciones orientadas a servicios que se ofrecen a través de internet, con un conjunto limitado de aspectos de calidad como son: la escalabilidad, portabilidad, así como la capacidad de integración con otras aplicaciones o servicios. Así pues, el CN puede ser descrito como un modelo que brinda acceso de red a un conjunto de recursos informáticos configurables, los cuales pueden ser accedidos por un usuario con esfuerzos mínimos y sin necesidad de terceros (Jain, Deshpande & Khanaa, 2017).

Diversos estudios secundarios (Bhatti & Rad, 2017; Karataş et al., 2017; Mostajabi, Safaei & Sahafi, 2021) han documentado que una de las problemáticas del CN gira en torno al desarrollo de las BD utilizadas por dicho modelo, en parte, debido al avance y crecimiento exponencial de los servicios que ofrecen, lo cual obliga a que sean diversos los elementos y aspectos de calidad que deben ser verificados.

3. Metodología de Investigación

El presente estudio utilizó como referencia la guía para realizar RSL propuesta por Kitchenham & Charters (2007). Las principales actividades de que constan las tres fases consideradas por la metodología son las siguientes:

1. *Planificar la RSL:*

- Identificar la necesidad de la revisión.
- Formular las preguntas de investigación.
- Definir la estrategia de búsqueda.
- Establecer los criterios para selección de los estudios primarios.
- Definir la estrategia de evaluación de la calidad de los estudios seleccionados.

2. *Realizar la RSL*

- Seleccionar los estudios primarios relevantes.
- Evaluar la calidad de los estudios primarios seleccionados.
- Extraer la información de los estudios primarios seleccionados.
- Sintetizar los hallazgos.

3. Reportar RSL

- Elaborar un informe técnico de la revisión.
- Evaluar el informe de la revisión.

Para incrementar el número los estudios relevantes, se incorporó la estrategia denominada bola de nieve hacia adelante (Wohlin, 2014) la cual sería aplicada al primer conjunto de Estudios Primarios Seleccionados (EPS).

4. Planificación del Estudio

4.1. Justificación de la necesidad del estudio

La evolución de manera acelerada de las Tecnologías de la Información y Comunicaciones dio origen, en el nuevo milenio, a un novedoso conjunto de paradigmas en la IS; en particular, el CN, el cual brinda acceso ubicuo bajo demanda a un conjunto compartido de recursos informáticos configurables, con los cuales es posible gestionar grandes volúmenes de información, de tipos de datos diversos. Su desarrollo en las últimas dos décadas ha sido gracias a numerosos estudios empíricos que se encuentran dispersos en diversas fuentes de información, para poder identificar nichos de oportunidad y continuar con el avance de la investigación en este innovador paradigma, es necesario realizar una selección, análisis e integración del conocimiento hasta ahora generado, por lo cual, se propone el presente estudio secundario, particularmente, mediante una RSL.

4.2. Formulación de las Preguntas de Investigación

Con base en los antecedentes del novedoso paradigma del CN, así como del análisis de algunos estudios secundarios previos (Bhatti & Rad, 2017; Karataş et al., 2017; Mostajabi, Safaei & Sahafi, 2021) los autores diseñaron un grupo de Preguntas de Investigación (PI), las cuales sirvieron para guiar la presente RSL:

PI-1 ¿Cuáles son las técnicas, procedimientos, o prácticas más utilizadas en el modelado de BD para el CN?

PI-2 ¿Cuáles son los principales modelos de datos que se usan en las BD para el CN?

PI-3 ¿Cuáles son los principales aspectos de calidad considerados por las arquitecturas y/o modelos propuestos para las BD utilizados para el CN?

PI-4 ¿Cuáles son los principales proveedores de herramientas para el CN?

PI-5 ¿Cuáles son las principales herramientas vinculadas con la gestión de BD para el CN?

4.3. Definición de la Estrategia de Búsqueda

Se realizó una búsqueda no sistemática de información sobre la temática del CN con el fin de identificar, tanto los conceptos clave vinculados con las PI, así como las fuentes donde suelen publicarse los estudios primarios vinculados con el tema.

Con base en la exploración del tema, se seleccionaron como fuentes de información: *IEEE Xplore*, *ACM Digital Library*, *ScienceDirect* y *SpringerLink*; así mismo, se decidió

agregar un repositorio —*Google Scholar*— que pudiera contener estudios relevantes de otras fuentes no consideradas. Así mismo, con base en las PI acordadas, se seleccionó un conjunto limitado de palabras clave que pudieran servir para identificar estudios primarios pertinentes con la temática del estudio: *cloud database, design, modeling, architecture*. Los conceptos considerados, junto con un conjunto de operadores lógicos, sirvieron para integrar una cadena de búsqueda genérica la cual será configurada en la fase de ejecución, en función de las características particulares de cada una de las 5 fuentes seleccionadas:

“Cloud database AND (modeling OR design)”

4.4. Criterios de Inclusión y Exclusión

Para discernir entre los estudios vinculados con la temática y los que realmente proporcionan información relevante, de acuerdo con las PI, se establecieron un conjunto de criterios de inclusión y de exclusión. Los criterios definidos para la selección de los estudios primarios son:

Criterios de Inclusión (CI):

- Artículos publicados en revistas y memorias de congresos especializados.
- Estudios empíricos relacionados con el diseño de BD para el CN.
- Publicaciones en idioma inglés.
- Estudios en la línea de tiempo de 2010 a 2021.

Criterios de Exclusión (CE):

- Estudios duplicados (manteniendo los de la fuente original).

4.5. Técnica de Bola de Nieve hacia adelante

Para incrementar y mejorar el conjunto de estudios pertinentes para la RSL, se decidió por el grupo de investigadores, utilizar la técnica denominada bola de nieve hacia adelante (Wohlin, 2014). Para la selección de un segundo conjunto de estudios con base en la aplicación de la técnica antes citada, se seleccionó *Google Scholar*, el cual es un motor de búsqueda de contenido y bibliografía científico-académica provista por *Google*, el cual, aparte de ser un repositorio, permite identificar, con cierta vigencia, los artículos que han citado a una publicación.

4.6. Estrategia de Evaluación de la Calidad de los EPS

El proceso de evaluación de la calidad en una RSL sirve para garantizar la validez y confiabilidad de los hallazgos descritos en los EPS, ponderando la importancia de cada uno de los n estudios (Kitchenham, Budgen & Brereton, 2015). Sin embargo, debido a la diversidad de metodologías utilizadas en los estudios primarios, no es posible disponer de un instrumento estandarizado, por lo que, en el caso del estudio, se procedió a adaptar —en un total de 10 ítems— el instrumento propuesto por Dyba & Dingsøyr (2008). Dicho instrumento considera los criterios: (1) Rigor, (2) Calidad del informe, (3) Relevancia y (4) Credibilidad. En función del grado de cumplimiento de cada uno de los criterios antes citados, el investigador asignará un valor cuantitativo —en una escala ordinal— con los siguientes valores:

- 1.0 – Para indicar que el artículo cumple con el mecanismo en su totalidad.
- 0.5 - Para indicar que el artículo cumple con el mecanismo en cierto grado.
- 0.0 – Para indicar que el artículo no cumple con el mecanismo en lo absoluto.

Luego de la evaluación del conjunto de EPS, cada estudio poseerá una valoración numérica en el intervalo de 0 - 10 puntos bajo una escala de tipo ordinal, con alguno de los valores cualitativos de calidad: Baja, Media y Alta. Para los autores, la valoración de cada uno de los EPS será utilizada para ponderar la importancia de los estudios individuales al momento de realizar, tanto el análisis, como la síntesis/agregación de los estudios seleccionados.

5. Ejecución del Estudio

5.1. Selección de los estudios primarios

Tomando como base la cadena genérica, se configuró en cada una de las BD, así como en el repositorio seleccionado. Las cadenas generadas se listan en la Tabla 1.

Fuente	Cadena Configurada
IEEE Xplore	cloud database AND (design OR architecture OR modeling)
ACM-DL	“cloud database” AND (“design” OR “architecture” OR “modeling”)
ScienceDirect	“cloud database AND (cloud computing OR modeling OR design OR architecture)”
SpringerLink	“cloud AND database AND (modeling OR architecture, OR design)”
Google Scholar	“cloud database AND (modeling OR design)” “cloud database AND (modeling OR architecture)”

Tabla 1 – Cadenas generadas para las 5 fuentes de información

Con las Cadenas Configuradas (CC) se procedió al proceso de búsqueda en cada una de las cinco fuentes de información, posteriormente se procedió a la aplicación de los CI, y finalmente, en un tercer momento, fueron aplicados los CE obteniendo al final un primer conjunto de 51 EPS. Dicho proceso se realizó entre los meses de noviembre y diciembre de 2021. En la Tabla 2 se puede observar la cantidad de estudios seleccionados por fuente, en cada una de las tres etapas del proceso.

Fuente	E1-CC	E2-CI	E3-CE
IEEE Xplore	6, 020	9	7
ACM-DL	423	15	15
ScienceDirect	914	17	16
SpringerLink	33	10	9
Google Scholar	14, 600	10	4

Tabla 2 – Número de estudios por fuente en cada etapa del proceso

Luego de realizar las tres etapas de la estrategia búsqueda, se utilizó la técnica de bola de nieve hacia adelante sobre el primer conjunto de 51 EPS, dicha actividad se realizó en las últimas dos semanas del mes de enero de 2022; con dicho proceso se incorporaron 21 nuevos estudios, teniendo al final del proceso, un total de 72 EPS para analizar.

5.2. Evaluación de la calidad de los EPS

La evaluación de la calidad para cada uno de los 72 EPS con el instrumento adaptado de Dyba & Dingsøyr (2008) permitió clasificar a: (1) 43 estudios con una valoración > 8 pts., (2) 25 estudios una valoración menor a 8 y mayor a 5, y (3) 4 estudios con una valoración igual o menor a 5 pts. La distribución cualitativa —de acuerdo con su calidad— de los 72 EPS se ilustra en la Figura 1.

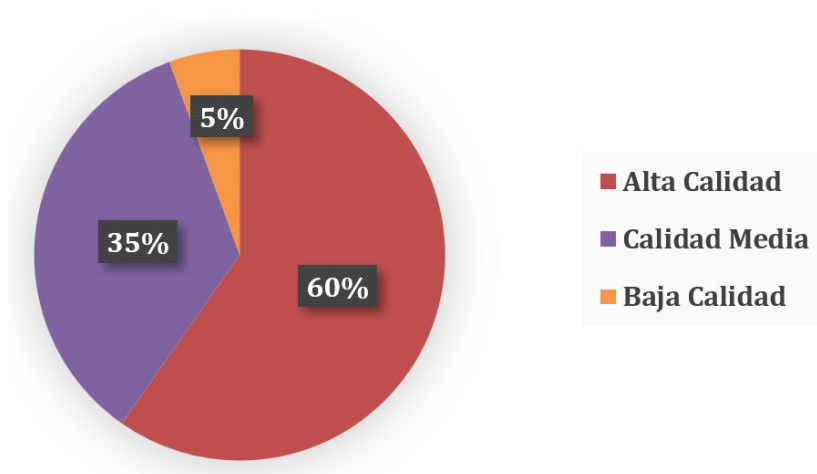


Figura 1 – Porcentaje de EPS de acuerdo con su calidad

6. Resultados del Estudio

La Figura 2 ilustra la distribución de los EPS en la ventana de tiempo considerada en el estudio, se puede identificar que no existe una tendencia clara en investigación sobre la temática del CN; no obstante, las aportaciones al cuerpo de conocimientos han sido constantes, y en particular, de 2019 a 2021 se ha tenido en promedio 10 publicaciones por año, lo cual es una señal de la vigencia en el interés de mantener la investigación sobre la temática que venimos analizando.

Luego de realizar las tres etapas de la estrategia búsqueda, se utilizó la técnica de bola de nieve hacia adelante sobre el primer conjunto de 51 EPS, dicha actividad se realizó en las últimas dos semanas del mes de enero de 2022; con dicho proceso se incorporaron 21 nuevos estudios, teniendo al final del proceso, un total de 72 EPS para analizar.

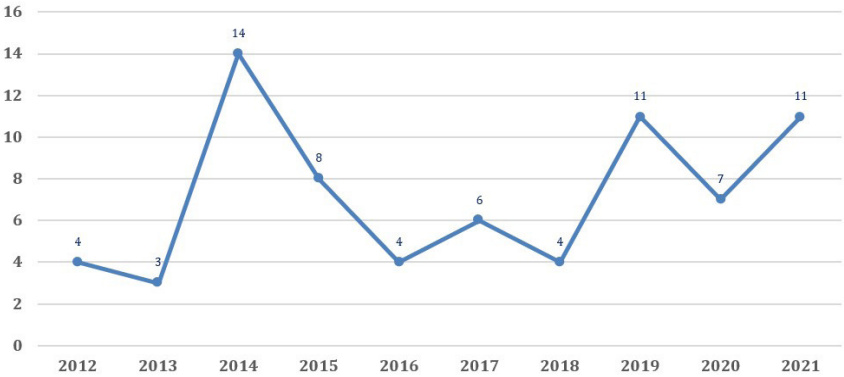


Figura 2 – Frecuencia de los EPS por año de publicación

Por otro lado, también resulta de interés para los investigadores conocer el tipo de investigación empírica realizada, en la Figura 3 se puede observar que el Estudio de Caso es la técnica más utilizada; así mismo, en función del propósito de la investigación, el CN está siendo estudiado empíricamente en todos los niveles: exploratoria, descriptiva, causa-efecto y de mejora (Robson, 2011).

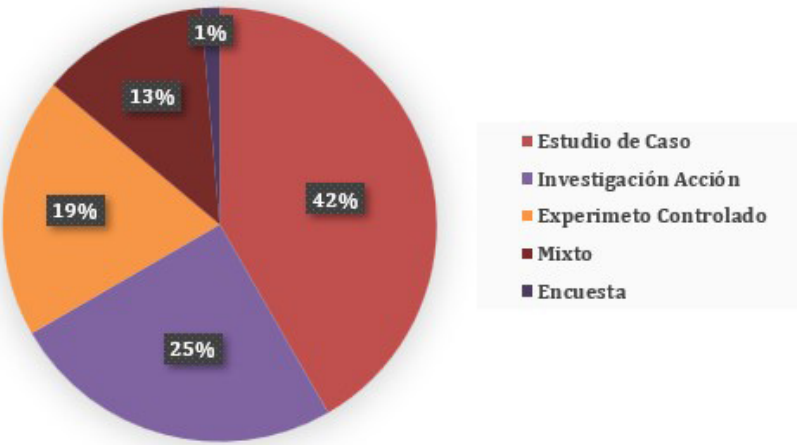


Figura 3 – Porcentaje de EPS por método empírico

Los hallazgos en función de las 5 PI se describen a continuación:

PI-1 *¿Cuáles son las técnicas, procedimientos, o prácticas más utilizadas en el modelado de BD para el CN?*

Con la revisión a los 72 EPS podemos comentar que existen diversas técnicas, procedimientos y prácticas utilizadas para proveer de calidad las BD para el CN, siendo las más citadas, las descritas en la Tabla 3.

Fuente	Descripción/Referencias
Cifrado de datos	Consiste en cifrar los conjuntos de datos mediante algoritmos de encriptación; mediante este proceso los datos se codifican en texto cifrado ilegible para ayudar a mantenerlos seguros. Normalmente debe apoyarse en otros procedimientos (p.e. intercambio secreto). (Ferretti et al., 2014B), (Awadallah & Samsudin, 2021), (Ferretti, Colajanni, & Marchetti, 2014B), (Wong et al., 2014), (Wu et al., 2021), (Kavin & Ganapathy, 2019).
Fragmentación	La Fragmentación de un elemento E, consiste en dividir sus atributos en diferentes fragmentos $E_1...E_n$, de tal manera que solo los atributos en el mismo fragmento son visibles en asociación, permitiendo de esta manera optimizar la eficiencia y la seguridad de las BD. (Branco et al., 2017), (Hwang & Fu, 2016)
Balanceo de cargas	Es un mecanismo mediante el cual se distribuye la carga de trabajo dinámica de manera uniforme entre todos los nodos. El balanceo o equilibrio de carga en la nube también se conoce como equilibrio de carga como servicio (LBaaS). (TengJiao et al., 2012), (Zhang et al., 2021)
Mecanismo de Reversión	El mecanismo de reversión es aplicado cuando se detectan anomalías consecutivas en los esquemas de persistencia o cuando el tiempo de respuesta se ralentiza luego de aceptar un determinado número de consultas; en ambos casos, las últimas transacciones son eliminadas y el esquema de persistencia regresa a un estado previo. (Tan et al., 2019), (Ferretti, Colajanni, & Marchetti, 2012).
Migración y replicación	Cuando las frecuencias de acceso de los elementos de datos conducen a una carga de trabajo desequilibrada en los nodos del sistema, la migración y replicación de arrendatarios/datos para distribuir la carga de trabajo a un conjunto flexible de sitios, es recomendable, lo anterior, con el fin de evitar violaciones a los acuerdos de servicio. (Abdel, Abo-Allian, & Badr, 2021), (TengJiao et al., 2012), (Xiao et al., 2021)
Algoritmos	El desarrollo y modificación de algoritmos específicos es una práctica común para el desarrollo de técnicas y procedimientos para crear u optimizar modelos que proveen aspectos vinculados con la calidad del esquema de persistencia de datos para el CN. (Mulani et al., 2015), (Kavin & Ganapathy, 2019), (Kayed & Omar, 2019), (Alomari & Noaman, 2019), (Tan et al., 2019), (Depoutovitch et al., 2020), (Vainshtein & Gudes, 2021), (Abdel, Abo-Allian, & Badr, 2021), (Antonopoulos et al., 2019), (Cao et al., 2018), (Kim & Song, 2014), (Dokeroglu, Bayir, & Cosar, 2015)
Desagregación de recursos	La desagregación de recursos se ejecuta en los centros de datos desagregados, en éstos los recursos están ubicados en diferentes nodos conectados a través de una red de alta velocidad, lo cual permite que cada tipo de recurso mejore su tasa de utilización y de recuperación a fallos, expanda su volumen y se actualice de forma independiente. (Cao et al., 2021)
DevOps	El propósito de utilizar DevOps para el CN consiste en facilitar el uso de metodologías de desarrollo propias de la IS. Las plataformas Cloud DevOps actualmente permiten a los equipos de desarrollo aprovechar los procesos DevOps para la integración de las herramientas necesarias. La comunidad de BD ha aceptado esta nueva dinámica de trabajo y se encuentra en proceso de adaptación. (Airaj, 2016)

Tabla 3 – Técnicas/Procedimientos más citadas en los EPS

PI-2 *¿Cuáles son los principales modelos de datos que se usan en las BD para el CN?*

Con la revisión de los EPS se encontraron únicamente dos estudios en los que explícitamente se describen modelos de datos particulares: (1) *HashMap* y (2) Un Modelo Basado en Columnas; no obstante, aunque los demás estudios no describieron los modelos de datos que utilizaron, en dichos EPS se habla de estructuras de datos primitivas de manera indirecta o simplemente se omiten, los autores del presente estudio consideran, que posiblemente en dichos estudios se da por hecho que las personas interesadas en el área de BD conocen el estándar o están familiarizadas con los modelos que comúnmente son usados. Cabe destacar que, de la revisión de los estudios previos, se pudo identificar que los autores hacen referencia a modelos de datos conocidos y usados de manera cotidiana, tales como los arreglos y apuntadores, lo cual sustenta la reflexión a la que se llegó en nuestro análisis.

El primer modelo citado es el *HashMap* (Obiniyi, Dzer, & Abdullahi, 2015), dicho modelo se deriva de las *hashtable* y se forma mediante la transformación de un rango de valores clave a un rango de valores de índice de matriz empleando la función que lleva su nombre (*hash function*). Un modelo *HashMap* tiene dos enfoques para su uso: cerrado o abierto; en el caso de estudio primario analizado, los autores utilizan el enfoque abierto, ya que indican, asegura un rendimiento óptimo. Esta estructura de datos es preferida porque el aleatorizar el orden de los datos, provee un acceso eficiente a la información con base en las llaves correspondientes del usuario. El segundo modelo es el descrito en (Wong et al., 2014); los autores mencionan el uso de un modelo de datos basado en columnas, el cual brinda suficiente flexibilidad como para manipular dichas columnas y aplicar las operaciones necesarias en los datos durante las peticiones del usuario, sin dejar de lado la seguridad.

PI-3 *¿Cuáles son los principales aspectos de calidad considerados por las arquitecturas y/o modelos propuestos para las BD utilizadas para el CN?*

Derivado del análisis de los EPS se pudo observar que los estudios no se enfocan en de manera particular en aspectos de calidad específicos, sino más bien en las necesidades que tienen las aplicaciones, servicios o las aplicaciones de software en la nube en general.

Los autores comúnmente recomiendan distribuir la arquitectura de la nube por capas (Alam et al., 2013; Wong et al., 2014) de manera que cada una conlleve responsabilidades específicas y provea garantías para aquellos aspectos clave del CN, como seguridad, consistencia, privacidad, garantías ACID —Atomicidad, Consistencia, Aislamiento y Durabilidad— y garantizando el control de acceso (John & Ramesh, 2017; Zhang et al., 2019; Chitra & Rani, 2014); sin embargo, también se reconoce que las arquitecturas pueden contribuir a la mejora del rendimiento de las BD al apoyar en tareas como la configuración de ajustes o el *buffer* (Tan et al., 2019), el control de cargas (TengJiao et al., 2012), entre otros aspectos que también resultan vitales para el funcionamiento óptimo del entorno.

Por otro lado, con el avance de la tecnología, los autores han tendido a enfocarse en arquitecturas lógicas con marcos de trabajo para mejorar/reforzar las características anteriormente mencionadas (Sahri, Moussa, Long, & Benbernou, 2014) empleando estrategias novedosas que incluyen el uso del aprendizaje máquina (Xiao et al., 2021) o adoptando estructuras basadas en sistemas distribuidos y BD en tiempo real en la

memoria (Stoja et al., 2014) por mencionar algunas innovaciones. También se identificó que los autores consideran que, por ejemplo, los modelos NoSQL poseen grandes ventajas sobre los SQL y ofrecen solución a múltiples problemáticas que surgen con las nuevas necesidades del software en la nube, aunque se habla también de integrar trabajos previos en el área de las BD en la nube para mejorar aquellos aspectos “flojos” en los modelos SQL, pues una gran mayoría del software en la nube emplea este modelo y es imposible que quede obsoleto, al menos en la actualidad.

Independientemente de los criterios/aspectos de calidad considerados, es necesario ofrecer toda la ayuda posible al ingeniero de software que diseña esquemas de persistencia para el CN, a fin de aumentar su productividad y contribuir a la generación de diseños pertinentes a dichos criterios (Blas, Leone y Gonnet, 2019).

PI-4 ¿Cuáles son los principales proveedores de herramientas para el CN?

Con la revisión y análisis a los EPS se pudo observar que, si bien los autores de los estudios presentan herramientas que contribuyen a la gestión de la información en el cómputo en la nube, la gran mayoría se enfoca principalmente en describir las herramientas en función de las necesidades de cada estudio particular, para facilitar tareas que se relacionan con el propósito de su investigación, o probar sus propuestas, por ejemplo, si se habla de la seguridad de un sistema, los autores presentan herramientas que permiten probar la privacidad de los datos al ejecutar peticiones sobre una base de datos cifrada, o herramientas por el estilo, pero siempre relacionadas con la temática y con el objetivo de facilitar las pruebas o las actividades a realizar. Un ejemplo de esto se da en Malhotra et al (2016), donde los autores emplean *CryptDB* como una herramienta para probar peticiones sobre datos cifrados.

En cuanto a los proveedores para el cómputo en la nube, la Figura 4 ilustra la distribución de las 44 menciones identificadas con el análisis de los EPS; cabe mencionar que la categoría denominada “Otros”, hace referencia a entidades como Apache y Yahoo!, que proveen herramientas “independientes” sin una entidad corporativa. Cabe destacar, que de entre los proveedores de entidades corporativas, el más citado fue Amazon con 12 menciones.

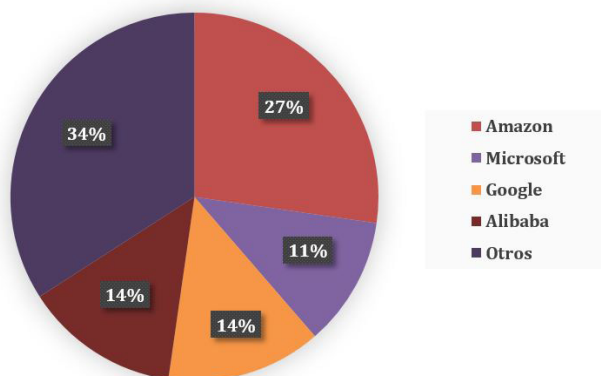


Figura 4 – Proveedores de herramientas para el CN

PI-5 *¿Cuáles son las principales herramientas vinculadas con la gestión de BD para el CN?*

El análisis de los EPS permitió identificar que solamente en siete estudios los autores citan herramientas específicas (ver Tabla 4) que se emplean para la gestión de las BD en la nube, sin embargo, resulta importante resaltar que: (1) las herramientas de las que hablan son las más populares o, al menos, las que más se repiten en los estudios previos analizados y (2) los autores hacen hincapié en ellas porque, tal y como mencionan otros estudios, dichas herramientas tienen características clave como: buen rendimiento, alta disponibilidad, elasticidad, seguridad y consistencia.

Herramienta	Referencias
Microsoft Windows Azure, Amazon Elastic Compute Cloud.	(Stoja et al., 2014)
Big Table de Google, Dynamo de Amazon.	(Sakr, 2013)
Apache Cassandra.	(Silva-Muñoz, Franzin, & Bersini, 2021)
Alibaba DB System.	(Li, 2019)
Amazon AWS, Microsoft Azure, Google’s Cloud Platform y Alibaba.	(Cao et al., 2021)
Alibaba Cloud RDS.	(Xiao et al., 2021)
Big Table de Google, Apache Cassandra, y Dynamo de Amazon.	(Cheng et al., 2014)
Microsoft Windows Azure, Amazon Elastic Compute Cloud.	(Stoja et al., 2014)

Tabla 4 – Herramientas para gestión de BD en la nube citadas en los EPS

Si bien estas características pueden no estar 100% aseguradas en todas las herramientas, se pueden seguir procedimientos o realizar mejoras para asegurar que el funcionamiento y la integridad de los datos no sean comprometidos.

7. Conclusiones y Trabajos Futuros

De acuerdo con el conjunto de 72 EPS analizados, es posible concluir que los aspectos vinculados con los esquemas de persistencia para el cómputo en la nube representan un paradigma aún vigente en el contexto de la investigación en IS.

En cuanto a las técnicas y/o procedimientos para el modelado de BD en la nube, el análisis a los EPS permitió identificar como las más citadas: Cifrado de Datos, Fragmentación, Balanceo de Cargas, Mecanismos de Reversión, Técnicas que hacen uso de Algoritmos, Procedimientos relacionados con la Migración y Replicación, Desagregación de Recursos, así como *DevOps*. Respecto de los modelos de datos utilizados, al parecer siguen siendo utilizados estructuras de datos primitivas y solo se pudo identificar dos modelos particulares: *HashMap* y un esquema basado en columnas. En relación con los aspectos de calidad, se pudo identificar que si bien los estudios no se enfocan de manera

particular en aspectos de calidad específicos, a través de las necesidades descritas para las aplicaciones y/o servicios, es posible identificar como aspectos de calidad que deben estar presentes en las BD para el cómputo en la nube: la disponibilidad, escalabilidad, elasticidad, rendimiento y seguridad, aunque aunado a éstos, es relevante tomar en cuenta otros elementos cuya influencia en la nube es alta, como por ejemplo los principios ACID que están estrechamente relacionados con dichas BD y con los acuerdos de nivel servicio, los cuales se relacionan con el funcionamiento de los sistemas de gestión de bases de datos. Por su parte, en cuanto a los proveedores de servicios/herramientas para el cómputo en la nube, si bien se identificaron diversas herramientas que están siendo utilizadas para actividades de investigación, buena parte de éstas son herramientas “independientes” sin una entidad corporativa, siendo Amazon, Microsoft, Google, y Alibaba, los 4 proveedores corporativos más citados o conocidos (populares). Así mismo, en cuanto a herramientas específicas para la gestión de esquemas de persistencia para el cómputo en la nube, solo en 7 estudios analizados fueron citadas de manera particular un limitado conjunto de herramientas, varias de éstas provistas por los proveedores corporativos denominados populares.

Finalmente, con la RSL se pudo identificar que el estudio del CN sigue en constante evolución, y en lo que compete a la IS, resulta necesario realizar mayor investigación empírica para comparar bondades y condiciones de uso de las diversas técnicas y procedimientos utilizados en el modelado de los esquemas de persistencia para aplicaciones en la nube, así como profundizar en propuestas de mejora —con base en aspectos de calidad— de las herramientas para la gestión de BD utilizadas en los entornos del CN.

Referencias

- Abdel R., Abo-Allan, A., & Badr, B. (2021). A Predictive Multi-Tenant Database Migration and Replication in the Cloud Environment. *IEEE Access*.
- Aguilar, R., Oktaba, H. & Juárez R. (2019). Introducción a la Ingeniería de Software. En Aguilar, R. (ed.) *Ingeniería de Software en México: Educación, Industria e Investigación*. Academia Mexicana de Computación, (pp. 11-14).
- Airaj, M. (2016). Enable cloud DevOps approach for industry and higher education.
- Alam, B., Doja, M. N., Alam, M., & Mongia, S. (2013). 5-Layered Architecture of Cloud Database Management System. *AASRI Procedia*, (pp. 194-199).
- Alomari, E., & Noaman, A. (2019). SeCloudDB: A Unified API for Secure SQL and NoSQL Cloud Databases. *Proceedings of the 2019 3rd International Conference on Cloud and Big Data Computing - ICCBDC*.
- Antonopoulos, P., Byrne, P., Chen, W., Diaconu, C., Kodandaramaih, R. T., Kodavalla, H. & Venkataramanappa, G. M. (2019). Constant Time Recovery in Azure SQL Database. *Proceedings VLDB Endow*, (pp. 2143-2154).
- Awadallah, R., & Samsudin, A. (2021). Using Blockchain in Cloud Computing to Enhance Relational Database Security. *IEEE Access*.

- Bhatti, H. J., & Rad, B. B. (2017). Databases in Cloud Computing: A Literature Review. *Information Technology and Computer Science*, 9(4), pp. 9-17.
- Blas, M., Leone, H. & Gonnet, S. (2019) Modelado y Verificación de Patrones de Diseño de Arquitectura de Software para Entornos de Computación en la Nube. *RISTI – Revista Ibérica de Sistemas e Tecnologías de Informação*, (35).
- Bourque, P. & Fairley, R. (2014). Guide to the Software Engineering Body of Knowledge (SWEBOK V3.0). IEEE Computer Society.
- Branco Jr, E. C., Monteiro, J. M., Reis, R., & Machado, J. C. (2017). A New Mechanism to Preserving Data Confidentiality in Cloud Database Scenarios. *Lecture Notes in Business Information Processing*, (pp. 261-283).
- Cao, W., Zhang, Y., Yang, X., Li, F., Wang, S., Hu, Q. & Gao, Y. (2021). PolarDB Serverless: A Cloud Native Database for Disaggregated Data Centers. *Proceedings of the International Conference on Management of Data*.
- Chitra, K., & Rani, J. B. (2014). DES: Dynamic and Elastic Scalability in Cloud Computing Database Architecture. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 5(1), pp.173-175.
- Cheng, X., Li, H., Wang, Y., Wang, T., & Yang, D. (2014). BF-Matrix: A Secondary Index for the Cloud Storage. WAIM 2014. *Lecture Notes in Computer Science*, (pp. 384-396).
- Depoutovitch, A., Chen, C., Chen, J., Larson, P., Lin, S., Ng, J. & He, Y. (2020). Taurus Database: How to be Fast, Available, and Frugal in the Cloud. *Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data*, (pp. 1463-1478).
- Dokeroglu, T., Bayir, M. A., & Cosar, A. (2015). Robust heuristic algorithms for exploiting the common tasks of relational cloud database queries. *Applied Soft Computing*, (pp. 72-82).
- DOMO. (2021). Domo Resource. Recuperado el 15 de Octubre de 2021, de <https://www.domo.com/learn/infographic/data-never-sleeps-8>
- Dyba T. & Dingsøyr, T. (2008). Strength of evidence in systematic reviews in software engineering. *Proceedings of the Second ACM-IEEE international symposium on Empirical software engineering and measurement*.
- Ferretti, L., Colajanni, M., & Marchetti, M. (2012). Supporting Security and Consistency for Cloud Database. *Lecture Notes in Computer Science*.
- Ferretti, L., Pierazzi, F., Colajanni, M., & Marchetti, M. (2014a). Performance and cost evaluation of an adaptive encryption architecture for cloud databases.
- Ferretti, L., Pierazzi, F., Colajanni, M., & Marchetti, M. (2014b). Scalable architecture for multi-user encrypted SQL operations on cloud database services. *IEEE Transactions on Cloud Computing*, (pp. 448-458).
- Hwang, G.-H., & Fu, S.-K. (2016). Proof of Violation for Trust and Accountability of Cloud Database Systems. *16th IEEE/ACM IS on CCGrid*, (pp. 425-433)

- Jain, N., Deshpande, A. V., & Khanaa, V. (2017). The Effectiveness Of Cloud Databases. *International Conference on Algorithms, Methodology, Models and Applications in Emerging Technologies*.
- John P, A. J., & Ramesh K, B. S. (2017). Proposed Architecture for Improving Security and Consistency of Data Transactions in Cloud Database using Tree-Based Consistency Approach. *IJCSIS*, 15(12), 118-126.
- Karataş, G., C. F., Doğan, G., Konca, C., & Akbulut, A. (2017). Multi-tenant architectures in the cloud: A systematic mapping study. *2017 International Artificial Intelligence and Data Processing Symposium (IDAP)*.
- Kavin, B. P., & Ganapathy, S. (2019). A secured storage and privacy-preserving model using CRT for providing security on cloud and IoT-based applications. *Computer Networks*, (pp. 181-190).
- Kayed, A., & Omar, S. (2019). Periodical key change for cloud mutable security protocol.
- Kim, T., & Song, S. (2014). Dynamic partition lock method to reduce transaction abort rates of cloud database. *Cluster Computing*, (pp. 233-242).
- Kitchenham, B. A., Budgen, D., & Brereton, P. (2015). *Assessing Study Quality*. In Evidence-Based Software Engineering And Systematic Reviews, (pp. 79-88). Florida: CRC Press.
- Kitchenham, B. & Charters, S. (2007). Guidelines for Performing Systematic Literature Reviews in Software Engineering. *Technical Report EBSE-2007-01, School of Computer Science and Mathematics*. Keele University.
- Li, F. (2019). Cloud-Native Database Systems at Alibaba: Opportunities and Challenges. *Proceedings VLDB Endow*, (pp. 2263-2272).
- Malhotra, S., Doja, M. N., Alam, B., & Alam, M. (2016). Cloud Database Management System Security Challenges and Solutions: an analysis. *CSI Transactions on ICT*, (pp. 199-207).
- Marinescu, D. (2018). *Cloud Computing: Theory and Practice*. MK Publishers.
- Mell, P. & Grance, T. (2011). The NIST definition of cloud computing.
- Mostajabi, F., Safaei, A. A., & Sahafi, A. (2021). A Systematic Review of Data Models for the Big Data Problem. *IEEE Access*.
- Mulani, N., Pawar, A., Mulay, P., & Dani, A. (2015). Variant of COBWEB Clustering for Privacy Preservation in Cloud DB Querying. *Procedia Computer Science*.
- Obiniyi, A. A., Dzer, R. M., & Abdullahi, S. (2015). Balancing Query Performance and Security on Relational Cloud Database: An Architecture. *IJCA*,. 34-39.
- Peralta, M., Salgado, C., Montejano, G. & Riesco D. (2016). Las Redes Sociales y la Nube: un nuevo Paradigma para los Procesos de Negocio. *RISTI- Revista Ibérica de Sistemas e Tecnología de Informação*, (8).
- Robson, C. (2011). *Real World Research: A Resource for Social Scientists and Practitioner-Researchers*, 3a Ed, Wiley.

- Sahri, S., Moussa, R., Long, D. D., & Benbernou, S. (2014). DBaaS-Expert: A Recommender for the Selection of the Right Cloud Database. *Foundations of Intelligent Systems*, (pp. 315-324).
- Sakr, S. (2013). Cloud-hosted databases: technologies, challenges, and opportunities. *Cluster Computing*, (pp. 487-502).
- Silva-Muñoz, M., Franzin, A., & Bersini, H. (2021). Automatic configuration of the Cassandra database using irace. *Peerj Computer Science*.
- Stoja, S., Jelačić, B., Vukmirović, S., Čapko, D., & Dalčeković, N. (2014). Architecture of Real-Time Database in Cloud Environment for Distributed Systems. *2nd International Conference on AI, Modelling and Simulation*,
- Tan, J., Zhang, T., Li, F., Chen, J., Zheng, Q., Zhang, P. & Zhang, R. (2019). iBTune: Individualized Buffer Tuning for Large-scale Cloud Database. *VLDB Endo*.
- TengJiao, W., ZiYu, L., BiShan, Y., Jun, G., Allen, H., DongQing, Y., & JinZhong, N. (2012). MBA: A market-based approach to data allocation and dynamic migration for cloud database. *Science China Information Sciences*.
- Vainshtein, Y., & Gudes, E. (2021). Use of Blockchain for Ensuring Data Integrity in Cloud Databases. *Lecture Notes in Computer Science*, (pp. 325-335).
- Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering. *Proceedings of the 18th ICE & A in SE*.
- Wong, W. K., Kao, B., Lok, C. D. W., Li, R., & Yiu, S. M. (2014). Secure Query Processing with Data Interoperability in a Cloud Database Environment. *Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data - SIGMOD '14*, (pp. 1395-1406).
- Wu, W., Xian, M., Parampalli, U., & Lu, B. (2021). Efficient privacy-preserving frequent itemset query over semantically secure encrypted cloud database. *World Wide Web (2021)*, pp. 607-629.
- Xiao, W., Yang, C., Wang, J., Zhu, X., Bao, W., Feng, X. & Liu, L. (2021). YISHAN: Managing Large-scale Cloud Database Instances via Machine Learning. *IEEE Transactions on Services Computing*.
- Zhao, L., Sakr, S., Liu, A., & Bouguettaya, A. (2014). Administración de Datos en la Nube. New York: Springer.
- Zhang, X., Hong, W., Chang, Z., Jin, S., Tan, J., Li, F. & Cui, B. (2021). ResTune: Resource Oriented Tuning Boosted by Meta-Learning for Cloud Databases. *Proceedings of the 2021 International Conference on Management of Data*.
- Zhang, J., Liu, Y., Zhou, K., Li, G., Xiao, Z., Cheng, B. & Li, Z. (2019). An End-to-End Automatic Cloud Database Tuning System Using Deep Reinforcement Learning. *Proceedings of the 2019 International Conference on Management of Data - SIGMOD '19*, (pp. 415-432).

Tendencias y Desafíos de la Computación de Alto Rendimiento en la Nube

Ezra F. Parra-González¹, Uziel Jaramillo-Avila¹, Pablo Salazar-Linares²,
Carlos A. Lara-Álvarez¹

**ezra.parra@cimat.mx; uziel.jaramillo@cimat.mx; salazar.pablo@ibm.com;
carlos.lara@cimat.mx**

¹ Centro de Investigación en Matemáticas A. C. Unidad Zacatecas. CP 98160, Parque Quantum. Zacatecas, Zacatecas, México.

² IBM. Carretera al Castillo 2200 El Quince, CP 45680 El Salto, Jalisco, México.

DOI: 10.17013/risti.49.131–146

Resumen: En los últimos años, la demanda de aplicaciones que requieren Computación de Alto Rendimiento (HPC) ha aumentado considerablemente. Aunque tradicionalmente se ejecutan en infraestructuras privadas, cada vez es más común usar la nube como alternativa, gracias a su flexibilidad. Sin embargo, la computación en la nube también enfrenta desafíos como la gestión de recursos y virtualización, lo que afecta su desempeño. Para abordar estos desafíos, en este artículo se revisan trabajos relacionados con HPC en la nube y se analizan las principales tendencias y desafíos. Finalmente, presentamos una categorización de las tendencias identificadas.

Palabras-clave: Computación de Alto Rendimiento; Computación en la nube.

Trends and Challenges of HPC in the Cloud

Abstract: In recent years, the demand for applications that require High-Performance Computing (HPC) has grown rapidly. These types of applications are traditionally executed on a company's or research center's private infrastructure; however, a different approach can also be used: the cloud. It is important to take advantage of the flexibility that decentralized computing power provides. However, like any other approach, cloud computing faces challenges such as resource management, virtualization, performance issues, etc. In this article, we review related works on HPC in the cloud to identify the main trends and challenges in this area. Finally, we propose a categorization of the identified main trends.

Keywords: High performance computing (HPC); Cloud Computing.

1. Introducción

La computación en la nube ha surgido como una alternativa al uso de recursos centralizados y se ha convertido en una solución viable para empresas y centros de

investigación debido a su aspecto económico (Rashid y Chaturvedi, 2019; Gong et al., 2010). Los proveedores de servicios en la nube ofrecen servicios de renta de recursos, como capacidad de almacenamiento y ciclos de almacenamiento (George y Pramila, 2021; Zhou et al., 2010), y una amplia variedad de servicios, tales como infraestructura como servicio (IaaS), software como servicio (SaaS), plataformas como servicios (PaaS), almacenamiento, copias de seguridad, bases de datos, análisis de datos, Internet de las cosas (IoT), Inteligencia artificial (IA) y aprendizaje automático (machine learning) (Bokhari et al., 2018; Meade et al., 2013; Mell y Grance, 2011).

Existen cuatro principales modelos de implementación de servicios en la nube: privados, públicos, comunitarios e híbridos (Bokhari et al., 2018; Dillon et al., 2010). En el contexto de la computación en la nube, existen diferentes paradigmas, siendo el más conocido el de Infraestructura como servicio (IaaS) (Chhabra et al., 2021), que está relacionado con la noción de cómputo en la nube (Malla y Christensen, 2020; Sotomayor et al., 2009). Existe una cantidad considerable de servicios de cómputo en la nube (Hoefer y Karagiannis, 2010), y casi todas las implementaciones se pueden plantear en términos de un servicio SaaS (Rimal et al., 2009).

Dado que cada tipo de aplicación requiere una cantidad diferente de recursos para desempeñarse de forma efectiva, resulta esencial revisar los estudios más relevantes relacionados con la computación de alto rendimiento (HPC) en la nube e identificar las principales tendencias y desafíos en su adopción. La revisión permitirá generar una categorización de los trabajos revisados según las principales tendencias detectadas y así identificar los tipos de implementaciones más relevantes según su ámbito de aplicación.

Con el fin de orientar el estudio, se plantearon las siguientes preguntas de investigación:

1. ¿Qué tendencias pueden ser identificadas en la adopción de HPC en la nube?
2. ¿Cuáles son los principales desafíos en la adopción de HPC en la nube?
3. ¿Cómo se pueden categorizar las principales aportaciones científicas en el área de la HPC en la nube?

La revisión de la literatura se llevó a cabo mediante una metodología narrativa, que permitió responder de manera puntual a las preguntas de investigación (Wiles, 2011). Para ello, se utilizaron bases de datos como Web of Science, Scopus y Google Scholar, utilizando términos clave relacionados con la computación en la nube y la HPC, tales como “high performance computing”, “cloud computing”, “IaaS”, entre otros. Posteriormente, se realizó una selección de los estudios relevantes y se procedió a su análisis crítico. Una vez identificadas las principales tendencias y desafíos en la adopción de HPC en la nube, se presenta una propuesta de categorización de los trabajos revisados según las principales tendencias detectadas y se expondrán las conclusiones obtenidas a partir de la revisión.

1.1. Aplicaciones de HPC

Aquellas aplicaciones que requieren una gran cantidad de recursos computacionales son conocidas como *Computación de Alto Rendimiento* (HPC). HPC es una disciplina de las tecnologías de la información con una gran cantidad de aplicaciones, con grandes requerimientos de recursos, en diferentes campos como la ciencia, economía, ingeniería

aeroespacial, etc. (Netto et al., 2019 ; Mauch et al., 2012); sin embargo, los sistemas HPC tradicionales son costosos y requieren una gran cantidad de recursos (Lynn et al., 2020), lo que limita su accesibilidad para muchos usuarios. En diversas aplicaciones, se requiere de modelos matemáticos para obtener un resultado satisfactorio, dichos modelos frecuentemente requieren grandes cantidades de recursos computacionales para llevar a cabo experimentos a gran escala o reducir la complejidad computacional en un momento dado (Vecchiola et al., 2009). Tradicionalmente, las aplicaciones de desempeño alto están basadas en sistemas de tipo redes o grupos (Duran-Limon et al., 2011). Ya que el cómputo en la nube ha experimentado un gran crecimiento en el modelo de tipo *“paga sobre la marcha”*, la comunidad científica ha mostrado considerable interés en su uso, dada la capacidad de sólo gastar el dinero correspondiente a los recursos computacionales requeridos en una aplicación dada (Jackson et al., 2010).

1.2. HPC en la nube

El uso de recursos computacionales bajo demanda proporciona una mejora al obtener resultados a problemas grandes en un tiempo aceptable (Wang et al., 2017). El interés en el uso de servicios de HPC en la nube crece conforme a su disponibilidad, poder de cómputo y desempeño incrementan (Expósito et al., 2012). Algunas de las compañías que ofrecen estos servicios lo hacen con tecnología patentada, aunque en algunos casos también se hace a través de código abierto (Dukaric y Juric, 2012; Wind, 2011; Cordeiro et al., 2010). Una parte fundamental del cómputo en la nube es el proporcionar infraestructura como un servicio, esto es logrado típicamente proporcionando máquinas virtuales a los usuarios para completar sus tareas (von Laszewski et al., 2012; Vila et al., 2019).

A pesar de que existen un gran número de iniciativas para el despliegue de HPC en la nube, hay un número de problemas asociados a las aplicaciones de gran desempeño (Benedict, 2013), siendo las principales; virtualización, desempeño, manejo de memoria, seguridad, dominio de aplicación e impactos económicos.

2. Tendencias

La revisión de trabajos realizada permitió de manera general identificar las siguientes tendencias en estudios referentes al HPC en la nube que de manera general se clasificaron en: aportaciones relacionadas a la virtualización, desempeño de los sistemas, aportaciones relacionadas al dominio de la aplicación, gestión de recursos y finalmente los relacionados a los aspectos económicos.

Virtualización: Es una técnica que permite la creación de múltiples instancias de sistemas operativos y aplicaciones en una misma infraestructura de hardware, lo que permite el uso eficiente de recursos y la mejora del desempeño de los sistemas. En el ámbito de la HPC en la nube, la virtualización ha sido objeto de un gran interés en la investigación, con un enfoque particular en el desarrollo de sistemas de peso ligero (Ren et al., 2017; Duran-Limon et al., 2011; Xavier et al., 2013; Lee et al., 2013; Dai et al., 2013; Chaudhary et al., 2008), peso pesado (Regola y Ducom, 2010; Youseff et al., 2006), hipervisores (Younge et al., 2011; Regola y Ducom, 2010; Perera et al., 2011; Aladyshev et al., 2018) y enfoques híbridos (Zhang y Panda, 2017; Stewart et al., 2012; Lee et al., 2013).

La virtualización de peso ligero y los contenedores son enfoques importantes para reducir los costos generales del uso de HPC en la nube (Ren et al., 2017; Colonnelli et al., 2020). Xavier et al. (2013) evaluaron el desempeño de la virtualización basada en contenedores para HPC, mientras que Lee et al. (2013) investigaron técnicas de optimización como la virtualización ligera y la afinidad de CPU. Guidi et al. (2021) demostraron que dar acceso directo a las I/O a las máquinas virtuales puede reducir la brecha en el desempeño entre implementaciones en la nube y nativas.

El hipervisor es un factor crítico en el desempeño experimentado por el sistema operativo invitado y las aplicaciones en máquinas de alto desempeño. Younge et al. (2011) evaluaron la factibilidad de entornos de virtualización HPC y determinaron que los hipervisores deben cumplir ciertas condiciones, siendo la máquina virtual basada en el núcleo (KVM) la mejor opción para implementaciones en la nube. Perera et al. (2011) identificaron características importantes para los clientes al seleccionar un hipervisor.

La virtualización híbrida puede minimizar los problemas causados por un solo enfoque de virtualización. Zhang y Panda (2017) propusieron una virtualización híbrida que utiliza procesos MPI almacenados en diferentes contenedores y máquinas virtuales, mientras que Stewart et al. (2012) presentaron un enfoque híbrido que ofrece un desempeño superior al de la virtualización completa, aunque no tan bueno como el de la para-virtualización o la virtualización del sistema operativo, con la capacidad adicional de virtualizar sistemas operativos invitados arbitrariamente.

Desempeño: En la investigación del HPC en la nube, se ha detectado una tendencia general hacia la mejora del rendimiento de los sistemas. Uno de los mayores desafíos en esta área es la necesidad de mejorar la eficiencia del almacenamiento, la entrada/salida (I/O), la resolución de problemas relacionados con la comunicación, como el ancho de banda, la tolerancia a fallos y las latencias.

Para mejorar la eficiencia del almacenamiento, se están investigando nuevos enfoques para reducir los tiempos de acceso a los datos (Kundu et al., 2010; Agarwal y Prasad, 2012; Chen et al., 2011). Ekanayake y Fox (2010) presentan experimentos en los que se comparan diversas aplicaciones con altos requerimientos de almacenamiento, indicando que muchas de ellas utilizan tareas paralelas para acceder a los datos almacenados en discos locales, aunque no está claro cómo funcionaría este enfoque usando máquinas virtuales.

Los costos de virtualización son especialmente visibles en la virtualización de I/O, donde el host o el monitor de la máquina virtual (VMM) debe revisar cada operación I/O. Esto no es favorable para aplicaciones HPC donde el desempeño de la comunicación es crítico (Maliszewski et al., 2019).

Ren et al. (2017) proponen una infraestructura de nube para aplicaciones HPC basada en la máquina virtual liviana “nOSV” que ejecuta aplicaciones HPC básicas y preserva la flexibilidad de la nube al permitir la virtualización anidada. Zhang y Panda (2017) proponen enfoques de virtualización para evitar cuellos de botella en cargas intensivas de E/S y sugieren una solución compuesta por una localidad de dos capas de alto rendimiento y una biblioteca MPI compatible con NUMA. Además, Expósito et al. (2013) muestran que el uso de discos efímeros proporciona un mejor rendimiento que los volúmenes EBS y que el rendimiento general puede verse limitado por un bajo rendimiento en la red.

Para mejorar el desempeño de I/O en aplicaciones HPC, Shan et al. (2008) proponen caracterizar y predecir el desempeño de I/O seleccionando requerimientos de I/O que correspondan a los patrones de aplicaciones específicas, mientras que Liu et al. (2011) sugieren que la configurabilidad de los servicios de nube pública puede proporcionar oportunidades para mejorar el rendimiento y/o reducir los costos de las aplicaciones de HPC. En particular, los clusters basados en la nube pueden permitir al usuario construir soluciones de I/O paralelas que satisfagan todas las necesidades de la aplicación, según lo analizado por Yu y Vetter (2008).

La variación en los retardos y el rendimiento inestable de la red tienen un fuerte impacto negativo en las aplicaciones científicas (Llopis et al., 2019). Incluso una granularidad de tiempo de decenas de milisegundos puede causar fluctuaciones en el rendimiento de una aplicación rápida de 1 Gb/s a cero (Wang et al., 2010). La configuración de parámetros, la pérdida de paquetes y la virtualización de la máquina son factores que afectan el rendimiento TPC.

El ancho de banda es una condición determinante para los entornos de computación en la nube, y la disponibilidad de interconexiones Infiniband en máquinas virtuales mejora el rendimiento para aplicaciones científicas (Brunetta y Borin, 2019; Ramakrishnan et al., 2012).

La escalabilidad de las aplicaciones HPC depende del soporte de las redes de comunicación en entornos virtualizados. Las latencias de arranque altas y los anchos de banda limitados limitan severamente la escalabilidad de programas intensivos en comunicación (Expósito et al., 2012).

La tolerancia a fallos permite a sistemas HPC con múltiples nodos en la nube completar la ejecución de aplicaciones intensivas, como el enfoque de redundancia en el nivel de proceso (PLR) presentado por Egwuotuoha et al. (2012, 2012b, 2013).

Dominio de aplicación: Otra forma de clasificar las aportaciones científicas en HPC en la nube agrupando es según el dominio de aplicación.

En el ámbito científico, la HPC en la nube se ha utilizado para acelerar la investigación en diversos campos, como la física, la química, la biología y la medicina. De manera general, los científicos han utilizado la HPC en la nube para modelar y simular sistemas complejos (Sharma y Jadhao, 2021; Vecchiola et al., 2009; Changchun et al., 2011; Hoffa et al., 2008; Jackson et al., 2010).

En el dominio de datos intensivos, la HPC en la nube se utiliza para manejar grandes conjuntos de datos y realizar análisis complejos en ellos. La HPC en la nube permite el procesamiento paralelo y distribuido, lo que acelera el tiempo de respuesta y reduce el costo total del procesamiento de datos (Abadi, 2009; Cooper et al., 2010; Shamsi et al., 2013).

En áreas tan particulares como la física de altas energías (Sobie et al., 2011), la HPC en la nube se ha utilizado para modelar y simular fenómenos altamente complejos, lo que ayuda a comprender mejor las problemáticas.

En el campo de la bioinformática, la HPC en la nube se enfoca a demostrar que la computación en la nube y la virtualización son tecnologías que pueden ser utilizadas

para abordar los desafíos en el análisis de datos biológicos, especialmente en términos de escalabilidad y eficiencia de los recursos de cómputo (Yin et al., 2019). Además, se discuten las posibles soluciones para abordar los desafíos y limitaciones actuales de estas tecnologías emergentes, incluyendo la necesidad de desarrollar algoritmos más eficientes y optimizados para su uso en plataformas de cómputo en la nube (Ekanayake et al., 2011; Youseff et al., 2006b; Arrais y Oliveira, 2010; Benedict, 2013; Calabrese y Cannataro, 2015).

En el ámbito de los videojuegos, la HPC en la nube se utiliza para el procesamiento de gráficos y la simulación en tiempo real. Investigaciones como la propuesta por Chen et al., (2014) realizan una evaluación de la calidad de servicio en los sistemas de juego en la nube y con sus experimentos demuestran que aplicaciones de alto rendimiento como los videojuegos se benefician del uso de la computación en la nube y el HPC. En general los desarrolladores de videojuegos pueden utilizar la HPC en la nube para crear mundos virtuales más realistas y detallados, lo que mejora la experiencia del usuario.

La simulación a gran escala es otro dominio de aplicación que fue detectado, en donde la HPC en la nube se utiliza para simular eventos complejos en una amplia gama de áreas, como la ingeniería, la física y la medicina. La HPC en la nube permite a los científicos y a los ingenieros realizar simulaciones complejas que antes no eran posibles debido a las limitaciones de hardware y software.

Gestión: Como ya se ha establecido en éste documento, la evolución de la HPC en la nube ha generado interés por ofrecer recursos bajo demanda y a escala. Sin embargo, la gestión de estos algunos recursos en la nube presenta una serie de desafíos.

En la revisión realizada se identificaron tres tendencias principales relacionadas a la gestión de HPC en la nube: seguridad, memoria y energía. En lo que respecta a la seguridad, es esencial garantizar la privacidad y la integridad de los datos y la infraestructura. El artículo de Grobauer et al., (2011) reporta detalladamente las vulnerabilidades en sistemas HPC en la nube y realizan un análisis de los riesgos de seguridad en los modelos de servicio (IaaS, PaaS y SaaS), identificando las principales amenazas y debilidades en cada uno de ellos. Vaquero et al., (2010) aborda otros aspectos relacionados a la seguridad como es la confidencialidad de los datos, la integridad de los sistemas y la disponibilidad del servicio. Los trabajos proponen estrategias de seguridad para la nube en diferentes áreas, desde la arquitectura y la gestión de identidad hasta la seguridad de la red y la virtualización.

En cuanto a la memoria, es crucial garantizar una gestión eficiente para maximizar el rendimiento en HPC en la nube. Los sistemas de HPC en la nube deben garantizar un acceso rápido y eficiente a la memoria de alta velocidad y proporcionar herramientas para optimizar el uso de la memoria y minimizar la latencia de acceso (Dong et al., 2013; Kundu et al., 2010; Wu et al., 2011).

Otro aspecto que requiere una gestión adecuada en el HPC en la nube es la energética. Administrar adecuadamente la energía permite reducir los costos y la huella de carbono de los centros de datos. Los sistemas de gestión de energía en la nube de HPC deben optimizar el uso de la energía y garantizar la disponibilidad de recursos para las aplicaciones de HPC de manera eficiente (Rodero et al., 2012; Anton et al., 2012).

Economía: La rentabilidad de las soluciones de HPC en la nube es otra área de interés para los investigadores. El objetivo es estudiar los desafíos relacionados al tema económico, incluyendo la gestión de costos y la garantía de un acuerdo de nivel de servicio (SLA) satisfactorio.

En cuanto a la gestión de costos, el análisis de costo-beneficio es una tendencia clave en la HPC en la nube. El análisis de costo-beneficio es una herramienta importante para determinar la viabilidad económica de una solución de HPC en la nube (Carlyle et al., 2010; Zhai et al., 2011). Además, la garantía de un acuerdo de nivel de servicio (SLA) satisfactorio es una tendencia clave en la HPC en la nube. Un SLA es un contrato entre el proveedor de la nube y el cliente que establece los niveles de servicio que se deben cumplir, como la disponibilidad del servicio, el tiempo de respuesta y la capacidad de procesamiento. Los proveedores deben asegurarse de que sus SLA sean claros, específicos y medibles, para que los clientes puedan confiar en la calidad y la consistencia del servicio de HPC en la nube. Asimismo, deben desarrollar herramientas y procesos para monitorear y medir el rendimiento de la nube y garantizar que se cumplan los SLA acordados (Carlyle et al., 2010; Zhai et al., 2011). Por ejemplo, Nasr et al., (2019) presentan un algoritmo para un algoritmo para la programación de tareas de HPC en la nube de forma rentable.

3. Desafíos

Las capacidades de la HPC en la nube han generado un especial interés en las industrias y organizaciones. Sin embargo, su implementación también plantea desafíos que deben ser abordados para su correcto funcionamiento. En este sentido, se han identificado cinco áreas principales en las que se presentan los desafíos más importantes en la HPC en la nube.

Problemas de seguridad y privacidad: Uno de los mayores desafíos de la computación en la nube es la seguridad y la privacidad. Cuando las empresas trasladan sus datos a la nube, le están confiando a un tercero su información confidencial, y eso puede generar inquietudes sobre violaciones de datos, acceso no autorizado y privacidad de datos (Chakraborti, 2022).

Dependencia de la conectividad a Internet: Los servicios de computación en la nube dependen de Internet para la comunicación entre el proveedor de la nube y el usuario. Esto significa que, si hay una interrupción en la conectividad a Internet, es posible que el usuario no pueda acceder a sus datos o aplicaciones. Esto puede ser un desafío importante para las empresas que requieren un acceso continuo a sus datos. Esto puede ser particularmente desafiante en países subdesarrollados ya que las velocidades de transferencia de datos pueden ser lentas (Razi y Batan, 2023). En esos casos es posible que no se pueda servicios como Platform as a Service (PaaS), Infrastructure as a Service (IaaS), and Software as a Service (SaaS).

Bloqueo del proveedor (Vendor lock-in): Cuando las empresas eligen un proveedor de la nube, están bloqueadas en la pila de tecnología de ese proveedor, lo que puede dificultar el cambio a un proveedor diferente en el futuro. Algunas de las soluciones propuestas es seleccionar el mejor proveedor de servicios de acuerdo con diferentes criterios (Mohamed

y Abdelsalam, 2020), o seleccionar al proveedor que logre un equilibrio entre el costo y la disponibilidad (Hari Kumar y Mala, 2022).

Problemas normativos y de cumplimiento: La computación en la nube puede generar problemas normativos y de cumplimiento para las empresas. Estas regulaciones pueden describir los estándares requeridos para tomar medidas de precaución para prevenir ataques (Parast et al. 2022).

Gestión de costos: La computación en la nube puede ser rentable, ya que las organizaciones pueden reducir gastos y aumentar la eficiencia compartiendo recursos y pagando solo por lo que necesitan (Razi y Batan, 2023). Sin embargo, también puede ser costosa si no se administra adecuadamente; por ejemplo, las empresas pueden incurrir en costo de tarifas de transferencia de datos, de almacenamiento u otros cargos.

4. Categorización de trabajos analizados

En la Figura 1 se muestran las categorías más relevantes obtenidas del análisis realizado. Cada categoría agrupa los trabajos revisados según su área de aplicación, lo que permite identificar las principales tendencias en el uso de HPC en la nube en diferentes campos. Las categorías identificadas fueron: virtualización, desempeño, dominio de aplicación, gestión y los trabajos relacionados al tema económico. Las categorías reflejan el amplio espectro de aplicaciones que pueden beneficiarse de la utilización de HPC en la nube.

5. Conclusiones

La computación en la nube se presenta como una solución para el procesamiento de cargas de trabajo de alto rendimiento. Aunque tanto las aplicaciones de HPC como las convencionales tienen problemas similares y diferentes, es tarea de los investigadores identificar los problemas que siguen siendo desafíos en el campo y mejorar el estado del arte. Este estudio propuso identificar las tendencias más importantes sobre la computación en la nube de alto rendimiento y destacar las preocupaciones abiertas y cerradas empleando un sistema de clasificación. Los resultados de esta revisión sugieren que, aunque la HPC basada en la nube ha logrado importantes avances, todavía hay áreas en las que se requiere más investigación. Al abordar estos problemas, se pueden mejorar aún más la eficiencia, la escalabilidad y la robustez de los sistemas de HPC basados en la nube. Además de identificar las áreas de oportunidad en el campo de la HPC basada en la nube, es importante considerar cómo estos vacíos pueden abordarse. Una posible estrategia es mediante el desarrollo de herramientas y recursos para la HPC en la nube. Por ejemplo, se podría crear una plataforma para evaluar el rendimiento y la eficiencia de diferentes proveedores de la nube y plataformas de virtualización para aplicaciones de HPC. También se podrían desarrollar herramientas de monitoreo y análisis para ayudar a los usuarios a comprender y optimizar el rendimiento de sus aplicaciones. En última instancia, el objetivo es mejorar la eficiencia y la accesibilidad de la HPC basada en la nube para una amplia gama de usuarios.

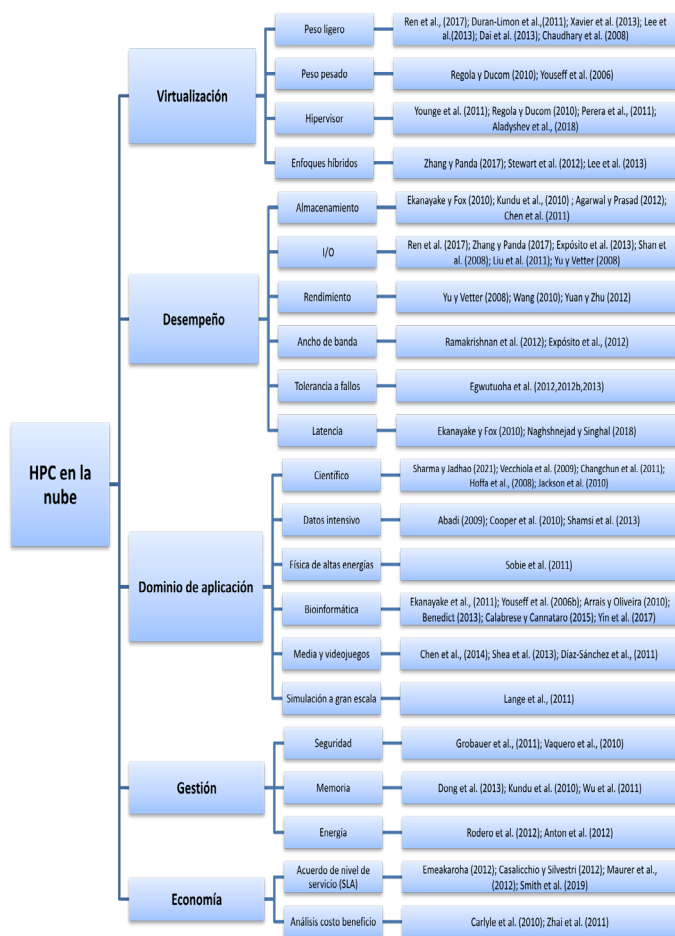


Figura 1 – Categorías de problemáticas específicas a diferentes tipos de HPC en la nube

Referencias

- Abadi, D. J. (2009). *Data management in the cloud: Limitations and opportunities*. IEEE Data Eng. Bull, 32(1), 3-12.
- Agarwal, D., & Prasad, S. K. (2012). *AzureBench: Benchmarking the Storage Services of the Azure Cloud Platform*. In Parallel and Distributed Processing Symposium Workshops & PhD Forum (IPDPSW), 2012 IEEE 26th International (pp. 1048-1057).
- Aladyshhev, O. S., Baranov, A. V., Ionin, R. P., Kiselev, E. A., & Shabanov, B. M. (2018). Variants of deployment the high performance computing in clouds. En 2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus) (pp. 1453-1457) Russia.

- Anton, B., Rajkumar, B., Young, C. L., & Albert, Z. (2012). *A taxonomy and survey of energy-efficient data centers and cloud computing systems*. Technical Report CLOUDS-TR-2010-3. arXiv preprint arXiv:1007.0066v2.
- Rashid A., Chaturvedi A., (2019). Cloud Computing Characteristics and Services: A Brief Review. International Journal of Computer Sciences and Engineering, 7(2), 421-426.
- Arrais, J.P. & Oliveira, J.L. (2010). *On the exploitation of cloud computing in bioinformatics*. Information Technology and Applications in Biomedicine (ITAB), 2010 10th IEEE International Conference on, 1-4.
- Benedict, S. (2013). *Performance issues and performance analysis tools for HPC cloud applications: a survey*. Computing, 95(2), 89-108.
- Bokhari, M.U., Makki, Q., & Tamandani, Y.K. (2018). A Survey on Cloud Computing. In V. Aggarwal, V. Bhatnagar, & D. Mishra (Eds.), Big Data Analytics . Advances in Intelligent Systems and Computing, vol 654 (pp. 197-212). Springer, Singapore.
- Brunetta, J. R., & Borin, E. (2019). *Selecting efficient cloud resources for HPC workloads*. In Proceedings of the 12th IEEE/ACM International Conference on Utility and Cloud Computing (pp. 155-164).
- Calabrese, B., & Cannataro, M. (2015). *Cloud computing in healthcare and biomedicine*. Scalable computing: practice and experience, 16 (1), 1-18. <https://doi.org/10.12694/scpe.v16i1.1057>
- Carlyle, A. G., Harrell, S. L., & Smith, P. M. (2010). *Cost-effective HPC: The community or the Cloud?*. In Cloud Computing Technology and Science (CloudCom), 2010 IEEE Second International Conference on (pp. 169-176).
- Casalicchio, E., & Silvestri, L. (2012). *Mechanisms for SLA provisioning in cloud-based service providers*. Computer Networks, 56(11), 2671-2686.
- Changchun S., Hiroshi N., Susan J., Kai S., Krishna M., & Yong Q. (2011). *HPC Cloud Applied To Lattice Optimization*. In Proceedings of 2011 Particle Accelerator Conference (pp. 3813-3815). IEEE.
- Chaudhary, V., Cha, M., Walters, J. P., Guercio, S., & Gallo, S. (2008). *A comparison of virtualization technologies for HPC*. In Advanced Information Networking and Applications, 2008. AINA 2008. 22nd International Conference on (pp. 861-868). IEEE.
- Chhabra, A., Singh, G., & Kahlon, K. S. (2021). *Multi-criteria HPC task scheduling on IaaS cloud infrastructures using meta-heuristics*. Cluster Computing, 24, 885-918. <https://doi.org/10.1007/s10586-021-03387-2>
- Chen, H., Cheng, K., Jheng, Z., Pan, Y., Wu, C. & Yu, H. (2011). *Easy Virtual Cluster with Cloud Storage for HPC Numerical Simulation*. Commerce and Enterprise Computing (CEC), 2011 IEEE 13th Conference on , vol., no., pp.197,200, 5-7 Sept. 2011. <https://doi.org/10.1109/CEC.2011.36>

- Chen, K.-T., Chang, Y.-C., Hsu, H.-J., Chen, D.-Y., Huang, C.-Y., & Hsu, C.-H. (2014). *On the quality of service of cloud gaming systems*. IEEE Transactions on Multimedia, 16(2), 480-495.
- Chakraborti, A., Curtmola, R., Katz, J., Nieh, J., Sadeghi, A. R., Sion, R., & Zhang, Y. (2022). *Cloud computing security: foundations and research directions*. Foundations and Trends in Privacy and Security, 3(2), 103-213.
- Colonnelli, I., Cantalupo, B., Merelli, I., & Aldinucci, M. (2020). *StreamFlow: cross-breeding cloud with HPC*. IEEE Transactions on Emerging Topics in Computing, 9(4), 1723-1737.
- Cooper, B.F., Silberstein, A., Tam, E., Ramakrishnan, R., & Sears, R. (2010). *Benchmarking cloud serving systems with YCSB*. Proceedings of the 1st ACM symposium on Cloud computing, 2010.
- Cordeiro, T., Damalio, D., Pereira, N., Endo, P., Palhares, A., Gonçalves, G., ... & Mångs, J. E. (2010). *Open source cloud computing platforms*. In Grid and Cooperative Computing (GCC), 2010 9th International Conference on (pp. 366-371). IEEE.
- Dai, Y., Qi, Y., Ren, J., Shi, Y., Wang, X., & Yu, X. (2013). *A lightweight VMM on many core for high performance computing*. In Proceedings of the 9th ACM SIGPLAN/SIGOPS international conference on Virtual execution environments (VEE '13) (pp. 111-120). ACM.
- Díaz-Sánchez D., Almenarez F., Marín A., Proserpio, D. & Cabarcos PA. (2011). *Media cloud: an open cloud computing middleware for content management*. IEEE Trans Consumer Electron 57(2):970-978
- Dillon, T., Wu, C., & Chang, E. (2010). *Cloud computing: Issues and challenges*. In Advanced Information Networking and Applications (AINA), 2010 24th IEEE International Conference on (pp. 27-33). IEEE.
- Dong O.S., Hong J.C., Jae H.P. & Cheol H. K. (2013). *Analysis of Memory Management Policies for Heterogeneous Cloud Computing*. Information Science and Applications (ICISA), 2013 International Conference on, pp.1,3, 24-26 June 2013 <https://doi.org/10.1109/ICISA.2013.6579403>
- Dukaric, R., & Juric, M. B. (2012). *Towards a unified taxonomy and architecture of cloud frameworks*. Future Generation Computer Systems.
- Duran-Limon, H. A., Silva-Bañuelos, L. A., Tellez-Valdez, V. H., Parlavantzas, N., & Zhao, M. (2011). *Using Lightweight Virtual Machines to Run High Performance Computing Applications: The Case of the Weather Research and Forecasting Model*. In Utility and Cloud Computing (UCC), Fourth IEEE International Conference on (pp. 146-153). IEEE.
- Egwutuoha, I.P., Chen, S., Levy, D., Selic, B., & Calvo, R. (2012). *A Proactive Fault Tolerance Approach to High Performance Computing (HPC) in the Cloud*. In Cloud and Green Computing (CGC), Second International Conference on (pp. 268-273). IEEE.

- Egwutuoha, I.P., Shiping C., Levy, D. & Selic, B. (2012). *A Fault Tolerance Framework for High Performance Computing in Cloud*. Cluster, Cloud and Grid Computing (CCGrid), 12th IEEE/ACM International Symposium on, 13-16 May 2012, (pp.709,710). <https://doi.org/10.1109/CCGrid.2012.80>
- Egwutuoha, I.P., Shiping, C., Levy, D., Selic, B. & Calvo, R. (2013). *Energy Efficient Fault Tolerance for High Performance Computing (HPC) in the Cloud*. Cloud Computing (CLOUD), IEEE Sixth International, June 28 2013-July 3 2013, (pp.762,769). <https://doi.org/10.1109/CLOUD.2013.69>
- Ekanayake, J., & Fox, G. (2010). *High performance parallel computing with clouds and cloud technologies*. In Cloud Computing (pp. 20-38). Springer Berlin Heidelberg.
- Ekanayake, J., Gunarathne, T. & Qiu J. (2011). *Cloud Technologies for Bioinformatics Applications*. IEEE Transactions on Parallel and Distributed Systems, 22(6).
- Expósito, R. R., Taboada, G. L., Ramos, S., González-Domínguez, J., Touriño, J., & Doallo, R. (2013). *Analysis of I/O performance on an amazon EC2 cluster compute and high I/O platform*. Journal of Grid Computing, 1-19.
- Expósito, R. R., Taboada, G. L., Ramos, S., Touriño, J., & Doallo, R. (2012). *Performance analysis of HPC applications in the cloud*. Future Generation Computer Systems.
- George, S. S., & Pramila, R. S. (2021). A review of different techniques in cloud computing. Materials Today: Proceedings, 46(Part 17), (pp. 8002-8008).
- Gong, C., Liu, J., Zhang, Q., Chen, H., & Gong, Z. (2010). *The characteristics of cloud computing*. In Parallel Processing Workshops (ICPPW), 2010 39th International Conference on (pp. 275-279). IEEE.
- Grobauer, B., Walloschek, T. & Stocker, E. (2011) *Understanding cloud computing vulnerabilities*. Secur Privacy IEEE 9(2),50–57
- Guidi, G., Ellis, M., Buluç, A., Yelick, K., & Culler, D. (2021). *10 years later: Cloud5computing is closing the performance gap*. In Companion of the ACM/ SPEC International Conference on Performance Engineering (41-48).
- Hari Kumar, P., & Mala, G. A. (2022). *H2RUN: An efficient vendor lock-in solution for multi-cloud environment using horse herd Runge Kutta based data placement optimization*. Transactions on Emerging Telecommunications Technologies, 33(9), e4541.
- Hoefer, C. N., & Karagiannis, G. (2010). *Taxonomy of cloud computing services*. In GLOBECOM Workshops (GC Wkshps), 2010 IEEE (pp. 1345-1350).
- Hoffa, C., G. Mehta, T. Freeman, E. Deelman, K. Keahey, B. Berriman & Good J. (2008). *On the Use of Cloud Computing for Scientific Workflows*. SWBES.
- Jackson, K., Ramakrishnan, L., Muriki, K., Canon, S., Cholia, S., Shalf, J., Wasserman, H., & Wright, N. (2010). *Performance Analysis of High Performance Computing Applications on the Amazon Web Services Cloud*. In 2nd IEEE International Conference on Cloud Computing Technology and Science. IEEE, 2010, (pp. 159-168).

- Kundu, A., Banerjee, C., Guha, S.K., Mitra, A., Chakraborty, S., Pal, C. & Roy, R. (2010). *Memory utilization in cloud computing using transparency*. Computer Sciences and Convergence Information Technology (ICCIT), 2010 5th International Conference on, 22-27, <https://doi.org/10.1109/ICCIT.2010.5711023>.
- Lange, J., Pedretti, K., Dinda, P., Bridges, P., Bae, C., Soltero, P. & Merritt, A. (2011). *Minimal Overhead Virtualization of a Large Scale Supercomputer*. Proceedings of the 2011 ACM SIGPLAN/SIGOPS. VEE Conference, 2011.
- Lee, P. F., Kaufmann, R., & Milojevic, D. (2013). *The Who, What, Why and How of High Performance Computing Applications in the Cloud*. IEEE Cloud Computing, 1(2), 24-31.
- Llopis, P., Lindqvist, C., Høimyr, N., van der Ster, D., & Ganz, P. (2019). *Integrating HPC into an agile and cloud-focused environment at CERN*. In EPJ Web of Conferences (Vol. 214, p. 07025). EDP Sciences.
- Liu, M., Zhai, J., Zhai, Y., Ma, X., & Chen, W. (2011). *One optimized I/O configuration per HPC application: leveraging the configurability of cloud*. In Proceedings of the 2nd ACM SIGOPS Asia-Pacific Workshop on Systems (APSys'11), Shanghai, China (pp. 1-5).
- Lynn, T., Fox, G., Gourinovitch, A., & Rosati, P. (2020). *Understanding the determinants and future challenges of cloud computing adoption for high performance computing*. Future Internet, 12(8), 135.
- Malla, S., & Christensen, K. (2020). *HPC in the cloud: Performance comparison of function as a service (FaaS) vs infrastructure as a service (IaaS)*. Internet Technology Letters, 3(1), e137.
- Maliszewski, A. M., Vogel, A., Griebler, D., Roloff, E., Fernandes, L. G., & OA, N. P. (2019, June). *Minimizing communication overheads in container-based clouds for HPC applications*. In 2019 IEEE Symposium on Computers and Communications (ISCC) (pp. 1-6). IEEE.
- Mauch, V., Kunze, M., & Hillenbrand, M. (2012). *High performance cloud computing*. Future Generation Computer Systems, 28(1), 76-90. <https://doi.org/10.1016/j.future.2011.02.005>
- Maurer, M., Emeakaroha, V. C., Brandic, I., & Altmann, J. (2012). *Cost-benefit analysis of an SLA mapping approach for defining standardized Cloud computing goods*. Future Generation Computer Systems, 28(1), 39-47.
- Meade, B., Manos, S., Sinnott, R., Fluke, C., van der Knijff, D., & Tseng, A. (2013). *Research Cloud Data Communities*. Journal of Grid Computing, 11(4), 601-618. <https://doi.org/10.1007/s10723-013-9262-8>
- Mell, P., & Grance, T. (2011). *The NIST definition of cloud computing*. National Institute of Standards and Technology, 53(6), 50.
- Mohamed, A. M., & Abdelsalam, H. M. (2020). *A multicriteria optimization model for cloud service provider selection in multicloud environments*. Software: Practice and Experience, 50(6), 925-947.

- Naghshnejad, M., & Singhal, M. (2018). Adaptive Online Runtime Prediction to Improve HPC Applications Latency in Cloud. En 2018 IEEE 11th International Conference on Cloud Computing (CLOUD) (pp. 762-769).
- Nasr, A. A., El-Bahnasawy, N. A., Attiya, G., & El-Sayed, A. (2019). *Cost-effective algorithm for workflow scheduling in cloud computing under deadline constraint*. Arabian Journal for Science and Engineering, 44, 3765-3780.
- Netto, M. A. S., Calheiros, R. N., Rodrigues, E. R., Cunha, R. L. F., & Buyya, R. (2019). HPC Cloud for Scientific and Business Applications: Taxonomy, Vision, and Research Challenges. ACM Computing Surveys, 51(1), Article 8.
- Parast, F. K., Sindhav, C., Nikam, S., Yekta, H. I., Kent, K. B., & Hakak, S. (2022). *Cloud computing security: A survey of service-based models*. Computers & Security, 114, 102580.
- Perera, P. Muditha, & Keppitiyagama, C. (2011). *A performance comparison of hypervisors*. In 2011 International Conference on Advances in ICT for Emerging Regions (ICTer) (pp. 107-112). IEEE.
- Ramakrishnan, L., Canon, R.S., Muriki, K., Sakrejda, I., & Wright, N.J. (2012). *Evaluating interconnect and virtualization performance for high performance computing*. SIGMETRICS Perform. Eval. Rev, 40(2), 55–60.
- Razi, M., & Batan, A. (2023). *Opportunities and Challenges of Cloud Computing in Developing Countries*. Artificial Intelligence in Society, 3(1), 1-8.
- Ren, J., Qi, Y., Dai, Y., Xuan, Y., & Shi, Y. (2017). nOSV: A lightweight nested-virtualization VMM for hosting high performance computing on cloud. Journal of Systems and Software, 124, 137-152.
- Regola, N., & Ducom, J. C. (2010). *Recommendations for virtualization technologies in high performance computing*. In Proceedings of the 2nd IEEE International Conference on Cloud Computing Technology and Science (CloudCom'10) (pp. 409-416). Indianapolis, IN, USA.
- Rimal, B. P., Choi, E., & Lumb, I. (2009). *A taxonomy and survey of cloud computing systems*. In INC, IMS and IDC, 2009. NCM'09. Fifth International Joint Conference on (pp. 44-51). IEEE.
- Rodero, I., Viswanathan, H., Lee, E.K., Gamell, M., Pompili, D., & Parashar, M. (2012). *Energy-efficient thermal aware autonomic management of virtualized HPC cloud infrastructure*. J. Grid Comput., 10(3), 447–473.
- Shamsi, J., Khojaye, M. A., & Qasmi, M. A. (2013). *Data-Intensive Cloud Computing: Requirements, Expectations, Challenges, and Solutions*. Journal of Grid Computing, 1-30.
- Shan, H., Antypas, K., & Shalf, J. (2008). *Characterizing and predicting the I/O performance of HPC applications using a parameterized synthetic benchmark*. In Proceedings of the 20th ACM/IEEE Supercomputing Conference (SC'08) (pp. 42:1-42:12). Austin, TX, USA.

- Sharma, P., & Jadhao, V. (2021). *Molecular Dynamics Simulations on Cloud Computing and Machine Learning Platforms*. In 2021 IEEE 14th International Conference on Cloud Computing (CLOUD) (pp. 751-753).
- Shea, R., Jiangchuan Liu, Ngai, E. C.-H., & Yong Cui. (2013). *Cloud gaming: architecture and performance*. IEEE Network, 27(4), 16-21.
- Smith, P., Harrell, S. L., Younts, A., & Zhu, X. (2019). *Community Clusters or the Cloud: Continuing cost assessment of on-premises and cloud HPC in Higher Education*. In Proceedings of the Practice and Experience in Advanced Research Computing on Rise of the Machines (learning) (pp. 1-4)
- Sobie, R. J., Agarwal, A., Anderson, M., Armstrong, P., Fransham, K., Gable, I., & Podaima, W. (2011). *Data intensive high energy physics analysis in a distributed cloud*. arXiv preprint arXiv:1101.0357.
- Sotomayor, B., Montero, R. S., Llorente, I. M., & Foster, I. (2009). *Virtual infrastructure management in private and hybrid clouds*. IEEE Internet Computing, 13(5), 14-22.
- Stewart, K. E., Humphries, J. W., & Andel, T. R. (2012). *An Automated Virtualization Performance Analysis Platform*. The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology, 9(3).
- Vaquero, L., Rodero-Merino, L., & Morán, D. (2010). *Locking the sky: A survey on IaaS cloud security*. Computing, 91(1), 93-118.
- Vecchiola, C., Pandey, S., & Buyya, R. (2009). *High-performance cloud computing: A view of scientific applications*. In Pervasive Systems, Algorithms, and Networks (ISPAN), 10th International Symposium on (pp. 4-16). IEEE.
- Von Laszewski, G., Diaz, J., Wang, F., & Fox, G. C. (2012). *Comparison of multiple cloud frameworks*. In Cloud Computing (CLOUD), 2012 IEEE 5th International Conference on (pp. 734-741). IEEE.
- Vila, S., Guirado, F., Lerida, J. L., & Cores, F. (2019). *Energy-saving scheduling on IaaS HPC cloud environments based on a multi-objective genetic algorithm*. The Journal of Supercomputing, 75(3), 1483-1495
- Wang, G. & Ng, T.E. (2010). *The Impact of Virtualization on Network Performance of Amazon EC2 Data Center*". INFOCOM 2010.
- Wang, L., Ranjan, R., Chen, J., & Benatallah, B. (2017). *Cloud computing: methodology, systems, and applications*. CRC press.
- Wind, S. (2011). *Open source cloud computing management platforms: Introduction, comparison, and recommendations for implementation*. In Open Systems (ICOS), 2011 IEEE Conference on (pp. 175-179). IEEE.
- Wiles, R., Crow, G., & Pain, H. (2011). *Innovation in qualitative research methods: A narrative review*. Qualitative Research, 11(5), 587-604.

- Wu, H., Tantawi, A. N., Diao, Y., & Wang, W. (2011). *Adaptive memory load management in cloud data centers*. IBM Journal of Research and Development, 55(6), 5:1-5:10. <https://doi.org/10.1147/JRD.2011.2170869>
- Xavier, M.G., Neves, M.V., Rossi, F.D., Ferreto, T.C., Lange, T., & De Rose, C.A.F. (2013). *Performance Evaluation of Container Based Virtualization for High Performance Computing Environments*. Euromicro International Conference on Parallel, Distributed and Network-Based Processing, (pp. 233-240).
- Yin, Z., Lan, H., Tan, G., Lu, M., Vasilakos, A. V., & Liu, W. (2017). *Computing Platforms for Big Biological Data Analytics: Perspectives and Challenges*. Computational and Structural Biotechnology Journal, 15, (pp. 403-411).
- Younge, A. J., Henschel, R., Brown, J. T., von Laszewski, G., Qiu, J., & Fox, G. C. (2011). *Analysis of virtualization technologies for high performance computing environments*. In *Cloud Computing (CLOUD)*. 2011 IEEE International Conference on (pp. 9-16). IEEE.
- Youseff, L., Wolski, R., Gorda, B., & Krintz, C. (2006). *Paravirtualization for HPC systems*. In: Proc. International Workshop on XEN in HPC Cluster and Grid Computing Environments (XHPC'06), Sorrento, Italy, (pp. 474-486).
- Youseff, L., Wolski, R., Gorda, B., & Krintz, C. (2006b). *Evaluating the performance impact of Xen on MPI and process execution for HPC systems*. In First International Workshop on Virtualization Technology in Distributed Computing (VTDC 2006). IEEE.
- Yu, W., & Vetter, J. S. (2008). *Xen-based HPC: A parallel I/O perspective*. In: Proc. 8th IEEE International Symposium on Cluster Computing and the Grid (CCGRID'08), Lyon, France, pp. 154-161.
- Yuan, W., & Zhu, Y. (2012). *An autonomic capacity management approach with cloud insight towards cost-efficient throughput optimization for high performance computing*. In Communication Systems and Network Technologies (CSNT), 2012 International Conference on (740-744). IEEE.
- Zhang, J., Lu, X., & Panda, D. K. (2017). *Designing Locality and NUMA Aware MPI Runtime for Nested Virtualization based HPC Cloud with SR-IOV Enabled InfiniBand*. In Proceedings of the 13th ACM SIGPLAN/SIGOPS International Conference on Virtual Execution Environments (187-200).
- Zhai, Y., Liu, M., Zhai, J., Ma, X., & Chen, W. (2011). *Cloud versus in-house cluster: evaluating Amazon cluster compute instances for running MPI applications*. In State of the Practice Reports (SC '11) (pp. 11, 10). ACM.
- Zhou, M., Zhang, R., Zeng, D., & Qian, W. (2010). *Services in the cloud computing era: A survey*. In UCS 4th International (pp. 40-46). IEEE.

Contrato Inteligente para la Gestión de Requerimientos en la Construcción de Software

Wilson Soto¹

wsoto@poligran.edu.co

¹ Institución Universitaria Politécnico Grancolombiano, Bogotá, 110231, Colombia.

DOI: 10.17013/risti.49.147–160

Resumen: Uno de los procesos esenciales en la industria de software, es la negociación de los términos de contratación para la construcción de un software. El contrato de construcción de software es un contrato entre un cliente y un desarrollador, donde el desarrollador crea y entrega una pieza específica y personalizada de software al cliente. El contrato de construcción de software es un acuerdo legal vinculante que contiene roles y responsabilidades, tiempos de entrega y pagos, criterios de aceptación, eventualidades en casos de fallas o retrasos y procedimientos para ajustes. En adición, en el contrato son especificados los requerimientos para la construcción del software. Este artículo propone el diseño y construcción de un contrato inteligente desplegado en una red blockchain para la gestión de los requerimientos dentro de una metodología ágil de construcción de software. Incluso el contrato inteligente propuesto puede ser usado para los pagos contractuales de cada requerimiento aceptado por el cliente.

Palabras-clave: Blockchain; Contratos Inteligentes; Ethereum; Procesos de Software.

Smart Contract for Requirements Management in Software Development

Abstract: One of the essential processes in the software industry is the negotiation of contract terms for the development of software. The software development contract is a contract between a customer and a developer, whereby the developer creates and delivers a specific, custom piece of software to the customer. The software development contract is a binding legal agreement that contains roles and responsibilities, delivery and payment times, acceptance criteria, eventualities in cases of failure or delay, and procedures for adjustments. In addition, the contract specifies the requirements for software development. This article proposes the design and implementation of a smart contract deployed in a blockchain network for requirement management within an agile software development methodology. Even the proposed smart contract is used for the contractual payments of each requirement accepted by the client.

Keywords: Blockchain; Ethereum; Smart Contracts; Software Processes.

1. Introducción

La tecnología blockchain es una de las tecnologías con mayor avance en los últimos años debido a sus principales características: infraestructura de computación distribuida, descentralización, inmutabilidad y una garantía de seguridad gracias a la criptografía.

En la tecnología blockchain, los mineros compiten para acumular transacciones, resolviendo el problema de dificultad computacional con el propósito de lograr consenso adicionando las transacciones con bloques a la red blockchain. La primera aplicación de la tecnología blockchain fue el Bitcoin, el cual sin un mecanismo centralizado los usuarios pueden transferir de forma segura dinero. Pero ahora varias plataformas de blockchain tales como Ethereum y Hyperledger han sido propuestas en otros campos más allá de las criptomonedas, como son los contratos inteligentes.

Los contratos inteligentes son programas descentralizados sobre la red blockchain. Las características de los contratos inteligentes son relevantes por el impacto que genera en diversas áreas. Un resumen completo de las aplicaciones basadas en contratos inteligentes se puede encontrar en (Hewa et al., 2021), donde explora entre otras: aplicaciones financieras (gestión de monedas, conocimiento del cliente, seguros, préstamos, procedimientos de auditoría y servicio de mercado de bolsa); cuidado de la salud y servicios relacionados (gestión de información de la salud, protección de datos de investigación clínica, monitoreo automático de pacientes y tratamientos); gestión de identidad y control de acceso (protección de identidad de datos, gestión de identidad descentralizada y políticas de seguridad en accesos de control); propiedad raíz (mejora de seguridad en procesos transaccionales, procesos de impuestos, comisiones por transacciones y tiempo en procesamiento); leyes y e-gobierno (ley, contratos, servicios públicos y democracia nacional); internet de las cosas (contratos inteligentes para compartir fuentes de IoT – *Internet of Things* –, vehículos independientes y ciudad inteligentes); servicios de telecomunicaciones (fuentes compartidas y autónomas en telecomunicaciones, control y gestión de acceso con contratos inteligentes, contratos inteligentes para servicios de *roaming*); gestión de logística (calidad del servicio en carga aérea y de mar, cadena de suministro en la agricultura y trazabilidad en la gestión de la cadena de suministros); contratos inteligentes en la industria (intercambio de energía, industria automotriz, protección ambiental, gestión de la construcción y gestión de tráfico aéreo).

Este artículo presenta el diseño, implementación, pruebas y despliegue de un contrato inteligente para la gestión de requerimientos en el proceso de desarrollo de software. El propósito de contrato inteligente es supervisar los requerimientos, los términos y condiciones del contrato de servicio y garantizar transparencia que no van a hacer modificados por terceros, todo como código programable, incluso las penalidades impuestas. En este sentido, los contratos inteligentes son una solución ideal para reemplazar los contratos clásicos y mejorar la eficiencia, efectividad y seguridad (Hewa et al., 2021). En definitiva, esta aproximación de contrato inteligente como herramienta de gestión de requerimientos dentro del proceso de desarrollo de software puede clasificarse como una aplicación de los contratos inteligentes en acuerdos contractuales para la construcción de software. Principalmente, en la etapa de ingeniería de requerimientos del ciclo de vida del software, donde la modificación de requerimientos es una actividad recurrente.

Esta investigación también provee una ayuda para las personas que quieran diseñar contratos inteligentes o mejorar los procesos de la industria del software.

El artículo esta estructurado de la siguiente forma: la siguiente sección muestra los antecedentes necesarios para el entendimiento de la investigación. La metodología explica el ambiente necesario, el proceso de diseño, la implementación y las pruebas del contrato inteligente como solución a un acuerdo contractual para la construcción de un software. La sección de resultados muestra el despliegue del contrato inteligente en una red blockchain. La última sección muestra las conclusiones de la investigación.

2. Antecedentes

Hoy en día la red *blockchain* es una de las tendencias tecnológicas en la industria. Esta red es una base de datos descentralizada y organizada por bloques que son inmutables. Uno de los usos más extensos de aplicaciones sobre la red blockchain son las criptomonedas, por ejemplo, *Bitcoin*. Pero otra aplicación importante es el contrato inteligente – en inglés *Smart Contract* (SC) -. Un contrato inteligente fue definido en (Szabo, 1994) como: “un protocolo de comunicación computarizada que ejecuta los términos de un contrato” o también se puede definir como una pieza de código que automáticamente hace respetar los términos entre las partes minimizando la participación de terceros. Muchos autores han contribuido en revisiones sistemáticas sobre las aplicaciones y trabajos a futuro sobre la tecnología blockchain. Por ejemplo, en (Casino et al., 2019) aparece un estudio exploratorio y un análisis descriptivo de la tecnología. La importancia de la investigación radica en la clasificación acorde a áreas temáticas y la taxonomía de aplicaciones propuesta.

Una de las redes más conocidas con la tecnología blockchain para desplegar contratos inteligentes es Ethereum.

2.1. Ethereum

Es una de las plataformas de *blockchain* más usadas hoy en día y reconocida por su criptomoneda nativa Ether, además de ser la primera en permitir desplegar contratos inteligentes. La red Ethereum también puede ser usada para la construcción de varias aplicaciones descentralizadas (Khan et al., 2021).

Las ventajas de la plataforma Ethereum entre otras son: sistema abierto, amplia comunidad desarrollo y disponibilidad en modo público y privado. En contraste, sus desventajas son: la sobrecarga de almacenamiento, demora en el tiempo de aprobación de las transacciones, costos altos por transacción, soporte de un solo lenguaje de programación (*solidity*) y limitaciones en la integración con otras tecnologías (Hewa et al., 2021).

2.2. Contratos inteligentes

El contrato inteligente o *smart contract* es una pieza de código que automáticamente hace respetar los términos entre las partes minimizando la participación de terceros. Un contrato inteligente desplegado en una red con tecnología *blockchain* garantiza su transparencia y seguridad.

Recientemente, varios artículos describiendo investigaciones relacionadas con contratos inteligentes se han presentado como en (Hu et al., 2021). En particular, el artículo muestra la literatura y recursos acerca de los contratos inteligentes entre 2008 y 2020. Además, resume varios desafíos e investigaciones futuras en el área. Otro artículo interesante es (Alharby et al., 2018), donde se clasificó por plataformas, áreas y aplicaciones los contratos inteligentes. Particularmente, se encontraron 40 artículos en el área de la ingeniería de software; 120 artículos relacionados con diversas aplicaciones entre las que se destacan: el internet de las cosas, computación en la nube, sector salud, sector financiero, comercio en la bolsa de valores y gestión de activos. Los autores en (Wang et al., 2018) sugirieron un estudio de contratos inteligentes, incluyendo *frameworks*, mecanismos de operación, plataformas y lenguajes de programación. Además, aplicación de escenarios, desafíos y tendencias de desarrollo futuro. Otro artículo con aportes interesantes es (Kim & Ryu, 2020) donde clasifican 391 artículos relacionados con contratos inteligentes en tres grupos: análisis dinámico, detección de vulnerabilidades y correctitud de programas en análisis estático.

Incluso, los contratos inteligentes se han aplicado como solución a un sistema de exámenes de una universidad (Kumar Samanta et al., 2021).

2.3. Aplicaciones descentralizadas

Una aplicación descentralizada – en inglés *Decentralized Application DApp* – es una aplicación que usa un contrato inteligente a través de una interfaz de presentación o *front-end*. Las aplicaciones descentralizadas se caracterizan por las siguientes cuatro propiedades: código abierto, soporte interno por criptomonedas, consenso descentralizado y puntos no centralizados de falla (Cai et al., 2018). Una metodología para construir una Dapp consiste en combinar el diseño e implementación de un contrato inteligente, el despliegue en una red blockchain y la construcción de una aplicación web para interactuar.

2.4. Contrato de desarrollo de software

En la industria del software, un contrato de desarrollo de software (a veces conocido como acuerdo o contrato de servicios maestro) es un contrato entre dos personas: un cliente y un proveedor de software. El cliente determina las características y funcionalidades requeridas y el proveedor construye el software a la medida a cambio de un pago o precio acordado.

Generalmente, un contrato de desarrollo de software incluye, pero no se limita a las siguientes secciones:

1. Compromisos del desarrollador
2. Responsabilidades del cliente
3. Funcionalidades del software
4. Remuneraciones y pagos
5. Términos y vencimiento del contrato
6. Cláusulas de confidencialidad
7. Derechos de propiedad intelectual

8. Garantías y descargos de responsabilidades
9. Cláusula de no competencia
10. Cláusula de funcionalidad
11. Condiciones generales

2.5. Ingeniería de requerimientos

En el proceso de desarrollo de software, los requerimientos son una etapa clave y frecuentemente la primera. Debido a la rigurosidad y al proceso sistemático que se debe realizar en esta etapa, se conoce como la etapa de la ingeniería de requerimientos. Usualmente, la ingeniería de requerimientos envuelve tres etapas: elicitación, especificación y validación. En especial, uno de los objetivos en la etapa de especificación es escribir los requerimientos en un documento. Este documento de requerimientos es un documento esencial en la construcción de software, incluso para pagar a las personas responsables que lo desarrollan (Sommerville, 2016).

3. Metodología

El contrato inteligente presentado en este artículo puede ser aplicado en una metodología ágil como en (Udvaros et al., 2023), donde el contrato inteligente en la metodología ágil scrum actúa como un artefacto de pila de producto – *producto backlog* - construido por el propietario de producto – *producto owner* -. Además de contener los requerimientos (historias de usuario) del software, las modificaciones realizadas a los requerimientos se registran en el contrato inteligente funcionando como una herramienta de gestión que garantiza la propiedad de la transparencia por estar desplegado en la red blockchain. Incluso el cliente puede validar las historias de usuario terminadas después de cada *sprint* y hacer los pagos contractuales correspondientes (Figura 1).

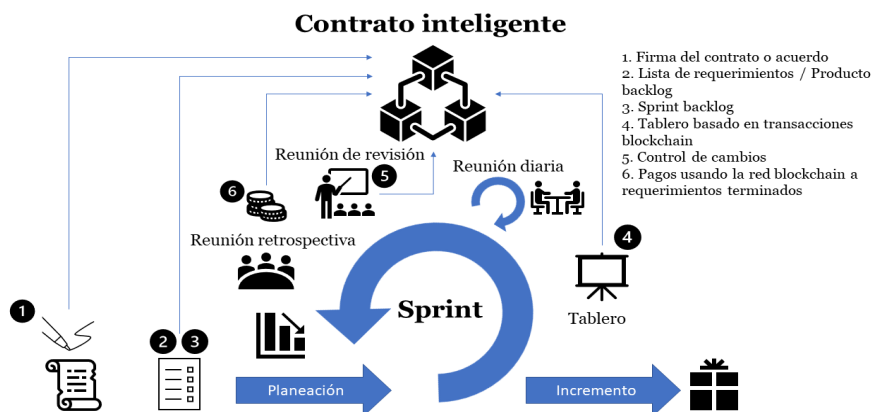


Figura 1 – Aplicación de contrato inteligente como herramienta de gestión de requerimientos en un marco de trabajo Scrum

A continuación, se describe cada una de las etapas del proceso de construcción del contrato inteligente.

3.1. Diseño

En esta fase fueron identificados los actores y las funciones. Los actores son el cliente y el desarrollador. Las funciones identificadas son: la firma del contrato, términos de terminación del contrato y la gestión de los requerimientos (cambios, pagos, tiempos de entrega, ajustes y criterios de aceptación) entre otros (Tabla 1).

Función	Parámetros	Retorno	Descripción
<i>signature</i>	-	-	Firma del acuerdo
<i>initTime</i>	-	uint	Inicio del acuerdo
<i>nowDate</i>	-	uint	Fecha actual
<i>endTime</i>	-	uint	Finalización del acuerdo
<i>showTotalTime</i>	-	uint	Tiempo del contrato
<i>addRequirement</i>	_reqID, _desc	-	Adición de requerimiento
<i>updateRequirement</i>	_reqID, _desc	(string, uint, , ,)	Modificación de requerimiento
<i>getRequirement</i>	_reqID	(string, uint, string, bool, unit)	Consulta de requerimiento
<i>validationRequirement</i>	_reqID	(string, uint, string, bool, ,)	Validación de requerimiento
<i>acceptanceCriteria</i>	_reqID, _desc	-	Criterios de aceptación por requerimiento
<i>validateMinimumPay</i>	valueUsd	-	Valor mínimo en dólares (USD) por transacción
<i>getVersion</i>	-	uint256	Versión de la dirección con la conversión ETH/USD
<i>getPrice</i>	-	uint256	Valor actual de 1 Ether (ETH) en dólares (USD)
<i>getConversionRate</i>	ethAmount	uint256	Conversión de USD a ETH.
<i>fundAccount</i>	-	-	Deposito en cuenta
<i>withdrawAccount</i>	-	-	Retiro en cuenta

Tabla 1 – Funciones del contrato inteligente propuesto

3.2. Implementación

Dos implementaciones fueron necesarias para construir el contrato inteligente. El código está disponible en el siguiente repositorio público: <https://github.com/Volfra/SmartContracts/tree/main/test>

La primera implementación, el contrato inteligente (SWAgreement.sol) usando el lenguaje solidity sobre el IDE Remix y Visual Studio Code y la segunda implementación, una aplicación descentralizada DApp usando el lenguaje Python (deploy.py - fund_and_withdraw.py) sobre Visual Studio Code.

3.3. Pruebas

Se requiere la instalación y configuración de Ganache para pruebas locales y la configuración de una billetera virtual para el despliegue en la red Ethereum. La billetera virtual fue creada en *Metamask* sobre la red de prueba Goerli (Figura 2).

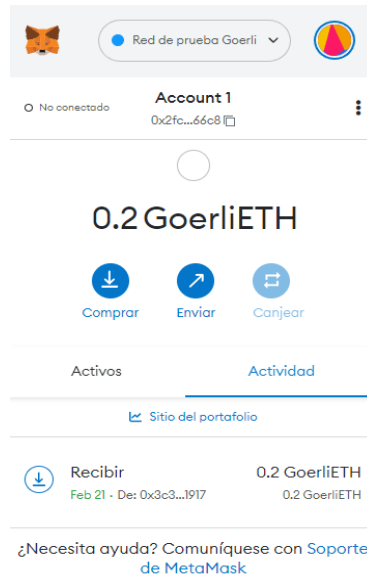


Figura 2 – Billetera virtual en *Metamask*

3.4. Despliegue

El contrato inteligente es desplegado en la red Ethereum. La red Ethereum proporciona una red de prueba llamada *Goerli*. Para ello es necesario abrir una cuenta sobre *Infura*. También es necesario *Brownie* para la aplicación descentralizada DApp. Otra herramienta usada es *Etherscan*, que explora la red Ethereum para validar las transacciones de los contratos inteligentes (<https://goerli.etherscan.io/>).

4. Resultados

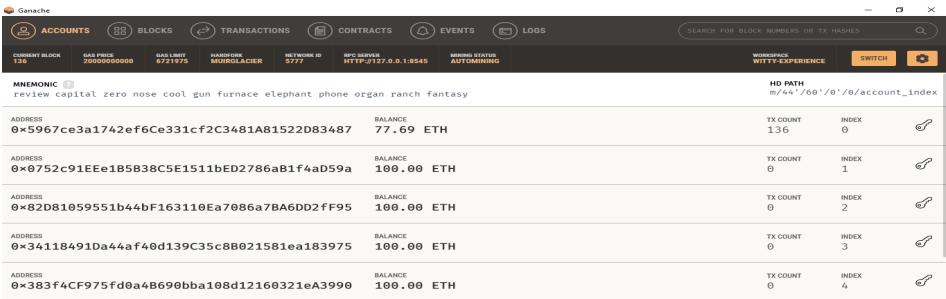
El contrato inteligente fue probado sobre un ejemplo de requerimiento basado en el *framework* Scrum – historia de usuario y criterios de aceptación –. Cada uno de los siguientes pasos genera una transacción y bloque en la red:

1. Despliegue y firma del contrato entre el desarrollador y el cliente.
2. Adición de la historia de usuario UH1.
3. Como cliente necesito hacer una compra así puedo tener comida en mi casa
4. Modificación de la historia de usuario UH1.
 - Como cliente necesito hacer una orden así puedo tener comida en mi casa

- 5. Adición de criterios de aceptación a la historia de usuario.
 - ¿Puedo ver el total acumulado de lo elegido hasta el momento?
 - ¿Puedo cambiar mi orden antes que pague?
 - ¿Puedo hacer una orden a cualquier hora (7x24)?
- 6. Consulta de la historia de usuario UH1.
- 7. Validación de la historia de usuario UH1.

Pago de la historia de usuario UH1 o requerimiento. El pago está alrededor de 164.3 USD. For the date, 1 ETH = \$1643,00 USD, therefore, the payment is around 0.1 ETH or 10000000000000000 Wei

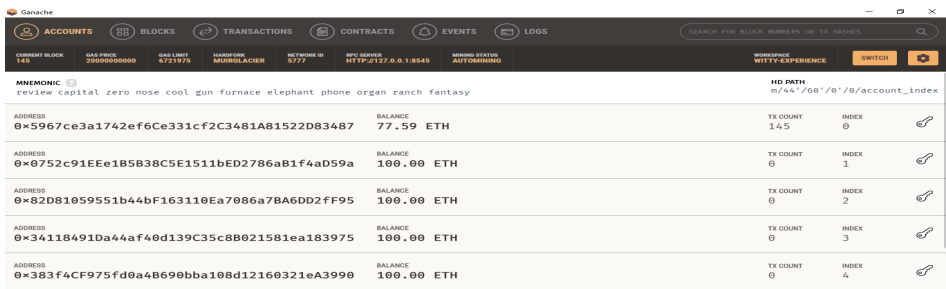
La primera prueba es sobre un ambiente local usando Ganache. La dirección con índice 0 es una cuenta con un monto de 77.69 ETH (figura 3(a)), que posterior a la ejecución de la prueba anterior, es decir, creado, modificado, validado y pagado el requerimiento su monto se reduce en 0.1 ETH, simulando un proceso completo (figura 3(b)). Además, se pueden observar todos los bloques (figura 4) y las transacciones realizadas por la prueba. En la figura se muestran los detalles de las transacciones para el despliegue del contrato y pago de requerimiento (figura 5). En la transacción del pago se puede observar el valor consignado a la dirección donde se desplegó el contrato.



The screenshot shows the Ganache application window. The 'ACCOUNTS' tab is selected. A table lists five accounts with their addresses and balances. The first account has a balance of 77.69 ETH. The other four accounts have a balance of 100.00 ETH. The HD path for each account is shown on the right.

ADDRESS	BALANCE	TX COUNT	INDEX
0x5967ce3a1742ef6Ce331cf2C3481A81522D83487	77.69 ETH	136	0
0x0752c91EEe185838C5E1511bED2786aB1f4aD59a	100.00 ETH	0	1
0x82D81059551b44bF163110Ea7086a7BA6DD2F95	100.00 ETH	0	2
0x34118491Da44af40d139C35c8B021581ea183975	100.00 ETH	0	3
0x383f4CF975fd0a4B690bba108d12160321eA3990	100.00 ETH	0	4

(a)

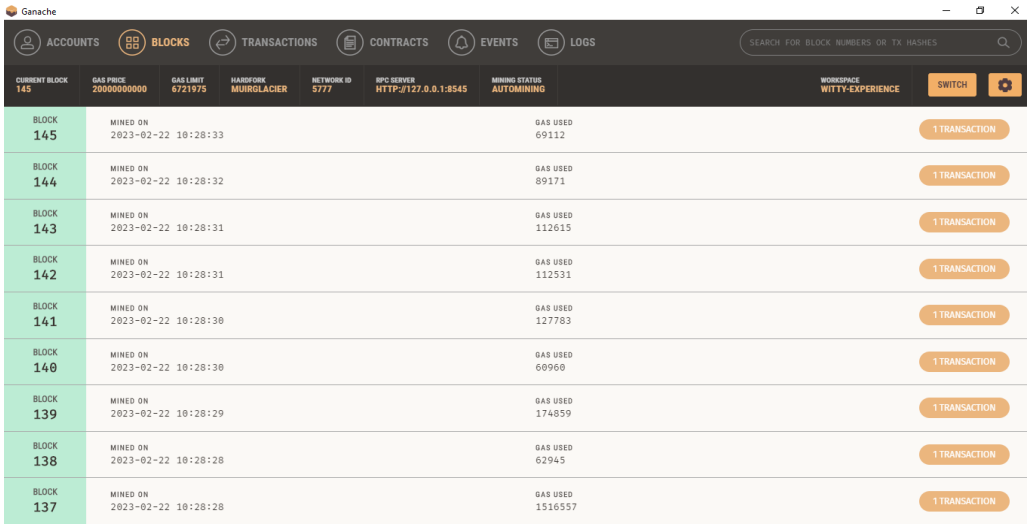


The screenshot shows the Ganache application window after a transaction. The 'ACCOUNTS' tab is selected. The table shows the same five accounts, but the balance of the first account has decreased to 77.59 ETH. The other four accounts still have a balance of 100.00 ETH. The TX COUNT for the first account has increased to 145.

ADDRESS	BALANCE	TX COUNT	INDEX
0x5967ce3a1742ef6Ce331cf2C3481A81522D83487	77.59 ETH	145	0
0x0752c91EEe185838C5E1511bED2786aB1f4aD59a	100.00 ETH	0	1
0x82D81059551b44bF163110Ea7086a7BA6DD2F95	100.00 ETH	0	2
0x34118491Da44af40d139C35c8B021581ea183975	100.00 ETH	0	3
0x383f4CF975fd0a4B690bba108d12160321eA3990	100.00 ETH	0	4

(b)

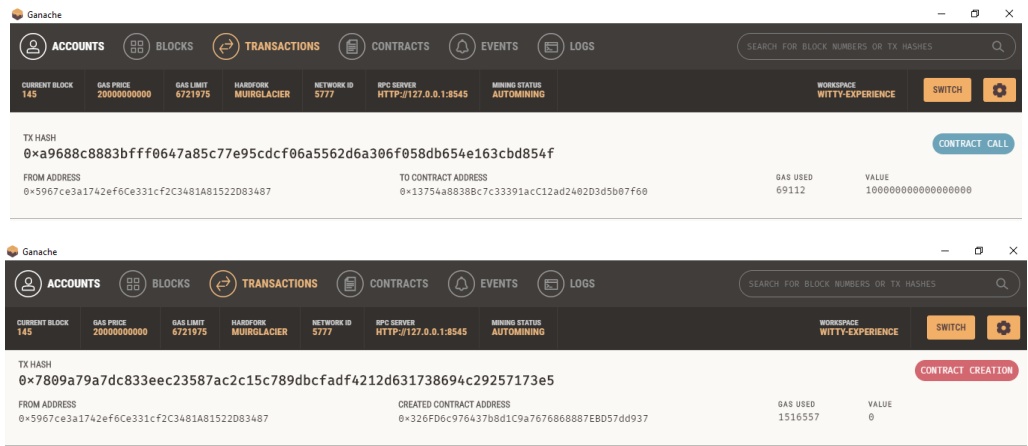
Figura 3 – Billeteras de prueba con Ganache. La billetera usada es la de índice 0 desde donde se hace el pago al contrato. (a) antes del pago. (b) después del pago.



The screenshot shows the Ganache application window. The top navigation bar includes tabs for ACCOUNTS, BLOCKS, TRANSACTIONS, CONTRACTS, EVENTS, and LOGS. The BLOCKS tab is active, displaying a table of mined blocks. The table has columns for BLOCK, MINED ON, and GAS USED. Each row represents a block, with the block number in the first column, the mining time in the second, and the gas used in the third. A '1 TRANSACTION' button is visible next to each row.

BLOCK	MINED ON	GAS USED
145	2023-02-22 10:28:33	69112
144	2023-02-22 10:28:32	89171
143	2023-02-22 10:28:31	112615
142	2023-02-22 10:28:31	112531
141	2023-02-22 10:28:30	127783
140	2023-02-22 10:28:30	60960
139	2023-02-22 10:28:29	174859
138	2023-02-22 10:28:28	62945
137	2023-02-22 10:28:28	1516557

Figura 4 – Bloques generados después de realizar la prueba sobre el ambiente de prueba local usando Ganache.



The screenshot shows the Ganache application window with the TRANSACTIONS tab active. It displays two transaction details. The first transaction is a 'CONTRACT CALL' with a TX HASH of 0xa9688c8883bfff0647a85c77e95cdf06a5562d6a306f058db654e163cbd854f. The second transaction is a 'CONTRACT CREATION' with a TX HASH of 0x7809a79a7dc833eec23587ac2c15c789dbcadfd4212d631738694c29257173e5. Both transactions show the FROM ADDRESS, TO CONTRACT ADDRESS, and the GAS USED and VALUE.

TX HASH	FROM ADDRESS	TO CONTRACT ADDRESS	GAS USED	VALUE
0xa9688c8883bfff0647a85c77e95cdf06a5562d6a306f058db654e163cbd854f	0x5967ce3a1742ef6ce331cf2c3481a81522083487	0x13754a8838bc7c3391ac12ad240203d5b07f60	69112	100000000000000000
0x7809a79a7dc833eec23587ac2c15c789dbcadfd4212d631738694c29257173e5	0x5967ce3a1742ef6ce331cf2c3481a81522083487	0x326f06c976437b8d1c9a7676868887EB057dd937	1516557	0

Figura 5 – Detalle de las transacciones en el despliegue del contrato y pago del requerimiento.

La segunda prueba es la ejecución del contrato inteligente sobre Infura y la red de pruebas Goerli. Las transacciones vistas desde la aplicación descentralizada (figura 6). La dirección del contrato desplegado en la red es: 0x81cDAD9BC22Da4dA3b6E62090e9a248Eb7753E9b

```

Brownie v1.19.3 - Python development framework for Ethereum
TestProject is the active project.
Running 'scripts\deploy.py::main'...
Wallet (Wei): 205095104534387702
Transaction sent:
0xaf0e0f60bb36c6a2618150c6bb3c308d3c354079766edc0a29b34d5cc6455bdf
    Gas price: 30.811919089 gwei    Gas limit: 1681852    Nonce: 11
    SWAgreement.constructor confirmed    Block: 8540260    Gas used: 1528957
(90.91%)
    SWAgreement deployed at: 0x81cDAD9BC22Da4dA3b6E62090e9a248Eb7753E9b
Address contract: 0x81cDAD9BC22Da4dA3b6E62090e9a248Eb7753E9b
Address wallet deploy contract client
0x2fcd44991ca6B22177Dc6b45e8699C6C389066c8
Address wallet client 0x2fcd44991ca6B22177Dc6b45e8699C6C389066c8
Transaction sent:
0xcd8b137fea385e85bf862290c9c87dc92e4c9f1df6d78312de277a0255a82984
    Gas price: 29.207733731 gwei    Gas limit: 74519    Nonce: 12
    SWAgreement.signature confirmed    Block: 8540268    Gas used: 70545
(94.67%)
Agreement signature Date (2023, 2, 23)
Transaction sent:
0xabaa6bf92fe985c02f115d1ebfebcb6fd62b145770753872e0e6ba8d40c9fe2ab
    Gas price: 27.550121552 gwei    Gas limit: 206644    Nonce: 13
    SWAgreement.addRequirement confirmed    Block: 8540282    Gas used: 187859
(90.91%)
Create User History ('UH1', 1677128100, 'Como cliente necesito hacer una
compra así puedo tener comida en mi casa', False, 1677128100, ())
Transaction sent:
0x27696f95a471922de6510d856f3dbb98911ecd295e17b0f75e7f6a55e6e57741
    Gas price: 26.739426503 gwei    Gas limit: 68461    Nonce: 14
    SWAgreement.updateRequirement confirmed    Block: 8540283    Gas used: 60060
(87.73%)
Update User History ('UH1', 1677128100, 'Como cliente necesito hacer una
orden así puedo tener comida en mi casa', False, 1677128100, ())
Transaction sent:
0xee474f920d19a1ff936273700f4c6b9be4b8dfec3f4flc63415022affe21567
    Gas price: 25.538481819 gwei    Gas limit: 149581    Nonce: 15
    SWAgreement.acceptanceCriteria confirmed    Block: 8540284    Gas used:
135983 (90.91%)

```

```

Transaction sent:
0x779fe54e33074d3dfdc2e4020292d61badb2253e3d572cd1f912071c722b371

  Gas price: 25.254563274 gwei   Gas limit: 130494   Nonce: 16

  SWAgreement.acceptanceCriteria confirmed   Block: 8540285   Gas used:
118631 (90.91%)

Transaction sent:
0xf1bbc43f64e2d72cf580cefc0dd396186cf72ef62fb1054d4c3cb5ae8f736564

  Gas price: 25.254563274 gwei   Gas limit: 130586   Nonce: 17

  SWAgreement.acceptanceCriteria confirmed   Block: 8540286   Gas used:
118715 (90.91%)

('UH1', 1677128100, 'Como cliente necesito hacer una orden así puedo tener
comida en mi casa', False, 1677128100, ((1, '¿Puedo ver el total acumulado
de lo elegido hasta el momento?'), (2, '¿Puedo cambiar mi orden antes que
pague?'), (3, '¿Puedo hacer una orden a cualquier hora (7x24)?'))))

Transaction sent:
0x3d892b3a5d1937f9908ba7639e08aa3406fac14b51b4f1a6d5e2381eff5b2caa

  Gas price: 23.188855851 gwei   Gas limit: 115798   Nonce: 18

  SWAgreement.validationRequirement confirmed   Block: 8540290   Gas used:
105271 (90.91%)

('UH1', 1677128100, 'Como cliente necesito hacer una orden así puedo tener
comida en mi casa', True, 1677128220, ((1, '¿Puedo ver el total acumulado
de lo elegido hasta el momento?'), (2, '¿Puedo cambiar mi orden antes que
pague?'), (3, '¿Puedo hacer una orden a cualquier hora (7x24)?'))))

Pago requerimiento UH1 x 164.3 USD (21.02.23) = 0.1 ETH

Transaction sent:
0xd149c74f466a96449dccc56c9f605863c4c7d8c2c034d88b22a09c5f2c39bd48

  Gas price: 22.904947585 gwei   Gas limit: 79543   Nonce: 19

  SWAgreement.fundAccount confirmed   Block: 8540295   Gas used: 75112
(94.43%)

```

Figura 6 – Transacciones generadas después de realizar la prueba sobre la red Goerli en Infura desde la aplicación descentralizada DApp.

Etherscan es usado para observar las transacciones del contrato en la red Ethereum (figura 7). Es posible identificar en las transacciones la creación del contrato, la firma y el pago del requerimiento. Todas las transacciones tienen una dirección – billetera del cliente -, el valor y la tarifa de la transacción. La tarifa de transacción, tarifa de gas o precio de gas que significa el valor a pagar por los recursos computacionales sobre la red Ethereum y es calculado al multiplicar el precio de gas por el gas usado por la transacción.

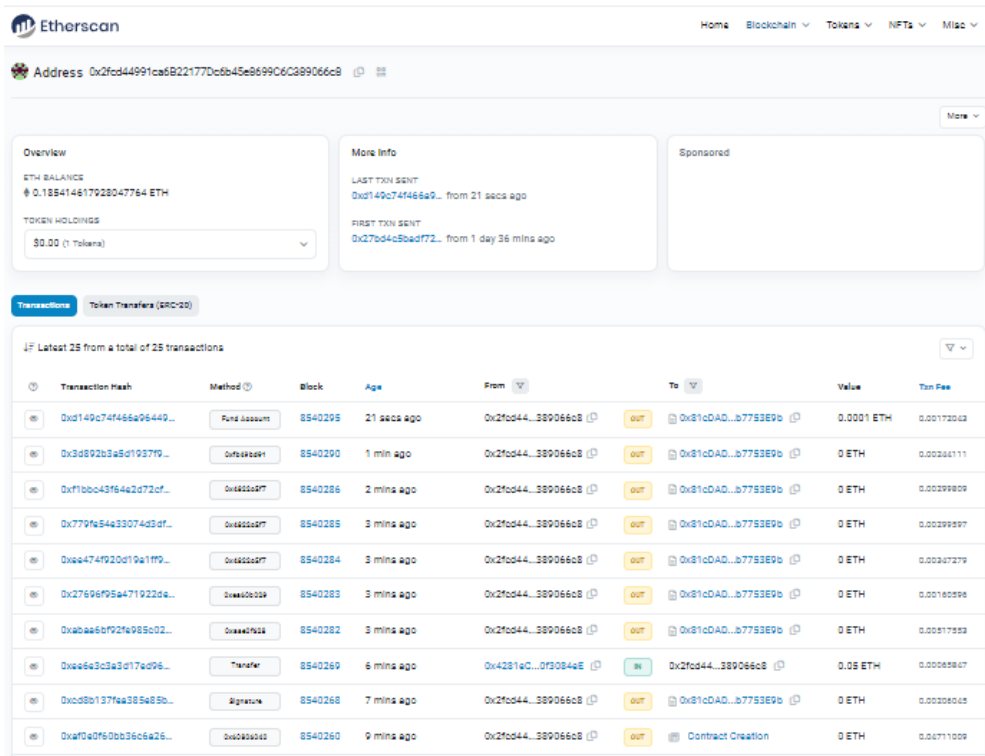


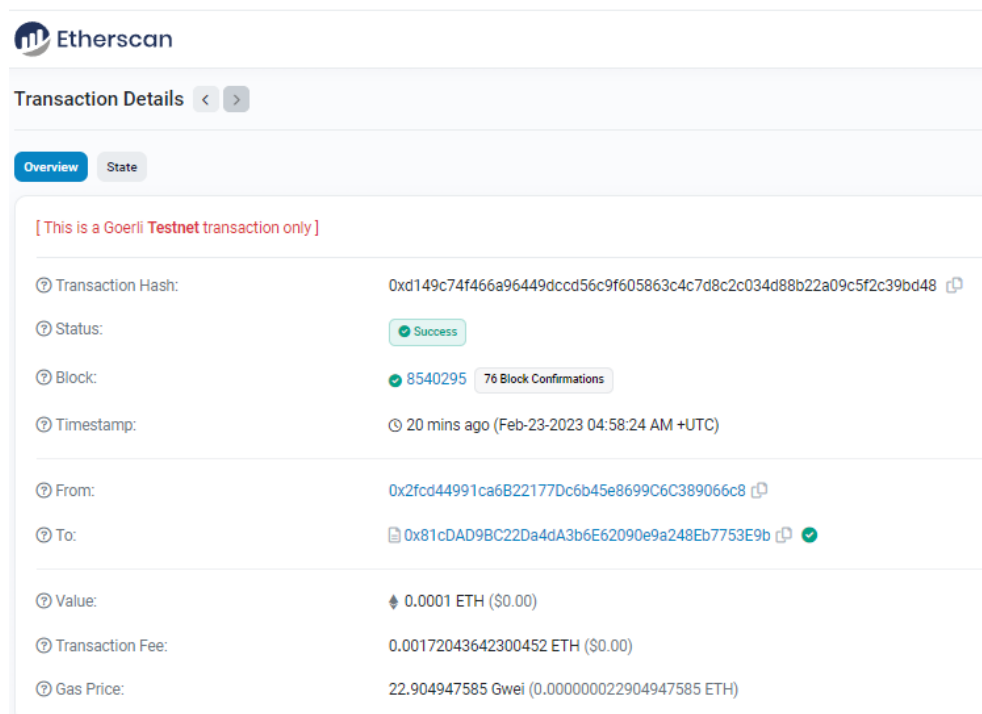
Figura 7 – Transacciones generadas después de realizar la prueba sobre la red Goerli en Infura. Las transacciones pueden ser vistas en el enlace <https://goerli.etherscan.io/address/0x2fcd44991ca6b22177dc6b45e8699c6c389066c8>

En el caso en particular, el valor total de todas las tarifas de transacciones es alrededor de 0,070238890 ETH o 7 USD.

También, Etherscan permite la inspección de los detalles de cada transacción, por ejemplo, el detalle de la transacción del pago del requerimiento (figura 8).

5. Conclusiones

Este artículo mostró un contrato de construcción de software como contrato inteligente sobre la red blockchain. Se incluyó el diseño, implementación, pruebas y despliegue del contrato inteligente sobre la red Ethereum. Una solución de este tipo puede llegar a ser una herramienta importante en la etapa de ingeniería de requerimientos dentro del ciclo de vida del software. Principalmente conociendo los desafíos en la ingeniería de requerimientos como su constante evolución y trazabilidad durante todo el ciclo, por lo tanto, herramientas de gestión que contribuyan con el mejoramiento de los procesos en la construcción de software son valiosas.



The screenshot shows the Etherscan interface for a transaction on the Goerli Testnet. The page title is 'Transaction Details' with navigation arrows. Below the title are two tabs: 'Overview' (selected) and 'State'. A red warning message states: '[This is a Goerli Testnet transaction only]'. The transaction details are listed as follows:

Transaction Hash:	0xd149c74f466a96449dcd56c9f605863c4c7d8c2c034d88b22a09c5f2c39bd48
Status:	Success
Block:	8540295 76 Block Confirmations
Timestamp:	20 mins ago (Feb-23-2023 04:58:24 AM +UTC)
From:	0x2fcd44991ca6B22177Dc6b45e8699C6C389066c8
To:	0x81cDAD9BC22Da4dA3b6E62090e9a248Eb7753E9b
Value:	0.0001 ETH (\$0.00)
Transaction Fee:	0.00172043642300452 ETH (\$0.00)
Gas Price:	22.904947585 Gwei (0.000000022904947585 ETH)

Figura 8 – Detalle de la última transacción después de realizar la prueba sobre la red Goerli en Infura. El detalle de la transacción puede ser visto en el enlace <https://goerli.etherscan.io/tx/0xd149c74f466a96449dcd56c9f605863c4c7d8c2c034d88b22a09c5f2c39bd48>

Los contratos inteligentes desplegados en una red blockchain y usados dentro de la etapa de ingeniería de requerimientos en el ciclo de vida de software tienen entre otras ventajas: precisión, transparencia, seguridad, confianza y no intermediación en la gestión de los requerimientos (tiempos de entrega y pagos, ajustes, riesgos, control de cambios y criterios de aceptación).

El trabajo futuro puede focalizarse en la implementación de contratos inteligentes en diversas etapas del ciclo de vida del software, como también evaluar el impacto de este tipo de herramientas en la calidad del software.

Recomendable investigar el uso de otro tipo de redes con tecnología blockchain para desplegar contratos inteligentes, como por ejemplo aquellas con transacciones sin tarifa.

Referencias

- Alharby, M., Aldweesh, A., & van Moorsel, A. (2018). Blockchain-based Smart Contracts: A Systematic Mapping Study of Academic Research (2018). International Conference on Cloud Computing, Big Data and Blockchain, ICCBB 2018. <https://doi.org/10.1109/ICCBB.2018.8756390>

- Cai, W., Wang, Z., Ernst, J. B., Hong, Z., Feng, C., & Leung, V. C. M. (2018). Decentralized Applications: The Blockchain-Empowered Software System. *IEEE Access*, 6, 53019–53033. <https://doi.org/10.1109/ACCESS.2018.2870644>
- Casino, F., Dasaklis, T. K., & Patsakis, C. (2019). A systematic literature review of blockchain-based applications: Current status, classification and open issues. *Telematics and Informatics*, 36, 55–81. <https://doi.org/10.1016/J.TELE.2018.11.006>
- Hewa, T., Ylianttila, M., & Liyanage, M. (2021). Survey on blockchain based smart contracts: Applications, opportunities and challenges. *Journal of Network and Computer Applications*, 177, 102857. <https://doi.org/10.1016/J.JNCA.2020.102857>
- Hu, B., Zhang, Z., Liu, J., Liu, Y., Yin, J., Lu, R., & Lin, X. (2021). A comprehensive survey on smart contract construction and execution: paradigms, tools, and systems. *Patterns*, 2(2), 100179. <https://doi.org/10.1016/J.PATTER.2020.100179>
- Karamitsos, I., Papadaki, M., Barghuthi, N. B. al, Karamitsos, I., Papadaki, M., & Barghuthi, N. B. al. (2018). Design of the blockchain Smart Contract: A Use Case for Real Estate. *Journal of Information Security*, 9(3), 177–190. <https://doi.org/10.4236/JIS.2018.93013>
- Khan, S. N., Loukil, F., Ghedira-Guegan, C., Benkhelifa, E., & Bani-Hani, A. (2021). blockchain smart contracts: Applications, challenges, and future trends. *Peer-to-Peer Networking and Applications*, 14(5), 2901–2925. <https://doi.org/10.1007/S12083-021-01127-0>
- Kim, S., & Ryu, S. (2020). Analysis of blockchain Smart Contracts: Techniques and Insights. *Proceedings - 2020 IEEE Secure Development, SecDev 2020*, 65–73. <https://doi.org/10.1109/SECDEV45635.2020.00026>
- Kumar Samanta, A., Bidyut, & Sarkar, B., & Chaki, N. (2021). A blockchain-Based Smart Contract Towards Developing Secured University Examination System. *Journal of Data, Information and Management* 2021 3:4, 3(4), 237–249. <https://doi.org/10.1007/S42488-021-00056-0>
- Sommerville, I. (2016). Software Engineering Tenth Edition. *Software Engineering Tenth Edition*, 111–133.
- Szabo, N., 1994. Smart contracts.
- Udvaros, J., Forman, N., & Avornicului, S. M. (2023). Agile Storyboard and Software Development Leveraging Smart Contract Technology in Order to Increase Stakeholder Confidence. *Electronics* 2023, Vol. 12, Page 426, 12(2), 426. <https://doi.org/10.3390/ELECTRONICS12020426>
- Wang, S., Yuan, Y., Wang, X., Li, J., Qin, R., & Wang, F. Y. (2018). An Overview of Smart Contract: Architecture, Applications, and Future Trends. *IEEE Intelligent Vehicles Symposium, Proceedings, 2018-June*, 108–113. <https://doi.org/10.1109/IVS.2018.8500488>

Critérios Editoriais

A RISTI (Revista Ibérica de Sistemas e Tecnologias de Informação) é um periódico científico, 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, 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.

El 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.



Revista Ibérica de Sistemas e Tecnologias de Informação
Revista Ibérica de Sistemas y Tecnologías de Información

©RISTI 2023 <http://www.risti.xyz>

