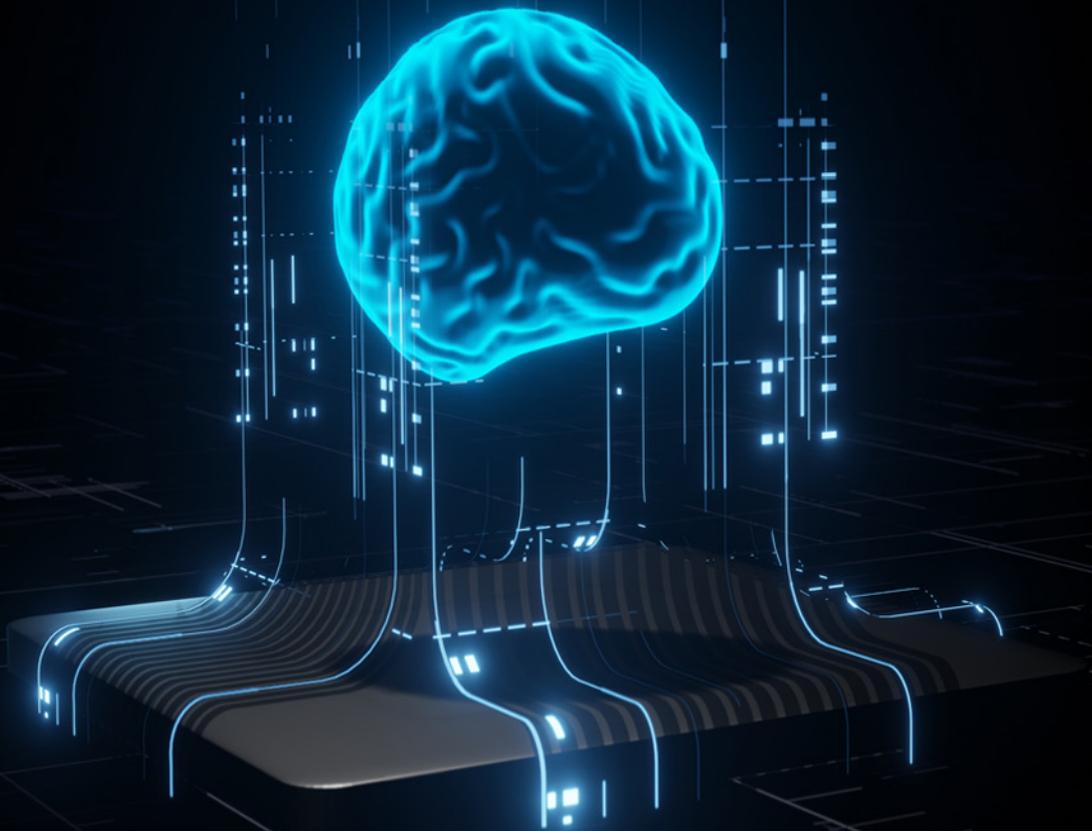


IN TER FASLS



Revista digital de la Carrera de Ingeniería de Sistemas de la Universidad de Lima

Edición n.º 16 // julio-diciembre 2022

Objetivo de *Interfases*

Interfases es una revista digital gestionada por la Universidad de Lima para la publicación de investigaciones originales en áreas temáticas relacionadas con las ciencias de la computación, gestión de datos, aprendizaje automático, sistemas de información, gestión del conocimiento y áreas afines. Se publican artículos científicos y avances de investigación, siempre que cumplan con el proceso de revisión por pares. La revista está indexada en CrossRef, Dialnet, Latindex y DOAJ: Directory of Open Access Journals, y se publica dos veces al año, la primera en julio y la segunda en diciembre. Sin embargo, a partir de julio del 2021, los manuscritos individuales se publicarán tan pronto como estén listos, añadiéndolos progresivamente al contenido de la edición en curso en la modalidad de publicación continua. Los artículos en publicación continua siguen el proceso de revisión por pares, y ya se pueden citar utilizando el año de publicación y el DOI.

Público objetivo

- Profesionales e investigadores relacionados con la ingeniería de sistemas y afines.
- Estudiantes y docentes del pregrado y posgrado de ingeniería de sistemas.
- Público interesado.

ISSN: 1993-4912 (electrónico)
doi: <https://doi.org/10.26439/interfases2022.n016>
Hecho el depósito legal en la Biblioteca Nacional
del Perú n.º 2020-09967

Periodicidad: semestral
Edición: n.º 16
Fecha de publicación: diciembre del 2022
Arbitraje editorial: revisión por pares
doble ciego
Directorios y catálogos: CrossRef, Dialnet,
Latindex y DOAJ

Las opiniones expresadas en los artículos
firmados son de exclusiva responsabilidad de los
autores. Los contenidos de la revista *Interfases* son
de acceso abierto y se encuentran bajo la licencia
Creative Commons.

Universidad de Lima
Av. Javier Prado Este 4600,
Urbanización Monterrico Chico,
Lima 33, Perú
Apartado postal 852, Lima 100, Perú
Teléfono: (511) 437-6767, anexo
30131
fondoeditorial@ulima.edu.pe
<http://www.ulima.edu.pe/>

Edición, diseño y diagramación:
Fondo Editorial

Carrera de Ingeniería de Sistemas
[http://www.ulima.edu.pe/pregrado/
ingenieria-de-sistemas](http://www.ulima.edu.pe/pregrado/ingenieria-de-sistemas)

Contacto: interfases@ulima.edu.pe

DIRECTOR

Julio Alejandro Padilla-Solís
jpadilla@ulima.edu.pe

EDITOR

Hernán Nina Hanco
hninaha@ulima.edu.pe

EDITOR ASOCIADO

Marco Antonio Sotelo Monge
Indra, España

EDITORES ASOCIADOS

Dr. Guillermo Antonio Dávila Calle
Universidad de Lima, Perú
gdavila@ulima.edu.pe

Dra. Ruth María Reátegui Rojas
Universidad Técnica Particular de Loja, Ecuador
rmreategui@utpl.edu.ec

COMITÉ EDITORIAL

Álvaro Talavera-López
ag.talaveral@up.edu.pe
Universidad del Pacífico, Perú

Carlos Mugruza-Vassallo
cmugruza@yahoo.com
Universidad Nacional Tecnológica de Lima Sur, Perú

César Beltrán-Castañón
cbeltran@pucp.pe
Pontificia Universidad Católica del Perú

Effie Lai-Chong Law
lai-chong.law@durham.ac.uk
Durham University, Inglaterra

Hugo Alatrasta-Salas
halatristas@up.edu.pe
Universidad del Pacífico, Perú

Ian D. Sanders
sandeid@unisa.ac.za
University of South Africa, Sudáfrica

Juan Gutiérrez-Cárdenas
jmgutier@ulima.edu.pe
Universidad de Lima, Perú

Michael Dorin
michael_andrew.dorin@stud-mail.uni-wuerzburg.de
University of St. Thomas, Estados Unidos

Nelly Condori Fernández
n.condori.fernandez@usc.es
Universidade de Santiago de Compostela, España

Víctor Ayma-Quirita
vayma@ulima.edu.pe
Universidad de Lima, Perú

EQUIPO DE GESTIÓN

Angelo Rodrigo Taco Jiménez
ataco@ulima.edu.pe
Universidad de Lima, Perú

REVISORES

Dr. Alejandro Apaza Tarqui (Universidad Nacional del Altiplano Puno, Perú)
Dr. Luis Rodrigo Barba Guamán (Universidad Técnica Particular de Loja, Ecuador)
Dra. Lorena Barona López (Escuela Politécnica Nacional, Quito, Ecuador)
Dr. Isaias Bianchi (Al-Farabi Kazakh National University, Kazajistán)
Mg. José Alberto Caballero Ortiz (Universidad de Lima, Perú)
Dr. Dennis Iván Candia Oviedo (Universidad Nacional de San Antonio Abad del Cusco, Perú)
Mg. Daniel Enrique Cárdenas Salas (Universidad de Lima, Perú)
Mg. William Alberto Chávez Espinoza (Universidad Pública de Navarra, España)
Dr. Edmanuel Cruz (Universidad Tecnológica de Panamá)
Dra. María de León Sigg (Universidad Autónoma de Zacatecas, México)
Dr. Ignacio Díaz-Cano (Universidad de Cádiz, España)
Dra. Simena Dinas (Universidad de San Buenaventura, Colombia)
Dr. Dario Francisco Dueñas Bustinza (Universidad Nacional Federico Villarreal, Perú)
Dr. Edwin Jonathan Escobedo Cárdenas (Universidade Federal de Ouro Preto, Brasil)
Mg. Jisbaj Gamarra Salas (Universidad de Granada, España)
Dra. Lillyana María Giraldo Marín (Universidad de Medellín, Colombia)

Mg. David Enrique González Pedroza (IEEE, Estados Unidos)
Dr. Juan Gutiérrez Cárdenas (Universidad de Lima, Perú)
Mg. Jorge Luis Irey Núñez (Universidad de Lima, Perú)
Dr. Jose Martin Lozano Aparicio (Universidad Nacional de Ingeniería, Perú)
Mg. Cesar Stuardo Lucho Romero (Pontificia Universidad Católica del Perú)
Mg. William-Rogelio Marchand-Niño (Universidad Nacional Agraria de la Selva, Perú)
Mg. Edgar André Manzano Ramos (Pontificia Universidad Católica del Perú)
Mg. Pilar Alexandra Moreno (Universidad Nacional Abierta y a Distancia, Colombia)
Mg. Braulio Oscar Murillo Veliz (Pontificia Universidad Católica del Perú)
Mg. Luis Beltrán Palma Ttito (Universidad Nacional de San Antonio Abad del Cusco, Perú)
Mg. Iván Darío Peñaranda Arenas (Universidad de Granada, España)
Mg. Ferdinand Edgardo Pineda Ancco (Pontificia Universidad Católica del Perú)
Dr. José Antonio Pow Sang Portillo (Pontificia Universidad Católica del Perú)
Mg. Hernán Alejandro Quintana Cruz (Universidad de Lima, Perú)
Mg. Lennin Paul Quiroz Villalobos (Universidad de Lima, Perú)
Mg. Lourdes Ramírez Cerna (Universidade Federal de Ouro Preto, Brasil)
Mg. Liseth Urpy Segundo Carpio (Universidade de São Paulo, Brasil)
Dr. Ángel Leonardo Valdivieso Caraguay (Escuela Politécnica Nacional, Ecuador)
Dr. Ronny Villafuerte Serna (Universidad Nacional de San Antonio Abad del Cusco, Perú)
Mg. Rannoverng Yanac Montesino (Universidad Nacional Agraria de la Selva, Perú)
Ing. Juan José Martínez Cámara (Universidad de Jaén, España)
Mg. Edilene Cavalcanti dos Anjos (Universidade Federal de Santa Catarina, Brasil)
Dra. Valentina Gomes Haensel Schmitt (Universidad de Lima, Perú)
Mg. Oswaldo Daniel Casazola Cruz (Universidad Nacional del Callao, Perú)

POLÍTICA EDITORIAL

ENFOQUE Y ALCANCE

Interfases es una revista digital gestionada por la Universidad de Lima para la publicación de investigaciones originales en áreas temáticas relacionadas con las ciencias de la computación, gestión de datos, aprendizaje automático, sistemas de información, gestión de conocimiento y áreas afines. Se publican artículos científicos y avances de investigación, siempre que cumplan con el proceso de revisión por pares. La revista *Interfases* está indexada en CrossRef, Dialnet, Latindex y DOAJ: Directory of Open Access Journals, y se publica dos veces al año, la primera en julio y la segunda en diciembre. Sin embargo, a partir de julio del 2021, los manuscritos individuales se publicarán tan pronto como estén listos, añadiéndolos progresivamente al contenido de la edición en curso en la modalidad de publicación continua. Los artículos en publicación continua siguen el proceso de revisión por pares, y ya se pueden citar utilizando el año de publicación y el DOI.

PROCESO DE REVISIÓN POR PARES

Los manuscritos originales e inéditos enviados a la revista *Interfases* siguen un proceso de evaluación en dos etapas.

En la primera, el editor examina el contenido para determinar si el manuscrito está alineado con el alcance y ha seguido las directrices para autores. Si el manuscrito no es aceptado, se devuelve al autor correspondiente con las razones detalladas que motivan la decisión adoptada por el editor.

Si el manuscrito es aceptado por el editor, este se envía a revisores externos expertos en el tema de investigación. Esta segunda evaluación corresponde a una revisión por pares doble ciego, donde el autor y los revisores son anónimos.

El revisor evalúa el contenido del manuscrito y, basándose en su experiencia y conocimiento, adopta una de las siguientes recomendaciones:

1. El manuscrito es aceptado sin cambios o con cambios mínimos.
2. El manuscrito se acepta con la condición de realizar cambios importantes, de acuerdo con las observaciones del revisor. La versión corregida del manuscrito debe ser aprobada en una segunda revisión.
3. El manuscrito no se acepta por las contribuciones limitadas del estudio u otras consideraciones informadas por el revisor.

Con base en los comentarios de los revisores, el editor informa la decisión al autor correspondiente, quien tiene hasta treinta días para realizar los cambios al manuscrito (recomendaciones 1 y 2), o argumenta por qué no se acepta (recomendación 3).

Cuando los revisores reciben el manuscrito corregido, tienen hasta veinte días para informar el resultado de la nueva evaluación; posteriormente, emiten su recomendación final. Una vez que el editor recibe la segunda ronda de revisiones, toma una decisión para publicar el manuscrito y luego se le notifica al autor correspondiente.

Cualquier objeción del autor con respecto a la decisión del editor o hacia los comentarios de los revisores será resuelta por el Comité Editorial como instancia final.

La revista se adhiere a los criterios establecidos por el Guidelines on Good Publication Practice del Committee on Publication Ethics (COPE), el cual establece las sanciones en caso de plagio.

DIRECTRICES PARA AUTORES/AS

ENVÍO DEL MANUSCRITO

Interfases publica tres tipos de artículos: trabajos de investigación (hasta 5000 palabras), avances en investigación (hasta 2800 palabras) y revisiones (hasta 1500 palabras).

Todos los artículos se envían del mismo modo. Una vez que el editor verifique que el contenido del manuscrito pertenece al ámbito de *Interfases*, lo remitirá a un proceso de revisión por pares. Este proceso (compuesto de dos rondas) toma aproximadamente de dos a tres meses, pero dependiendo de la complejidad del manuscrito, podría extenderse.

Los manuscritos enviados a *Interfases* no deben haberse publicado previamente ni estar en consideración para su publicación en otra revista.

Página del título

La página del título debe incluir:

- Un título conciso e informativo (hasta 30 palabras).
- El nombre completo de cada autor, incluyendo la afiliación institucional, la dirección de correo electrónico y el código ORCID.
- Resumen de 200-250 palabras. El resumen debe indicar la naturaleza y contribución del estudio. Evite las abreviaturas no definidas, las ecuaciones matemáticas o las referencias bibliográficas en el texto del resumen.
- Palabras clave (de 3 a 5) separadas por comas. Las palabras clave deben tomarse de la taxonomía de la IEEE Computer Society: <https://www.computer.org/digital-library/journals/sc/tsc-taxonomy-list>

Texto

Los trabajos enviados deben haber sido redactados en un documento Word (.doc o .docx), y aquellos aceptados para ser publicados deben usar la plantilla de *Interfases* LATEX que estará disponible pronto.

Al redactar el manuscrito, usar la opción de numeración automática para numerar las páginas. Por favor, evite el uso de funciones de campo. Utilice la función de tabla, no una hoja de cálculo pegada, para hacer tablas. Si escribe su manuscrito con Word, use el editor de ecuaciones o MathType para las ecuaciones.

Tablas

Las tablas son el núcleo de los nuevos hallazgos reportados en la corriente principal de la ciencia; por lo tanto, incluya las tablas que considera que son estrictamente necesarias. Todas las tablas se numeran utilizando números arábigos (por ejemplo, Tabla 1, Tabla 2...) e incluyen un título que señala la relevancia de los datos presentados.

Las tablas se mencionan en el orden en que aparecen en el manuscrito. Además del número, el título y los datos, las tablas pueden incluir una nota para detallar la fuente de información, así como explicaciones adicionales que no están incluidas en el manuscrito.

Abreviaturas

Use abreviaturas solo si son necesarias para mejorar la legibilidad de su documento. Debe definir cada abreviatura en la primera mención y, después, usarla de manera consistente.

Conclusiones

Recuerde que las conclusiones no son la versión narrativa y textual de las tablas incluidas en la sección Resultados. Por el contrario, reseñan y sintetizan los principales argumentos del artículo. Las conclusiones se extraen de los hallazgos y proporcionan una respuesta adecuada a la pregunta de investigación. Además, incluyen las limitaciones del estudio y sugieren nuevas preguntas y aplicaciones para futuros estudios.

Referencias

Las citas y las referencias deberán indicarse de acuerdo con las normas APA. Según la norma señalada, las referencias, enlistadas al final de la publicación, se presentarán de la siguiente forma:

a. Libros

Apellido Autor, A. A., & Apellido Autor, A. A. (Año de la publicación). *Título del libro* (número de la edición). Nombre de la editorial. DOI o URL

b. Artículos de revistas o capítulos de un libro

Apellido del (los) autor(es), A. A. (Año de publicación). Título del artículo o el capítulo. Nombre de la revista o el libro (en cursiva), número de la revista (en cursiva), páginas en las que se encuentra el artículo o el capítulo.

c. Libros electrónicos:

Apellido del (los) autor(es), A. A. (Año de publicación). Título del texto electrónico (en cursiva). <http://...> (dirección web)

d. Artículos de revistas electrónicas:

Apellido del (los) autor(es), A. A. (Año de publicación). Título del artículo. Nombre de la revista (en cursiva), número de la revista (en cursiva), páginas en las que se encuentra el artículo. <http://...> (dirección web)

e. Ponencias en congresos o simposios:

Apellido del (los) expositor(es), A. A. (Año, [indicar día] de [indicar mes]). Título de la ponencia (en cursiva). Conferencia presentada en el [nombre del evento]. <http://...> (dirección web)

Material suplementario electrónico

Los autores pueden incluir archivos de texto (para tablas y figuras) y hojas de cálculo como material complementario. Sin embargo, para datos de investigación, es recomendable archivarlos en repositorios de datos. Para el código de *software*, los autores pueden usar plataformas como GitHub o similares.

Si los originales contienen fotografías o reproducciones de obras pictóricas, estas se entregarán aparte en archivos TIFF o JPG, con 300 píxeles de resolución (DPI). Si contienen gráficos, cuadros, dibujos, flujogramas u otros elementos, estos deben entregarse igualmente en un archivo aparte y en el programa original en que fueron creados (por ejemplo: Excel, Illustrator, etcétera).

Lista preliminar para la preparación de envíos

Los artículos deberán respetar el siguiente formato:

- a. Formato de página A4.
- b. Título del artículo, centrado, en negrita, con letra Times New Roman en doce puntos.
- c. Títulos del texto, en negrita, con letra Times New Roman en doce puntos, dejando dos líneas en blanco antes del párrafo.
- d. Texto del cuerpo con letra Times New Roman en doce puntos, con espacio y medio de interlineado.

Declaración de privacidad

Los nombres y las direcciones de correo electrónico introducidos en esta revista se usarán exclusivamente para los fines establecidos en ella y no se proporcionarán a terceros o para su uso con otros fines.

ÍNDICE

PRESENTACIÓN DE LA SECCIÓN DOSIER	15
Gestión del conocimiento y técnicas de inteligencia artificial <i>Guillermo Antonio Dávila Calle</i> <i>Ruth María Reátegui Rojas</i>	
PRESENTACIÓN DE LA SECCIÓN GENERAL	17
<i>Julio Alejandro Padilla-Solís</i>	
DOSIER	19
Retos y oportunidades en la accesibilidad de datos	20
<i>Mónica Calva</i>	
Modelamiento de tópicos utilizando mensajes de Twitter relacionados con el cáncer cervical	40
<i>Ruth María Reátegui Rojas</i>	
O papel da capacidade absorptiva na inovação: evidências de práticas desenvolvidas por empresas de base tecnológica inseridas em redes colaborativas	52
<i>Graciele Tonial</i> <i>Nathalia Berger Werlang</i> <i>Alessandra Cassol</i>	
Capital humano: a chave da gestão do conhecimento em organizações públicas - Uma abordagem integrativa	72
<i>Danisson Luiz dos Santos Reis</i> <i>Paulo Roberto Moura</i> <i>Gregório Varvakis</i> <i>Neri dos Santos</i>	

ARTÍCULOS DE INVESTIGACIÓN	94
Reforzamiento de habilidades aritméticas de fracciones y divisiones en niños de primaria mediante un juego serio educativo	9
<i>Renato Coronado Álvarez</i>	
Sistema de monitoreo de ambiente para el control de las condiciones ambientales de las granjas acuícolas de la selva del Perú basado en internet de las cosas	138
<i>Adrian Ernesto Lino Villaran</i>	
Protocolo básico de ciberseguridad para pymes	167
<i>Olda Bustillos Ortega</i>	
<i>Javier Rojas Segura</i>	
Aplicación de <i>machine learning</i> para campañas de <i>marketing</i> en la banca comercial	185
<i>Ganímedes T. Rosales Reyes</i>	
<i>Xavier Alberto Gutierrez Coral</i>	
<i>Augusto Enrique Hayashida Marchinares</i>	
ARTÍCULOS DE REVISIÓN	199
Diseño de interfaces de sistemas interactivos utilizando técnicas de <i>machine learning</i> : una revisión del diseño y la usabilidad	200
<i>Julio Vladimir Quispe Sota</i>	
Dispositivos <i>wearables</i> y los riesgos a la privacidad: una revisión de la literatura	213
<i>Angelo Rodrigo Taco Jimenez</i>	
Una revisión de las implementaciones de sistemas para la identificación de tendencias de la diabetes	229
<i>Rocio Isabel Benites Loja</i>	
<i>Marco Antonio Coral Ygnacio</i>	
ARTÍCULO DE DIVULGACIÓN	250
Formulación de requisitos para un sistema de información de apoyo al proceso de inducción de estudiantes universitarios: caso UNAD	251
<i>Pilar Alexandra Moreno</i>	
<i>Juan Olegario Monroy Vásquez</i>	
DATOS DE LOS AUTORES	276

PRESENTACIÓN DE LA SECCIÓN DOSIER

Gestión del conocimiento y técnicas de inteligencia artificial

doi: <https://doi.org/10.26439/interfases2022.n016.6172>

En años recientes, el conocimiento se ha constituido en el principal recurso estratégico (Dávila & Dos Anjos, 2021). En consecuencia, gestionarlo adecuadamente ha contribuido al desarrollo de las sociedades (Angelidou, 2015), la competitividad de las organizaciones (Dávila et al., 2019) y el bienestar de los individuos (Kianto et al., 2016). El desafío de la gestión del conocimiento es integrar el conocimiento individual especializado y diverso en los bienes y productos finales de una organización o territorio (Grant, 1996).

Gracias a su carácter multidisciplinario, la gestión del conocimiento ha incorporado y adaptado teorías y herramientas de diversas áreas, tales como ciencias sociales aplicadas, ciencias humanas e ingenierías, entre otros. En las ciencias de la computación, la inteligencia artificial (IA), por su naturaleza multidisciplinaria, tiene como esencia el conocimiento. Por tanto, técnicas de procesamiento de lenguaje natural, representación del conocimiento y visión computacional han sido utilizadas para soportar procesos de adquisición y captura de conocimiento. Además, procesos de distribución del conocimiento son soportados por técnicas como minería de datos, descubierta de conocimiento y agentes inteligentes (Liebowitz, 2001). Procesos de combinación de conocimiento tácito y de creación de conocimiento —fundamentalmente predictivo— usan frecuentemente *big data* o algoritmos de *machine learning* (Sumbal et al., 2017).

Este apartado especial está abierto a artículos científicos (de revisión, estudios de caso y cuantitativos) que aborden el uso de herramientas de IA y sus diversas ramas para gestionar conocimiento individual, organizacional o territorial. Sin embargo, serán especialmente valorados artículos empíricos orientados a mejorar la productividad organizacional, o tópicos críticos para las sociedades en tiempos actuales, tales como sustentabilidad, salud, educación, movilidad, etcétera.

En primer lugar, nos ubicamos en la sección de Gestión del conocimiento y técnicas de inteligencia artificial, donde destaca el aporte de Mónica Calva de la Universidad Técnica Particular de Loja, en Ecuador, quien realiza una revisión sistemática acerca de la importancia e impacto del acceso a datos de salud. En segundo lugar, Ruth María Reátegui Rojas, también de la Universidad Técnica Particular de Loja, utiliza en su investigación el modelado de tópicos como herramienta para identificar tuits relacionados con el cáncer cervical. En tercer lugar, las autoras de Brasil, Graciele Tonial, Nathalia Berger

Werlang y Alessandra Cassol, presentan un estudio enfocado en pymes brasileñas y sus prácticas para la mejora de la gestión del conocimiento y la facilitación de generación de innovación. Finalmente, para cerrar la sección, Danisson Luiz dos Santos Reis, Paulo Roberto de Moura, Gregório Varvakis y Neri dos Santos, también de Brasil, elaboran un artículo cuyo propósito es la contribución a cómo la investigación académica aborda la relación entre el capital humano, la gestión del conocimiento y el sector público.

Guillermo Antonio Dávila Calle
Universidad de Lima

Ruth María Reátegui Rojas
Universidad Técnica Particular de Loja

REFERENCIAS

- Angelidou, M. (2015). Smart cities: A conjuncture of four forces. *Cities*, 47, 95-106.
- Dávila, G. A., Andreeva, T., & Varvakis, G. (2019). Knowledge management in Brazil: What governance mechanisms are needed to boost innovation? *Management and Organization Review*, 15(4), 857-886.
- Dávila, G. A., & Dos Anjos, E. C. (2021). Configurations of knowledge management practices, innovation, and performance: Exploring firms from Brazil. *International Journal of Innovation Management*, 25(06), 2150065.
- Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal*, 17(S2), 109-122.
- Kianto, A., Vanhala, M., & Heilmann, P. (2016). The impact of knowledge management on job satisfaction. *Journal of Knowledge Management*, 20(4), 621-636.
- Liebowitz, J. (2001). Knowledge management and its link to artificial intelligence. *Expert Systems with Applications*, 20(1), 1-6.
- Sumbal, M. S., Tsui, E., & See-to, E. W. (2017). Interrelationship between big data and knowledge management: An exploratory study in the oil and gas sector. *Journal of Knowledge Management*, 21(1), 180-196.

PRESENTACIÓN DE LA SECCIÓN GENERAL

doi: <https://doi.org/10.26439/interfases2022.n016.6173>

Uno de los objetivos de la revista *Interfases* es promocionar la comunicación científica en áreas temáticas relacionadas con ingeniería de sistemas, ciencias de la computación, ingeniería del *software*, sistemas de información, tecnologías de información, ciberseguridad, ciencia de datos y áreas afines. En ese sentido, presentamos la edición número 16, correspondiente al periodo agosto-diciembre del año 2022, formada por diez artículos que lograron pasar por un riguroso proceso de revisión por pares. Los autores proceden de Colombia, Ecuador, Perú y Brasil.

En primer lugar, tenemos el artículo de investigación de Renato Coronado Álvarez de la Universidad de Lima, que hace una reflexión sobre la dificultad de aprender matemáticas para los niños y cómo la tecnología puede ayudar a resolver este problema. Para ello, propone el uso de un juego serio educativo para el cual empleó heurísticas para el diseño de interfaces, y para la mecánica, aplicó el *framework* MDA. En segundo lugar, Adrian Ernesto Lino Villaran, también de la Universidad de Lima, explora el sector de la acuicultura en el Perú y aporta un sistema de monitoreo de ambiente semiautónomo para la zona oriental del país, el cual demostró tener la capacidad de medición de condiciones ambientales de manera precisa y autónoma. En tercer lugar, Olda Bustillos Ortega y Javier Rojas Segura de la Universidad Nacional de las Américas, en Costa Rica, hacen una investigación orientada a la ciberseguridad de las pymes; su principal objetivo es brindar un protocolo básico de ciberseguridad como defensa frente a un ataque de terceros y asegurar la continuidad del negocio. En cuarto lugar, el artículo de Ganímedes T. Rosales Reyes, Xavier Alberto Gutierrez Coral y Augusto Enrique Hayashida Marchinares de la Universidad Nacional Mayor de San Marcos está orientado a la aplicación de *machine learning* en *telemarketing*, y compara cuatro modelos. Luego, despliega un modelo basado en el modelo elegido y, finalmente, recomienda una hoja de ruta para la arquitectura que lo soporte. En quinto lugar, presentamos el trabajo de revisión sistemática sobre los enfoques planteados para el diseño de interfaces de usuario utilizando técnicas de

machine learning, por Julio Vladimir Quispe Sota, catedrático de la Universidad Nacional de San Antonio Abad del Cusco. En sexto lugar, se encuentra el artículo de revisión de literatura sobre los dispositivos *wearables*, cuáles son sus usos y principalmente a qué riesgos están expuestos los usuarios respecto a su privacidad y seguridad, desarrollado por el autor Angelo Rodrigo Taco Jimenez de la Universidad de Lima. En séptimo lugar, los autores de la Universidad Católica Sedes Sapientiae, Rocio Isabel Benites Loja y Marco Antonio Coral Ygnacio, hacen una revisión sistemática sobre los métodos y la problemática en la construcción de sistemas de identificación de tendencias en la diabetes. En octavo lugar, el artículo de divulgación desarrollado por Pilar Alexandra Moreno y Juan Olegario Monroy Vásquez, ambos docentes de la Universidad Nacional Abierta y a Distancia, en Colombia, reconoce la importancia de que los ingresantes cuenten con un sistema y herramientas de inducción que les sirvan de apoyo para su ingreso, adaptación y transición de la formación presencial a la virtual.

Finalmente, expresamos nuestro enorme reconocimiento a todos los investigadores que sometieron a revisión sus manuscritos en la presente edición de *Interfases*. En ese sentido, valoramos profundamente su esfuerzo y dedicación al desarrollo científico, a cuya difusión convoca en esta oportunidad la revista *Interfases*. También, queremos agradecer a nuestro equipo editorial y de revisores, quienes constantemente promueven la mejora continua de la revista asegurando la calidad del material publicado.

Julio Alejandro Padilla-Solís
Director de la revista *Interfases*

DOSIER

RETOS Y OPORTUNIDADES EN LA ACCESIBILIDAD DE DATOS

MÓNICA CALVA
mmcalva@utpl.edu.ec
<https://orcid.org/0000-0002-7279-3842>
Hospital UTPL, Ecuador

RESUMEN. En el área de la salud, existe una gran variedad y cantidad de información sensible y valiosa que no es explotada en su totalidad en beneficio de las personas, debido a la falta de acceso a los datos. Varios factores influyen en la accesibilidad de los datos en organizaciones privadas o públicas, y para determinarlos se ha hecho una revisión bibliográfica; de los 300 artículos consultados, 23 se incluyeron en esta revisión. Entre los factores resultantes que influyen en la accesibilidad de la información, destacan gobiernos abiertos, sistemas de información de salud, leyes de protección de datos personales, ética en el uso de la información médica y lagos de datos. Facilitar la accesibilidad a los datos de salud traería una mejora de servicios y planes de tratamiento, ahorro de recursos económicos estatales, e impulsaría la educación y la investigación.

PALABRAS CLAVE: accesibilidad, datos, salud, protección, ética

CHALLENGES AND OPPORTUNITIES IN DATA ACCESSIBILITY

ABSTRACT. There is a great variety and amount of sensitive and valuable healthcare information that could be used to benefit people, but is not exploited due to limited access to data. This article aims to determine factors influencing data accessibility in private and public organizations through a literature review. Of the 300 related articles identified, 23 were included in this review. Among the resulting factors influencing information accessibility, the following stood out: open governments, health information systems, personal data protection laws, ethics in the use of medical information, and data lakes. Making health data more accessible would improve services and treatment plans, save state resources and promote education and research.

KEYWORDS: accessibility, data, health, protection, ethics

1. INTRODUCCIÓN

Una gran cantidad de datos se genera a través de aplicaciones, ya sea directamente por usuarios o por organizaciones. Muchos de ellos no son aprovechados en su totalidad. En el ámbito de la salud, hay un sinnúmero de fuentes de información de pacientes, por ejemplo, aplicaciones de salud, relojes inteligentes, dispositivos médicos, de monitoreo, signos vitales, equipos de imagen, sistemas de registros clínicos electrónicos, etcétera. Pero si no están enlazados a un sistema de información, no aportan valor o beneficio a las organizaciones. Frecuentemente, las técnicas y prácticas para manipulación de datos están enfocadas en datos estructurados, por lo que se pierde una extensa cantidad de información no estructurada de notas clínicas de texto libre, que puede ser de gran ayuda en el ámbito de la investigación (Jones et al., 2020) y la educación.

Como consecuencia de la necesidad de manipular estas grandes cantidades de datos, surgen iniciativas de marcos de trabajo sobre la creación de almacenes de datos: lagos de datos. Los lagos de datos son un concepto nuevo que ha cobrado relevancia en los últimos años, cuyo objetivo es ser un depósito centralizado, seleccionado y seguro que almacena todos los datos de las organizaciones, tanto en su forma original como en la generada a partir de análisis (Bauer et al., 2021), y con un costo reducido, a diferencia de los almacenes de datos.

Sin embargo, existen preocupaciones sobre el uso ético y legal que pueda darse a los datos recolectados. Por ello, deben analizarse las regulaciones legales para la protección de los datos personales y su privacidad, así como los enfoques éticos usados en proyectos de investigación con este tipo de datos.

Por otro lado, los gobiernos abiertos son iniciativas que han puesto sobre la mesa la discusión acerca de la protección de datos, la accesibilidad y las regulaciones para proteger los derechos de las personas. También han comenzado el debate sobre la confianza de los titulares de los datos en el uso que se les dé a estos. Los gobiernos son los entes promotores para que se establezcan políticas claras para proteger los derechos fundamentales y para generar un valor a los Estados a partir de estos datos.

En esta revisión bibliográfica se plantea identificar cuáles son los factores que influyen en la accesibilidad a los datos. En la segunda sección se describen algunos de estos factores y sus implicaciones en el área de la salud. En la tercera sección se presentan sugerencias para investigadores y trabajos futuros; y, finalmente, en la cuarta sección, las conclusiones.

2. FACTORES QUE INFLUYEN EN LA ACCESIBILIDAD A LOS DATOS

Los artículos más relevantes que se han incluido en la revisión bibliográfica se detallan en la Tabla 1.

Tabla 1

Principales artículos de la revisión bibliográfica

Autores	Artículo	Objetivo/propósito	Resultados
Alsahafi, Y. A., Gay, V., & Khwaji, A. A.	Factors Affecting the Acceptance of Integrated Electronic Personal Health Records in Saudi Arabia: The Impact of e-Health Literacy	Identificar los factores que pueden influir en la aceptación y el uso de un sistema integrado de registros médicos personales en Arabia Saudita.	Bajo nivel de conocimiento entre los ciudadanos saudíes sobre la implantación nacional de un sistema integrado de registros médicos personales y su utilidad.
Avila-Garzon, C.	Applications, Methodologies, and Technologies for Linked Open Data: A Systematic Literature Review	Realizar una revisión bibliográfica sistemática sobre datos abiertos enlazados (LOD).	La mayoría de los estudios realizados se centran en el uso de tecnologías y herramientas de la web semántica aplicadas a contextos como la biología, las ciencias sociales, las bibliotecas, la investigación y la educación; pero falta investigación sobre una metodología estandarizada para la gestión de LOD. Se pueden utilizar muchas herramientas para la gestión de LOD, pero la mayoría de ellas carecen de interfaces fáciles de usar para la consulta de conjuntos de datos.
Bauer, D., Froese, F., Garcés-Erice, L., Giblin, C., Labbi, A., Nagy, Z. A., Pardon, N., Rooney, S., Urbanetz, P., Vetsch, P., & Wespi, A.	Building and Operating a Large-Scale Enterprise Data Analytics Platform	Ejecución de una plataforma de procesamiento de datos a gran escala para aplicar la analítica a los datos corporativos a escala en una instancia de nube privada de OpenStack.	Lecciones aprendidas de la construcción y el funcionamiento del <i>datalake</i> empresarial: automatización, familiaridad de las herramientas, desagregación, uso compartido de recursos, cumplimiento continuo, disponibilidad de datos.

(continúa)

(continuación)

Autores	Artículo	Objetivo/propósito	Resultados
Boudreau LeBlanc, A., Williams-Jones, B., & Aenishaenslin, C.	Bio-Ethics and One Health: A Case Study Approach to Building Reflexive Governance. <i>Frontiers in Public Health</i>	Sentar las bases de un marco metodológico en bioética empírica, cuyos rudimentos se han aplicado a un estudio de caso para construir la gobernanza reflexiva en One Health. Empoderar a los individuos y a las asociaciones en un diálogo con la sociedad, operación que se lleva a cabo mediante un enfoque de estudio de casos sobre sistemas de intercambio de datos.	Un proceso que "reensambla" el paradigma de One Health desde la perspectiva de la bioética global para crear puentes entre la persona y el ecosistema a través de la ética pragmática.
Buhr, L., Schicktanz, S., & Nordmeyer, E.	Attitudes Toward Mobile Apps for Pandemic Research Among Smartphone Users in Germany: National Survey	Explorar el potencial de aceptación de las aplicaciones orientadas a la investigación (ROA) en la población alemana.	Mayor confianza en las organizaciones gubernamentales (gobierno federal o estatal, oficina regional de salud), públicas (seguro de salud obligatorio) o financiadas por el gobierno (institutos de investigación) para el almacenamiento de datos y la distribución de la aplicación, frente a las organizaciones privadas (instituciones de investigación privadas, clínicas, seguros de salud, empresas de tecnología de la información [TI]). Tener un título universitario aumentaba significativamente la probabilidad de utilizar una aplicación para la pandemia, mientras que tener un origen migratorio la disminuía significativamente. La mayoría (653/778, 83,9 %) de los usuarios de teléfonos inteligentes estaban dispuestos a proporcionar los datos de sus aplicaciones para la investigación financiada por el Estado.

(cont.)

(continuación)

Autores	Artículo	Objetivo/propósito	Resultados
Calva, M., & Piedra, N.	Cycle of Transformation of Open and Linked Medical Data to Improve the Training of Health Professionals	Propuesta de un ciclo de transformación de datos médicos a datos abiertos enlazados anónimos.	Impulsar mejoras para que los entornos educativos de datos abiertos enlazados contribuyan a optimizar las capacidades de los profesionales de salud. Esto implica mejoras en interfaces de usuario gráficas utilizables, que sean fáciles de modificar y de configurar, con datos de pacientes que sirvan para realizar simulaciones clínicas realistas con fines de entrenamiento y evaluación. Apoyo del profesorado, capacitación y soporte financiero. Evaluación de satisfacción del usuario, tanto del profesorado como de los estudiantes, en estos entornos.
Cohen, G., Gerke, S., & Kramer, D. B.	Ethical and Legal Implications of Remote Monitoring of Medical Devices	Analizar las implicaciones legales de las normativas de privacidad de Estados Unidos y Europa para el acceso de los pacientes a los datos resumidos o brutos de los dispositivos.	Los pacientes estadounidenses pueden tener poco acceso a sus datos brutos recopilados y conservados por los fabricantes de dispositivos en Estados Unidos, en virtud de la Regla de Privacidad de la Ley de Portabilidad y Responsabilidad de los Seguros Médicos, mientras que el Reglamento General de Protección de Datos (RGPD) de la Unión Europea concede un mayor acceso a los datos recopilados por los dispositivos cuando el procesamiento de los datos personales entra en el ámbito territorial del RGPD.
Comandè, G., & Schneider, G.	Can the GDPR make data flow for research easier? Yes it can, by differentiating! A careful reading of the GDPR shows how EU data protection law leaves open some significant flexibilities	Demostrar que el RGPD puede funcionar como facilitador de la investigación.	Los pilares reguladores de la ley europea de protección de datos, el primero relacionado con la protección del derecho fundamental a la protección de datos y el segundo relativo a la promoción de la libre circulación de datos personales, dan lugar a una arquitectura de regímenes de protección de datos por capas,

(continúa)

(continuación)

Autores	Artículo	Objetivo/propósito	Resultados
	for data protection-sound research activities		que refuerza o relaja los derechos de los interesados y las salvaguardias de protección de datos con respecto a las actividades de tratamiento basadas en intereses públicos o meramente económicos. Cada uno de los regímenes de protección de datos identificados da forma a distintos "puntos normativos habilitantes" para el tratamiento de datos personales sensibles con fines de investigación.
Dunbar, E., Olsen, H. E., Salomon, E., Bhatt, S., Mutuku, R., Wasunna, B., Edwards, J., Kolko, B., & Holeman, I.	Towards Responsible Data Practices in Digital Health: A Case Study of an Open Source Community's Journey	Plantear una agenda de investigación de diseño que surgió de las necesidades prácticas de una comunidad de código abierto centrada en el <i>software</i> de salud digital para la salud comunitaria en países de ingresos bajos y medios.	Resultados de las etapas formativas de su proceso de diseño en varias fases, que incluyen una revisión bibliográfica de alcance y una reformulación de las políticas y procedimientos institucionales. Una de las principales aportaciones de este estudio de caso es que ofrece un ejemplo del "trabajo previo" institucional necesario para dar sentido al complejo panorama de la protección de datos y trazar un camino para diseñar un <i>software</i> que apoye mejor las prácticas de datos responsables. Este estudio de caso también articula el importante papel de los diseñadores y ejecutores de la sanidad digital en la puesta en práctica de los derechos de los datos de los pacientes
Eder, J., & Shekhovtsov, V. A	Data Quality for Federated Medical Data Lakes	Proponer una arquitectura que ayude a los investigadores a identificar de forma eficiente y eficaz las colecciones de material y datos relevantes con calidad documentada	Este estudio describe el panorama de los biobancos como lagos de datos médicos federados, por ejemplo, las colecciones de muestras y sus anotaciones en la federación europea de biobancos (Biobanking and Biomolecular Resources Research Infrastructure - European Research Infrastructure Consortium, BBMRI-ERIC),

(continúa)

(continuación)

Autores	Artículo	Objetivo/propósito	Resultados
.		para sus proyectos de investigación, respetando al mismo tiempo las estrictas normas de privacidad.	y desarrolla un modelo conceptual que captura la información de los esquemas con anotaciones de calidad. Este artículo analiza en profundidad las dimensiones de la calidad de los conjuntos de datos para la investigación médica y propone representaciones tanto de los metadatos como de la documentación de la calidad de los datos con el objetivo de ayudar a los investigadores a identificar de forma eficaz y eficiente los conjuntos de datos adecuados para los estudios médicos.
Galeone, C., & Bonzi, R	Open Data: Applications in Social Care and Health	Ofrecer una visión general de los datos abiertos y su aplicación en el ámbito social y sanitario. Proporcionar una definición de datos abiertos y discutir el estado actual de la situación en el mundo. Por último, se centra en la situación italiana en relación con los datos sanitarios abiertos.	Se ha producido un aumento considerable de la cantidad de datos abiertos, que están cada vez más disponibles debido al elevado número de portales web y conjuntos de datos gratuitos. Sin embargo, no son tantos los conjuntos de datos totalmente abiertos. Existen muchos obstáculos que afectan negativamente a la difusión de los datos abiertos, como la actualización inadecuada y poco oportuna de los datos; la posibilidad de que la estructura de los datos cambie cuando se publica una actualización; la dificultad de encontrar una codificación única y estándar, con la consiguiente actividad de cotejo costosa y no siempre factible (Bacon & Goldacre, 2020, como se citó en Galeone & Bonzi, 2020). La escasa disponibilidad de datos completamente abiertos afecta en particular a los datos sanitarios públicos de todo el mundo, sin distinciones relacionadas con el nivel de desarrollo económico y tecnológico del país (World Wide Web Foundation, 2018, como se citó en Galeone & Bonzi, 2020).

(continúa)

(continuación)

Autores	Artículo	Objetivo/propósito	Resultados
Gefenas, E., Lekstutiene, J., Lukaseviciene, V., Hartlev, M., Mourby, M., & Cathaoir, K.	Controversies between Regulations of Research Ethics and Protection of Personal Data: Informed Consent at a Cross-Road	Explorar las siguientes controversias: ¿debe considerarse el consentimiento como base jurídica prioritaria para el tratamiento de datos personales en la investigación de datos sanitarios? ¿Puede el consentimiento amplio seguir siendo un fundamento jurídico adecuado para el biobanco? ¿Cuál debería ser el papel de las disposiciones de ética de la investigación que difieren de las normas del RGPD, y cuál debería ser el papel y la función de los comités de ética de la investigación en el entorno cambiante de la investigación de datos sanitarios?	Necesidad de cooperación entre los comités de ética y los organismos de protección de datos. El consentimiento desempeña un papel importante en el mantenimiento de la participación pública en las iniciativas de intercambio de datos. Aunque todavía hay debates en curso sobre cuál sería la mejor opción para el intercambio de datos: el uso del consentimiento dinámico, el consentimiento amplio o la donación de datos sin consentimiento, parece que las preferencias del público por un cierto nivel de control sobre sus datos, independientemente de la base legal en la que se basa el RGPD, apoyan un papel continuo del consentimiento para mantener la confianza en la investigación, muestra el respeto a los participantes en el estudio, y es una necesidad por defecto bajo las directrices de ética de la investigación.
Healthcare Information and Management Systems Society	2021 HIMSS Healthcare Cybersecurity Survey	Ofrecer una visión del estado de la ciberseguridad en la sanidad basada en las opiniones de 167 profesionales de la ciberseguridad sanitaria. Las organizaciones sanitarias se enfrentan a un sinfín de retos, como presupuestos ajustados, infraestructuras	Los resultados de la Encuesta de Ciberseguridad Sanitaria 2021 de HIMSS sugieren que las organizaciones sanitarias todavía tienen importantes retos que superar. Estos obstáculos al progreso incluyen presupuestos de seguridad ajustados, huellas heredadas cada vez mayores y el creciente volumen de ciberataques y compromisos. Además, los controles de seguridad básicos no se han implementado completamente en muchas organizaciones. Pero quizá la mayor vulnerabilidad

(continúa)

(continuación)

Autores	Artículo	Objetivo/propósito	Resultados
		envejecidas y el aumento de los ataques de ingeniería social y <i>ransomware</i> .	sea el factor humano. Las organizaciones sanitarias deberían hacer más para apoyar a los profesionales de la ciberseguridad sanitaria y sus programas de ciberseguridad.
Jones, K. H., Ford, E. M., Lea, N., Griffiths, L. J., Hassan, L., Heys, S., Squires, E., & Nenadic, G.	Toward the Development of Data Governance Standards for Using Clinical Free-Text Data in Health Research: Position Paper	Elaborar recomendaciones para la creación de normas de gobernanza de datos que se integren en los marcos existentes para el uso de datos personales, a fin de permitir que los datos de texto libre se utilicen de forma segura para la investigación en beneficio de los pacientes y del público.	Guía autorizada sobre la gobernanza de los datos para la reutilización de los datos de texto libre, a fin de garantizar la transparencia pública en los flujos y usos de los datos, para tratar los datos de texto libre desidentificados como potencialmente identificables con un uso limitado a los refugios de datos acreditados. Mejora continua para comprender las relaciones entre la eficacia de la desidentificación y los riesgos de reidentificación, y comunicar a todas las partes interesadas.
Kandasamy, K., Srinivas, S., Achuthan, K., & Rangan, V. P.	Digital Healthcare -Cyberattacks in Asian Organizations: An Analysis of Vulnerabilities, Risks, NIST Perspectives, and Recommendations	Investigar los recientes ciberataques en instituciones sanitarias asiáticas.	Se ha descubierto que hay cinco tipos de ciberataques que dominan las instituciones sanitarias asiáticas. Como parte de este estudio se realiza un análisis detallado de estos ataques, sus vulnerabilidades y los riesgos asociados. Los ciberataques a instituciones sanitarias asiáticas se analizan a través del lente del marco de riesgos del Instituto Nacional de Normas y Tecnología (NIST). Finalmente, se incluyen algunas recomendaciones para ser utilizadas como medidas de mitigación de riesgos por las organizaciones sanitarias asiáticas y los investigadores para gestionar y mejorar la creciente situación de los ciberataques.

(continúa)

(continuación)

Autores	Artículo	Objetivo/propósito	Resultados
Keefer, P., & Scartascini, C.	Confianza: la clave de la cohesión social y el crecimiento en América Latina y el Caribe	Analizar la confianza en varios niveles sociales y cómo afecta al crecimiento en América Latina y el Caribe.	Diferentes medidas y relaciones de la confianza con economía, civismo, políticas públicas, transformación digital, sociedad, etcétera. Recomendaciones.
Red Iberoamericana de Protección de Datos	Recomendaciones generales para el tratamiento de datos en la inteligencia artificial	Recomendaciones para los desarrolladores de productos de IA, con el fin de orientarlos para que desde el diseño del producto se tenga en cuenta las exigencias de las regulaciones sobre tratamiento de datos personales.	Se establecen diez recomendaciones para desarrolladores de productos de IA.
Rooney, S., Bauer, D., Garcés-Erice, L., Urbanetz, P., Froese, F., & Tomic, S.	Experiences with Managing Data Ingestion into a Corporate Datalake	Describir la gestión de datos, transferencia e ingesta de datos en la plataforma de <i>datalake</i> .	Explicar la experiencia en el diseño, construcción y gestión de un gran <i>datalake</i> corporativo.
Shah, S. M., & Khan, R. A.	Secondary Use of Electronic Health Record: Opportunities and Challenges	Examinar sistemáticamente varios usos secundarios de la historia clínica electrónica con el objetivo de destacar cómo afectan a la privacidad de los pacientes. Analizar de forma crítica las normativas del RGPD y la HIPAA, y destacar sus posibles áreas de mejora, teniendo en cuenta el creciente uso de la tecnología y los diferentes usos secundarios de la historia clínica electrónica.	Se describen varios usos secundarios de la historia clínica electrónica (HCE), como la investigación clínica, la vigilancia de la salud pública y las auditorías clínicas para proporcionar servicios de atención sanitaria eficaces, oportunos y de calidad a los pacientes. También se han descrito los desafíos asociados a los usos secundarios de la HCE. Además, se analizan en detalle diferentes cuestiones éticas y de privacidad derivadas de la filtración de datos de HCE. Riesgos para la seguridad de los datos y la privacidad de los pacientes relacionados con los usos secundarios de la HCE, especialmente cuando sus datos se almacenan en la nube, se transmiten a través de la red y se comparten e intercambian con

(continúa)

(continuación)

Autores	Artículo	Objetivo/propósito	Resultados
			múltiples partes interesadas. Adicionalmente, se estudian importantes normativas de privacidad, como el RGPD y la HIPAA, desde la perspectiva del uso secundario de la HCE. Se señalan las posibles áreas de mejora en estas normativas para hacerlas más eficaces en la protección de la privacidad y la seguridad de los datos, y para hacerlas robustas frente a las crecientes técnicas asistidas por la IA en el análisis de datos y los ciberataques.

De esta revisión se han obtenido los siguientes aspectos que constituyen un reto u oportunidad para mejorar la accesibilidad a los datos médicos en instituciones públicas o privadas.

2.1 Ley de protección de datos

Las leyes y políticas de protección de datos proporcionan estándares para regular el procesamiento de datos que protegen los derechos de las personas y establecen obligaciones para los responsables y custodios de los datos (Cohen et al., 2020). Estas leyes recalcan cada vez más que las personas tienen derecho a no estar sujetas a decisiones guiadas únicamente por procesos automatizados por todas las implicaciones éticas que esto conlleva.

Estados Unidos ha liderado la promulgación de normativa para protección de datos personales, pues en 1996 publicó la Ley de Portabilidad y Responsabilidad de los Seguros de Salud (HIPAA, por sus siglas en inglés) para normar el acceso y privacidad de los datos de salud de los pacientes. Sin embargo, de acuerdo con Cohen et al. (2020), la mayoría de los pacientes de este país pueden tener poco acceso a sus propios datos digitales de salud, bajo esta ley.

Otra regulación vigente desde el 2020 en Estados Unidos en el estado de California es la Ley de Privacidad del Consumidor de California (CCPA), que propone medidas para proteger mejor la privacidad de los datos de los consumidores. Esta ley tiene un amplio alcance y cubre algunos de los datos que no están considerados por la HIPAA. Pero la CCPA no se aplica a la información de salud protegida que recopilan las entidades cubiertas o sus socios comerciales bajo la HIPAA (Cohen et al., 2020).

Por otro lado, el Reglamento General de Protección de Datos (RGPD) de la Unión Europea, que entró en vigor en el 2018, busca proteger los datos personales y la forma en que las organizaciones los procesan, almacenan y, finalmente, destruyen, cuando ya no son requeridos. Algunos países de la región, como Ecuador, que no disponían de leyes de protección de datos, lograron aprobar en el año 2021 una ley de protección de datos basada en el RGPD de la Unión Europea.

Si bien la HIPPA proporciona acceso a los registros clínicos del paciente, las leyes que otorgan mayor acceso a los datos recopilados son el RGPD de la Unión Europea y la CCPA en Estados Unidos (Cohen et al., 2020).

El RGPD proporciona una regulación altamente sofisticada de las actividades de procesamiento de datos con fines de investigación, según la naturaleza de las actividades de investigación realizadas. Esto fomenta la confianza en los titulares de datos y genera un alto nivel de protección de los derechos fundamentales de los interesados (Comandè & Schneider, 2021). También impulsa la puesta en marcha de proyectos de ciencia ciudadana, que permiten incorporar al ciudadano común en proyectos de investigación, los cuales deben contar con una normativa específica para la protección de datos personales (Pierce & Evram, 2022).

En ámbitos como la informática, hay iniciativas de regulaciones diseñadas específicamente para administrar el uso de datos personales, como la inteligencia artificial. Por ejemplo, la Red Iberoamericana de Protección de Datos, que consta de 22 autoridades de protección de datos en Portugal y España, y en México y otros países de Centroamérica, Sudamérica y el Caribe, ha emitido un reglamento que norma el acceso y uso de los datos en inteligencia artificial (IA) denominado *Recomendaciones generales para el tratamiento de datos personales en inteligencia artificial* (2019), que tiene como objetivo principal evitar amenazas o vulneraciones de los derechos humanos de los titulares de los datos.

2.2 Sistemas de información de salud

Los sistemas de registros médicos electrónicos de las instituciones de salud (EHR) proporcionan beneficios a los pacientes, porque son capaces de almacenar y gestionar historiales completos de salud de cada paciente con información de diferentes áreas, como imagen, laboratorio, farmacia, tratamientos clínicos, cirugías y demás procedimientos complementarios, además de información social, financiera y de seguros. Esta centralización de la información permite mejorar la calidad de atención y toma de decisiones, ya que reduce los errores en el análisis de datos médicos. Igualmente, facilita la incorporación de herramientas para la toma de decisiones, como los sistemas de soporte de decisiones (DSS); y hace posible el acceso a los datos desde diferentes ubicaciones y plataformas (Shah & Khan, 2020).

Pero existen riesgos como la seguridad y privacidad de los datos de los pacientes. De acuerdo con la encuesta de ciberseguridad en el área de salud realizada por Healthcare Information and Management Systems Society (HIMSS) en el 2021, el 65 % de los encuestados reportaron incidentes de seguridad significativos en sus organizaciones. Ataques de *phishing* y *ransomware* fueron los más típicos y frecuentes (HIMSS, 2021). Los datos médicos están siendo demandados por el cibercrimen y tienen un gran valor en el mercado negro (Kandasamy et al., 2022).

Por otra parte, los registros médicos personales (e-PHR) son cada vez más comunes. Los propios ciudadanos están tomando la iniciativa del uso de plataformas donde se aloja la información de diversos proveedores de salud. Estas iniciativas traen ventajas para los ciudadanos porque facilitan la centralización de todo su historial de salud en un solo lugar, lo que les permite mantener y controlar más fácilmente sus datos clínicos, minimizando errores y duplicidad de información (Alsañafi et al., 2022). A pesar de ello, siguen latentes los retos para salvaguardar los datos y la privacidad en estos entornos.

Un aspecto que preocupa son los usos secundarios que se puede dar a la información en los EHR o e-PHR, como compartir información para la investigación clínica, la vigilancia automatizada de enfermedades, auditorías clínicas para mejorar la calidad, enviar la información a diferentes agencias gubernamentales, de seguros, etcétera. Estos usos secundarios también plantean riesgos para la seguridad y privacidad de los datos. Si bien se incluyen mecanismos como el consentimiento informado o las leyes de protección de datos, las técnicas de vanguardia como el aprendizaje automático (ML), análisis de datos (DA) y piratería hacen que sea aún más difícil proteger completamente la privacidad de un individuo/paciente (Shah & Khan, 2020). Es indispensable, por lo tanto, que estos sistemas dispongan de protocolos seguros de almacenamiento, gestión y transmisión de datos.

2.3 Ética en el uso de la información médica

Las técnicas de inteligencia artificial aplicadas en la investigación en el área de la salud enfrentan cuestiones éticas que deben ser analizadas para evitar una vulneración de la seguridad de los datos y la privacidad de las personas. Desde 1979, se han establecido principios de respeto, justicia y beneficencia en la investigación con personas, descritos en el Informe de Belmont (The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1979) que plantea los principios éticos que deben guiar a investigadores y custodios de datos.

Adicionalmente a las leyes de protección de datos, los comités de ética de la investigación cumplen un rol importante para delinear el trabajo de los equipos de

investigación. A pesar de los informes, leyes y reglamentos internos de comités de ética, de acuerdo con Gefenas et al. (2022), falta una cooperación entre los comités de ética de investigación y los organismos rectores de protección de datos, con el fin de establecer las justificaciones éticas y legales de los estudios de investigación con datos personales. De esta manera, se evitarían contradicciones entre normas y se proporcionaría seguridad y privacidad a los datos personales.

Para lograr este objetivo, se deben crear normativas en el ámbito gubernamental. Los gobiernos deben establecer un marco legal y ético para la investigación y el desarrollo, así como facilitar mecanismos de comunicación para que los ciudadanos estén informados del beneficio de la investigación basada en datos, de manera que hagan posible la colaboración a través de asociaciones público-privadas académicas (Boudreau LeBlanc et al., 2022). Una tarea pendiente en América latina y el Caribe es crear el ambiente para generar confianza en la institucionalidad de los Estados; en un reporte publicado por el Banco Interamericano de Desarrollo (BID) sobre la confianza en la región, los porcentajes en América latina son muy bajos, comparados con países europeos (Keefer & Scartascini, 2022): nueve de cada diez encuestados no confía en los demás. Y cuando se trata de instituciones de gobierno, menos de una de cada tres personas confía en las instituciones gubernamentales. En cambio, en países de Europa, por ejemplo, una encuesta realizada en Alemania sobre el uso de aplicaciones móviles en esta era de pandemia revela que los ciudadanos confían más en instituciones estatales que en privadas, y un 78,8 % de los encuestados están dispuestos a donar datos a instituciones estatales o institutos de investigación para la investigación médica (Buhr et al., 2022). De acuerdo con el informe, la desconfianza genera bajos índices de desarrollo en innovación o investigación. Esto se convierte en una limitante para facilitar el acceso a los datos.

2.4 Repositorios de información de salud: lagos de datos

Almacenar información de diferentes fuentes es un proceso engorroso y costoso. Con el advenimiento del *big data*, es un reto conseguir grandes repositorios para guardar información diversa y de diferentes fuentes. Los almacenes de datos cumplen estas funciones a través de procesos de ETL (*extract, transform, load*), que implican altos costos en la creación, mantenimiento y gestión de estos entornos. Este es uno de los motivos por los cuales los lagos de datos se están posicionando como una alternativa para almacenar grandes cantidades de información en su forma y formato original, sin costos altos por la gestión de los datos (Eder & Shekhovtsov, 2021).

Se ha demostrado que los lagos de datos permiten a las organizaciones procesar datos a una escala que antes no hubiera sido factible sin sistemas especializados y costosos. Facilitan a las organizaciones almacenar datos sin procesos complejos de

ETL, pero sí asociados a un conjunto de políticas, para ser procesados por múltiples equipos y de acuerdo con sus necesidades (Rooney et al., 2019).

El objetivo principal de un lago de datos es hacer posible una explotación completa de su contenido para aportar valor a los datos de las organizaciones. Los lagos de datos dentro de un ambiente de un sistema de información deben ser considerados como un nuevo componente del sistema. La diferencia entre los almacenes de datos y los lagos de datos es que estos últimos son ágiles y flexibles. Tienen sus propias reglas para el gobierno de datos, en términos de gestión de la calidad de los datos, la seguridad de los datos, el ciclo de vida de los datos y los metadatos. Laurent et al. (2020) proponen como un componente clave de la arquitectura del lago de datos el catálogo de metadatos para garantizar la coherencia de las fuentes de datos y su eficiente gestión, evitando así los denominados “pantanos de datos”.

2.5 Datos abiertos, gobiernos abiertos

Los datos abiertos son fundamentales para los Estados y las instituciones privadas, academia e institutos de investigación, porque fomentan la transparencia y participación de sus miembros, impulsan el aprendizaje y la investigación. Uno de los primeros gobiernos que introdujeron el concepto de gobierno abierto fue el de Barack Obama (Estados Unidos), y ha marcado el punto de inicio para establecer las directrices para la apertura y facilidad de acceso a las actividades gubernamentales y administrativas, con el fin de impulsar acciones prácticas y garantizar un control generalizado sobre la gestión de los bienes públicos por parte de los ciudadanos (Galeone & Bonzi, 2020).

De acuerdo con la Carta Internacional de Datos Abiertos (Open Data Charter, 2015), desarrollada en el 2015 por algunos gobiernos, sociedad civil y expertos de orden mundial, se estableció un conjunto de seis principios básicos de datos abiertos:

1. *Datos abiertos por defecto.* Los gobiernos deben publicar todo tipo de datos, excepto aquellos que justificadamente se mantienen cerrados por motivos de seguridad o protección de datos.
2. *Oportunos y exhaustivos.* Son datos completos, relevantes, publicados en su forma original, sin modificaciones.
3. *Accesibles y utilizables.* Son datos fáciles de acceder, en diferentes formatos, gratuitos y legibles por máquinas.
4. *Comparables e interoperables.* Mayor acceso a conjuntos de datos de calidad y estandarización para mejorar la interoperabilidad.
5. *Para mejorar la gobernanza y la participación ciudadana.* Destaca la importancia de compartir habilidades técnicas y experiencias entre países, y la

transparencia en los métodos de recopilación, estandarización y procesos de publicación.

6. *Para el desarrollo inclusivo y la innovación.* Fomentar la reutilización de datos abiertos para fines comerciales y no comerciales, con el fin impulsar sectores de la educación, salud, medioambiente, etcétera.

Todos los datos recopilados pueden terminar en silos de información, sin aportar al objetivo por el que fueron recogidos. Para evitar esto, las tecnologías de web semántica proporcionan prácticas y estándares para publicar y conectar los datos posibilitando la interoperabilidad de datos, información y sistemas (Calva & Piedra, 2020). Cada vez más se evidencia cómo los datos vinculados se han ido incorporando en diferentes iniciativas o proyectos que permiten mejorar servicios educativos y de salud para los ciudadanos, investigación médica, intervenciones ambientales y climáticas, así como un uso más eficiente de los escasos recursos energéticos. De acuerdo con la revisión bibliográfica de Avila-Garzon (2020), el mayor uso que se ha dado a los datos abiertos enlazados (LOD) es en las áreas de biología y ciencias sociales.

En su revisión bibliográfica, Galeone y Bonzi (2020) mencionan esta evolución significativa sobre la cantidad de datos de salud disponibles gratuitamente para el público en los últimos años. Las investigaciones se centran en el uso de tecnologías de web semántica, como el marco de descripción de recursos (RDF), para que los datos sean accesibles y utilizables. Por ejemplo, la creación de vocabularios RDF como Observational Health Data Science and Informatics (OHDSI), conectado a recursos vinculados; conjuntos de datos vinculados sobre el cáncer y recursos biomédicos; y aplicaciones de la investigación del cáncer El ReVeaLD (un explorador visual en tiempo real y agregador de datos vinculados). En el ámbito farmacéutico, Drugle es un buscador semántico que lista y analiza información sobre medicamentos disponibles en internet a través de una simple consulta (Galeone & Bonzi, 2020). Esto puede ser beneficioso para dar un valor agregado a sistemas de registros clínicos o sistemas de aprendizaje de salud, y para fortalecer la interoperabilidad entre diferentes organismos.

3. SUGERENCIAS PARA INVESTIGADORES Y TRABAJOS FUTUROS

Son diversos y complejos los factores que afectan el acceso a la información médica; por ello, se sugiere generar grupos de investigación con vinculación social, en donde participen los gobiernos, instituciones privadas y la academia, con el fin de impulsar observatorios de datos o proyectos piloto para crear un ecosistema de datos abiertos médicos.

La academia, con sus grupos de investigación, debe usar todos los mecanismos vigentes para protección de datos personales para salvaguardar la integridad y seguridad de la información personal. Es necesaria una estrecha comunicación con los comités de bioética, manejadores de datos y propietarios de datos.

Es indispensable que los investigadores busquen nuevas formas de comunicar y socializar los proyectos, resultados y conclusiones de las investigaciones, para que tengan un mayor alcance de difusión. Una ciudadanía bien informada y educada facilita e impulsa la ejecución de proyectos sociales en el ámbito sanitario.

Los trabajos futuros deben estar enfocados en encontrar mejores formas de construir almacenes de datos abiertos colaborativos, a bajo costo y con interfaces fáciles de modificar, configurar y usar por el usuario, para asegurar su uso y aplicabilidad en entornos reales de trabajo. Adicionalmente, los Estados deben planificar políticas nacionales estandarizadas que cumplan con estándares de informática médica para la recolección, transformación y procesamiento de información sanitaria; esto contribuirá en gran medida a tener almacenes de datos o lagos de datos abiertos, oportunos y exhaustivos, accesibles y utilizables, comparables e interoperables.

4. CONCLUSIONES

El reporte de confianza del BID (Keefer & Scartascini, 2022) muestra cómo América Latina y el Caribe manifiestan niveles bajos de confianza institucional. Por lo tanto, los gobiernos deben enfocarse en generar políticas en el ámbito de la salud acordes a la realidad de sus sociedades y que produzcan resultados visibles para el bienestar ciudadano. Con buena política pública, con alianzas estratégicas con la sociedad civil y la academia, se puede avanzar en la planificación de proyectos nacionales para datos abiertos en salud. Por otro lado, los gobiernos deben crear un plan de comunicación e información ciudadana para concientizar a la población de las ventajas y oportunidades de los gobiernos abiertos, y de cómo la ciudadanía puede contribuir como donante de datos para proyectos de investigación e innovación. Asimismo, se debe asegurar que las leyes, reglamentos y normativas permitan un alto nivel de confianza y privacidad de la información personal. Las leyes como el RGPD o la HIPPA que adopten los países deben ser ajustadas a las realidades de cada sociedad.

Si bien existe normativa, reglamentos y leyes para protección de datos personales, un aspecto clave es el actuar ético de controladores, manejadores, custodios de datos e investigadores. La investigación y los enfoques de estructuras tecnológicas deben estar centradas en el ser humano, particularmente en comunidades con niveles más bajos de alfabetización de datos, alfabetización

digital o alfabetización en general (Dunbar et al., 2021). Solo así, se puede minimizar los riesgos en la vulneración de los derechos de la persona.

REFERENCIAS

- Alsahafi, Y. A., Gay, V., & Khwaji, A. A. (2022). Factors affecting the acceptance of integrated electronic personal health records in Saudi Arabia: The impact of e-health literacy. *Health Information Management Journal*, 51(2), 98-109. <https://doi.org/10.1177/1833358320964899>
- Avila-Garzon, C. (2020). Applications, methodologies, and technologies for linked open data: A systematic literature review. *International Journal on Semantic Web and Information Systems*, 16(3), 53-68. <https://doi.org/10.4018/IJSWIS.2020070104>
- Bauer, D., Froese, F., Garcés-Erice, L., Giblin, C., Labbi, A., Nagy, Z. A., Pardon, N., Rooney, S., Urbanetz, P., Vetsch, P., & Wespi, A. (2021). Building and operating a large-scale enterprise data analytics platform. *Big Data Research*, 23. <https://doi.org/10.1016/j.bdr.2020.100181>
- Boudreau LeBlanc, A., Williams-Jones, B., & Aenishaenslin, C. (2022). Bio-ethics and One Health: A case study approach to building reflexive governance. *Frontiers in Public Health*, 10. <https://doi.org/10.3389/fpubh.2022.648593>
- Buhr, L., Schick Tanz, S., & Nordmeyer, E. (2022). Attitudes toward mobile apps for pandemic research among smartphone users in Germany: National survey. *JMIR MHealth and UHealth*, 10(1). <https://doi.org/10.2196/31857>
- Calva, M., & Piedra, N. (2020). Cycle of transformation of open and linked medical data to improve the training of health professionals. En *Proceedings of the 15th Latin American Conference on Learning Technologies (LACLO)* (pp. 1-10). <https://doi.org/10.1109/LACLO50806.2020.9381158>
- Cohen, G., Gerke, S., & Kramer, D. B. (2020). Ethical and legal implications of remote monitoring of medical devices. *The Milbank Quarterly*, 98(4), 1257-1289.
- Comandè, G., & Schneider, G. (2021). Can the GDPR make data flow for research easier? Yes it can, by differentiating! A careful reading of the GDPR shows how EU data protection law leaves open some significant flexibilities for data protection-sound research activities. *Computer Law & Security Review*, 41. <https://doi.org/10.1016/j.clsr.2021.105539>
- Dunbar, E., Olsen, H. E., Salomon, E., Bhatt, S., Mutuku, R., Wasunna, B., Edwards, J., Kolko, B., & Holeman, I. (2021). Towards Responsible Data Practices in Digital Health: A case study of an open source community's journey. En *Extended*

- Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems (CHI EA '21)* (artículo 42, pp. 1-8). Association for Computing Machinery. <https://doi.org/10.1145/3411763.3443438>
- Eder, J., & Shekhovtsov, V. A. (2021). Data quality for federated medical data lakes. *International Journal of Web Information Systems*, 17(5), 407-426. <https://doi.org/10.1108/IJWIS-03-2021-0026>
- Galeone, C., & Bonzi, R. (2020). Open data: Applications in social care and health. *Statistica Applicata - Italian Journal of Applied Statistics*, 32(2), 125-139. <https://doi.org/10.26398/IJAS.0032-008>
- Gefenas, E., Lekstutiene, J., Lukaseviciene, V., Hartlev, M., Mourby, M., & Cathaoir, K. (2022). Controversies between regulations of research ethics and protection of personal data: Informed consent at a cross-road. *Medicine, Health Care and Philosophy*, 25(1), 23-30. <https://doi.org/10.1007/s11019-021-10060-1>
- Healthcare Information and Management Systems Society. (2021). *2021 HIMSS Healthcare Cybersecurity Survey*. <https://www.himss.org/resources/himss-healthcare-cybersecurity-survey>
- Jones, K. H., Ford, E. M., Lea, N., Griffiths, L. J., Hassan, L., Heys, S., Squires, E., & Nenadic, G. (2020). Toward the development of data governance standards for using clinical free-text data in health research: Position paper. *Journal of Medical Internet Research*, 22(6). <https://doi.org/10.2196/16760>
- Kandasamy, K., Srinivas, S., Achuthan, K., & Rangan, V. P. (2022). Digital healthcare - cyberattacks in Asian organizations: An analysis of vulnerabilities, risks, NIST perspectives, and recommendations. *IEEE Access*, 10, 12345-12364. <https://doi.org/10.1109/ACCESS.2022.3145372>
- Keefer, P., & Scartascini, C. (2022). *Confianza: la clave de la cohesión social y el crecimiento en América Latina y el Caribe*. Banco Interamericano de Desarrollo. <http://dx.doi.org/10.18235/0003911>
- Laurent, A., Laurent D., & Madera, C. (2020). *Data Lakes*. John Wiley & Sons.
- Open Data Charter. (2015). *Carta Internacional de Datos Abiertos*. <https://opendatacharter.net/principles-es/>
- Pierce, R., & Evram, M. (2022). Getting it right: Implementing data protection in citizen science research. *Insights: The UKSG Journal*, 35. <https://doi.org/10.1629/UKSG.538>
- Red Iberoamericana de Protección de Datos. (2019). *Recomendaciones generales para el tratamiento de datos en la inteligencia artificial*. <https://www.redipd.org/>

sites/default/files/2020-02/guia-recomendaciones-generales-tratamiento-datos-ia.pdf

Rooney, S., Bauer, D., Garcés-Erice, L., Urbanetz, P., Froese, F., & Tomic, S. (2019). Experiences with managing data ingestion into a corporate datalake. En *Proceedings-2019 IEEE 5th International Conference on Collaboration and Internet Computing (CIC)* (pp. 101-109). <https://doi.org/10.1109/CIC48465.2019.00021>

Shah, S. M., & Khan, R. A. (2020). Secondary use of electronic health record: Opportunities and challenges. *IEEE Access*, 8, 136947-136965. <https://doi.org/10.1109/ACCESS.2020.3011099>

The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1979). *The Belmont Report. Ethical principles and guidelines for the protection of human subjects of research*. <https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/read-the-belmont-report/index.html>

MODELAMIENTO DE TÓPICOS UTILIZANDO MENSAJES DE TWITTER RELACIONADOS CON EL CÁNCER CERVICAL

RUTH MARÍA REÁTEGUI ROJAS

rmreategui@utpl.edu.ec

<https://orcid.org/0000-0002-7304-4413>

Universidad Técnica Particular de Loja, Ecuador

RESUMEN. El cáncer cervical es un problema de salud mundial que genera una gran cantidad de información que circula por las redes sociales. El modelado de tópicos permite automáticamente identificar aquellos que tratan sobre un tema en un conjunto de documentos. En el presente trabajo se ha aplicado el modelamiento de tópicos para identificar los tópicos de un conjunto de tuits que tratan sobre al cáncer cervical. Se aplicó el algoritmo LDA y la métrica de la coherencia para la evaluación. Se identificaron siete tópicos relacionados con el efecto de las vacunas contra el VPH, la relación que existe entre el VPH y otras enfermedades, las formas de prevención como vacunas y test de Papanicolaou, programas que prestan servicios médicos para la prevención y eliminación de esta enfermedad, historias de mujeres que han padecido de cáncer cervical y estudios dirigidos a mujeres latinas.

PALABRAS CLAVE: text mining, Twitter, cáncer cervicouterino, modelado de tópicos

TOPIC MODELING USING TWITTER MESSAGES RELATED TO CERVICAL CANCER

ABSTRACT. As a global health problem, cervical cancer generates much information that circulates through social networks. Modeling allows us to automatically identify the topics that deal with a specific subject matter in a set of documents. This research used the LDA algorithm and the coherence metric for topic modeling and identified seven topics in a set of tweets on cervical cancer. The topics were related to the effect of HPV vaccines, the relationship between HPV and other diseases, forms of prevention such as vaccines and Papanicolaou tests, programs that provide medical services for the prevention and elimination of this disease, stories of women who have had cervical cancer and studies aimed at Latina women.

KEYWORDS: text mining, Twitter, cervical cancer, topic modeling

1. INTRODUCCIÓN

El cáncer cervical es el cuarto tipo de cáncer más frecuente en las mujeres a nivel mundial, pero genera un mayor impacto en países de bajos recursos económicos. Este tipo de cáncer afecta al cuello del útero y el 99 % de las veces se debe al virus del papiloma humano (VPH) (Organización Mundial de la Salud, 2022). Entre los tipos de VPH considerados de alto riesgo están el 16, 18, 31, 33, 45, 52 y 58, mientras que los tipos considerados de bajo riesgo son el 6 y el 11.

El cáncer cervical es prevenible y curable si es diagnosticado a tiempo. Como método de prevención primaria, existe la vacuna contra el VPH que se recomienda aplicarla antes del inicio de la vida sexual. Además, como método secundario, se tiene el tamizaje y tratamiento de lesiones precancerosas. A pesar de la existencia de estos métodos, la desigualdad económica, social, cultural y racial incide en el acceso a servicios médicos, incluso en los países desarrollados (Hernández et al., 2021).

Puesto que el cáncer cervical es un problema de salud mundial, son muchos los programas, estudios y eventos que se realizan en torno a esta enfermedad. A través de las redes sociales (por ejemplo, Twitter), se puede observar cómo el personal del área de la salud, los investigadores e incluso los pacientes comparten temas de interés. Las redes sociales se han convertido en una fuente popular para la difusión y debate de información sobre salud (Lyu et al., 2021), pues contienen opiniones que permiten comprender la percepción popular sobre temas como diagnóstico médico, medicamentos, instalaciones y reclamos (Asghari et al., 2019); además, actúan como un sensor del comportamiento de una persona (Hinduja et al., 2022) o incluso para establecer comunidades de soporte para ciertas enfermedades (Jansli, 2022).

Debido a la gran cantidad de información que puede circular por las redes sociales, la aplicación de la inteligencia artificial y sus diferentes ramas, como la minería de texto y procesamiento de lenguaje natural, hace posible extraer información para inferir conocimiento relevante en torno a un determinado tema. El modelado de tópicos (*topic modeling*, en inglés) permite automáticamente identificar temas o tópicos de un conjunto de documentos con formato no estructurado. La asignación latente de Dirichlet (LDA, por sus siglas en inglés) es un modelo probabilístico generativo de tópicos, en el cual los documentos se representan como una combinación aleatoria sobre tópicos latentes, y cada tópico es representado por una distribución de probabilidades de palabras (Blei et al., 2003).

El modelado de tópicos ha sido utilizado ampliamente para obtener información de las redes sociales en temas relacionados con la salud. Musaeu et al. (2019) examinan las conversaciones de Twitter sobre la salud cardiovascular para

evaluar el enfoque temático de estas conversaciones, así como el papel de varios departamentos de salud estatales. Alahmari et al. (2022) aplican modelamiento de tópicos para analizar la red social Twitter en temas relacionados con el cáncer. Además, Joshi et al. (2022) proponen un modelo basado en LDA para extraer e identificar las reacciones adversas de medicamentos basado en *posts* provenientes de las redes sociales. El trabajo presentado por Sidana et al. (2018) se enfoca en el monitoreo de la salud de las personas a lo largo del tiempo en la red Twitter y utiliza algunos modelos de tópicos.

En los últimos años, también se han desarrollado varios trabajos que utilizan el modelado de tópicos para el análisis de los tuits relacionados con la pandemia del COVID-19. El estudio de Ma et al. (2021) aplicó dos métodos de modelado de tópicos: LDA y Top2Vec, para analizar comentarios de usuarios con respecto a las dudas sobre la vacunación. Kwok et al. (2021) y Bogdanowicz y Guan (2022) utilizaron el modelo de tópicos para identificar temas que se discuten respecto al COVID-19 tanto en Australia como en Estados Unidos, respectivamente. Lyu et al. (2021) emplearon LDA para identificar percepciones, actitudes y preocupaciones relacionadas con el vapeo, y para discernir posibles desinformaciones y conceptos erróneos sobre esta acción durante la pandemia. Otro reciente estudio publicado por Oliveira et al. (2022) usó el mismo algoritmo LDA para analizar nuevos temas en tuits y luego proceder a identificar emociones en respuestas a tuits de noticias.

En vista de la utilidad del modelado de tópicos en el área de la salud, en el presente trabajo se propone la aplicación del modelo LDA para identificar temas relacionados con el cáncer cervical en los mensajes de la red social Twitter.

2. METODOLOGÍA

2.1 Obtención de datos

Los datos fueron recolectados de la red social Twitter en los meses de noviembre y diciembre del 2021. Se buscó tuits con las palabras *cervical cancer* en el idioma inglés, puesto que se constató que hay una mayor interacción en esta lengua. Se obtuvo un total de 630 registros únicos, y para el análisis se utilizó el contenido textual de los tuits.

2.2 Preprocesamiento

En este trabajo se utilizó el lenguaje Python, que proporciona varias librerías relacionadas con el procesamiento de lenguaje natural y la aplicación del algoritmo LDA. Durante el preprocesamiento, se usaron las librerías NLTK y Spacy. Algunos pasos de esta fase se detallan en la Tabla 1.

que cada una aparece en el documento y su frecuencia en el conjunto de documentos. De esta forma, no solo las palabras más comunes serán las más relevantes.

2.4 Identificación de tópicos

Una vez obtenidos los valores TF-IDF, se procedió a aplicar el modelo LDA definido en la librería Gensim. LDA maneja los parámetros alfa y beta; para el presente trabajo, se consideraron los valores que vienen establecidos por defecto. Alfa se refiere a la distribución de documentos por tema; y beta, a la distribución de palabras por tema (Blei et al., 2003). Otro de los parámetros es el valor k , que hace referencia al número de tópicos que se desea obtener. Para conocer cuál es el mejor valor de k , se utilizó la métrica de coherencia.

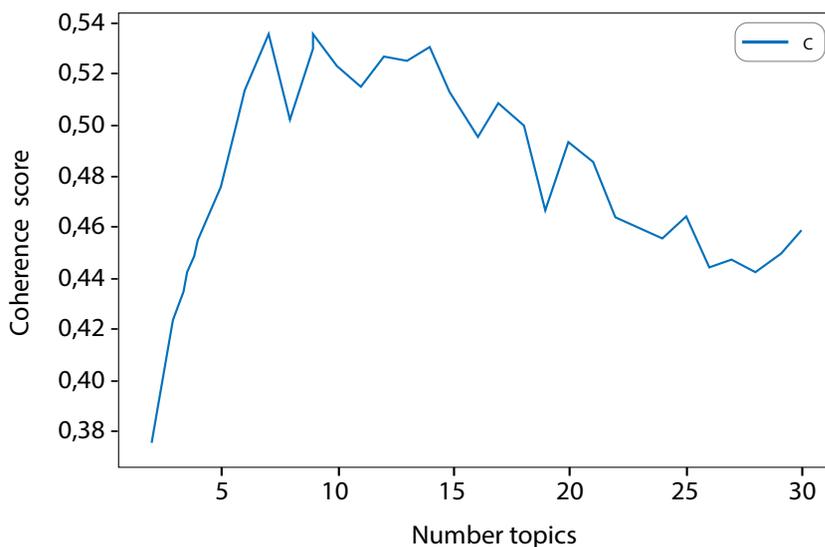
2.5 Evaluación

La métrica de coherencia permite calificar un tema o tópico midiendo el grado de similitud semántica entre las palabras de alto puntaje en el tema. Se utilizó, igualmente, la librería Gensim que se configuró para que se use la métrica 'c_v'. Esta métrica trabaja con una ventana deslizante y combina la medida del coseno de similitud con la información mutua normalizada por puntos (Röder et al., 2015; Syed & Spruit, 2017).

La Figura 2 muestra los valores de la coherencia para valores de k que van de 2 a 30. Se puede observar que el valor más alto de la coherencia (0,54) se obtiene cuando $k = 7$. Por tanto, se procedió a identificar siete tópicos.

Figura 2

Métrica de coherencia



3. RESULTADOS

Una vez identificado el mejor número de tópicos $k = 7$, las palabras que corresponden a cada tópico se encuentran en la Tabla 2. Los números que preceden a cada término son la probabilidad de que cada término corresponda a dicho tópico.

La Tabla 3 muestra la cantidad de tuits por cada tópico; se ha tomado en cuenta el tópico dominante en cada tuit. Se puede observar que la mayoría de los tuits corresponden a los tópicos 2 y 3.

Tabla 2

Tópicos identificados y palabras por tópico

Número de tópico	Palabras por cada tópico
1	0,014**"vaccination" + 0,013**"may" + 0,012**"study" + 0,012**"prevention" + 0,012**"smoke" + 0,011**"reduce" + 0,010**"need" + 0,010**"support" + 0,010**"risk" + 0,010**"late"
2	0,024**"vaccine" + 0,019**"hpv" + 0,017**"woman" + 0,013**"get" + 0,011**"pap" + 0,011**"tell" + 0,011**"live" + 0,011**"year" + 0,010**"prevent" + 0,010**"help"
3	0,024**"screening" + 0,018**"perspective" + 0,017**"health" + 0,011**"campaign" + 0,010**"survivor" + 0,009**"available" + 0,008**"join" + 0,008**"still" + 0,007**"develope" + 0,007**"session"
4	0,020**"life" + 0,016**"elimination" + 0,013**"test" + 0,010**"breast" + 0,009**"save" + 0,009**"clinical" + 0,009**"important" + 0,009**"come" + 0,009**"radiomic" + 0,008**"screen"
5	0,018**"hispanic" + 0,017**"go" + 0,017**"cervicalcancer" + 0,014**"give" + 0,014**"treatment" + 0,012**"pregnant" + 0,010**"week" + 0,009**"know" + 0,008**"oncology" + 0,008**"cause"
6	0,018**"new" + 0,018**"photographer" + 0,016**"be" + 0,014**"stage" + 0,011**"work" + 0,009**"story" + 0,009**"journey" + 0,008**"release" + 0,007**"australian" + 0,007**"challenge"
7	0,023**"global" + 0,016**"doctor" + 0,014**"die" + 0,014**"day" + 0,012**"reason" + 0,010**"could" + 0,010**"early" + 0,009**"action" + 0,009**"leader" + 0,008**"earlystage"

Tabla 3

Documentos por tópico

Número de tópico	Número de tuits
1	78
2	161
3	105
4	74
5	89

(continúa)

(continuación)

Número de tópico	Número de tuits
6	73
7	50
Total	630

4. DISCUSIÓN DE LOS RESULTADOS

Una de las dificultades de este tipo de modelos es la interpretación humana que se puede dar a los diferentes tópicos encontrados. Con base en los términos que forman los tópicos y una revisión manual de algunos tuits que tienen como dominante el tópico que se va a analizar, se presenta a continuación la interpretación de cada uno de los tópicos. La Tabla 4 ofrece algunos ejemplos de tuits por tópico.

- *Tópico 1.* Entre las palabras que conforman este tópico están *vaccination, may, study, prevention, smoke, reduce, need, support, risk, late*. Este tópico principalmente abarca tuits que tratan sobre los efectos de la vacunación contra el cáncer cervical. También varios textos hacen referencia a estudios relacionados con el cáncer cervical.
- *Tópico 2.* Algunas de las palabras que se encuentran en este tópico son *vaccine, hpv, woman, get, pap, tell, live, year, prevent, help*. Dentro de este tópico se encuentran tuits que mencionan algún tipo de relación entre los virus de VPH y el virus de la inmunodeficiencia humana (VIH). Además, se comenta sobre algunas formas de prevención como la vacuna y los test de Papanicolaou.
- *Tópico 3.* Algunas de las palabras que forman parte de este tópico son *screening, perspective, health, campaign, survivor, available, join, still, develop, session*. Dentro de este tópico se encuentran tuits que comentan sobre programas o instituciones que prestan servicios para la prevención del cáncer cervical; entre los servicios que se mencionan está el tamizaje (*screening*, en inglés).
- *Tópico 4.* Entre las palabras que conforman este tópico están *life, elimination, test, breast, save, clinical, important, come, radiomic, screen*. Este tópico hace referencia a los exámenes de prevención para el cáncer cervical y de seno. Menciona algunos estudios o programas en países de bajos recursos para eliminar el cáncer cervical, así como estudios científicos basados en radiómica.
- *Tópico 5.* En este tópico se encuentran palabras como *hispanic, go, cervical-cancer, give, treatment, pregnant, week, know, oncology, cause*. Este tópico

abarca tuits que hablan sobre algunos estudios o programas dirigidos a mujeres latinas o hispanas.

- **Tópico 6.** Entre las palabras que conforman este tópico están *new, photographer, be, stage, work, story, journey, release, Australian, challenge*. Este tópico abarca tuits que tratan de historias de mujeres que han padecido de cáncer cervical. Además, algunos textos comentan sobre nuevos estudios, programas o historias relacionados con el cáncer cervical.
- **Tópico 7.** Algunas de las palabras que conforman este tópico son *global, doctor, die, day, reason, could, early, action, leader, earlystage*. Este tópico considera tuits que tratan sobre programas o acciones globales para eliminar el cáncer cervical. Además, nombran algunos casos de personas que han fallecido por este tipo de cáncer.

Tabla 4

Tuits por tópicos dominantes

Tópico	Tuits
1	<p>The effects of the national HPV vaccination programme in England, UK, on cervical cancer and grade 3 cervical intraepithelial neoplasia incidence: a register-based observational study https://t.co/HrRnTAL9wz</p> <p>A new study reports a 38 % drop in cervical cancer incidence and a 43 % decline in mortality among young women and girls after HPV vaccination was introduced in the U.S. @Medscape @slusom</p>
2	<p>Women living with HIV/AIDS have up to 5 times higher risk of acquiring cervical cancer compared to those without. This World AIDs day, let us regularize cervical cancer screening. We also call upon Government and partners to integrate HIV/AIDs and Cancer services. https://t.co/XvDvRDZNbh</p> <p>Women with HIV should be screened regularly for HPV. Women have shown a higher preference to urine for HPV testing for cervical cancer. Read more in our blog - #WorldAIDSDay</p>
3	<p>Cancer is preventable, treatable, & beatable. Erie County Cancer Services provides free breast, cervical, & colon cancer screenings & education to eligible un- & under-insured men & women. More info at: https://t.co/2MfXrISsGI</p> <p>"Title X is the only federal grant program that funds family planning and counseling programs to help patients access contraception, as well as breast and cervical cancer screenings, screenings and treatments for sexually transmitted ... https://t.co/QhaEl73tMr https://t.co/Q83tQT9r5l"</p>
4	<p>"Day by day women are getting aware about Breast and Cervical Cancer early detection and prevention. Canapp is Playing an important role in educating women about Breast Self Exam Thanks to @madhuchauhanbjp for the initiative and helping us and encouraging women to come for checkup https://t.co/icPrTmlrFW"</p> <p>"Machine learning based evaluation of clinical and pretreatment (18)F-FDG-PET/CT radiomic features to predict prognosis of cervical cancer patients https://t.co/QqIFXO2hVH"</p>

(continúa)

(continuación)

Tópico	Tuits
5	<p>Hispanic women born outside the U. S. have less frequent cervical cancer screenings than Hispanic women born in the U. S.</p> <p>Hispanic women who live in Tarrant County are more likely to die from cervical cancer than Hispanic women who live in Dallas County. for @FortWorthReport https://t.co/cZAMzSw5bW</p>
6	<p>An inspiring story from HK-based Australian photographer Julia Broad, who has just gone through a tough cancer recovery journey https://t.co/DlpuHdCCm9</p> <p>A GoFundMe has been started for an #EnfieldNS single mom of four, Donna Boutilier, who has had her cervical cancer return. Here's how you can help her out: #EastHants #Halifax #NS https://t.co/lGj5Qxbxvo</p>
7	<p>Global leaders call for cervical cancer elimination on Day of Action https://t.co/FK8f0UzShu https://t.co/hkczFjorjc</p> <p>CA587:A621horeography written by: William Masters As soon as Albert Andrew Fischer's wife died of cervical cancer, after a final, twenty-two-day hospitalization, he felt relief. Such relief derived not from the end of his https://t.co/AjBteLwFne</p>

5. CONCLUSIONES

Actualmente, en las redes sociales como Twitter, profesionales de la salud, pacientes e investigadores comparten información de su preferencia. Esta puede ser analizada automáticamente con técnicas de inteligencia artificial y generar conocimiento de interés para el área de la salud. Aplicando el algoritmo LDA, fue posible identificar cuáles son los temas principales o tópicos que se comenta en Twitter. Entre los temas de interés están el efecto de las vacunas contra el VPH (tópico 1), la relación que existe entre el VPH y otras enfermedades como el sida (tópico 2) y el cáncer de seno (tópico 4), formas de prevención como vacunas y test de Papanicolaou (tópico 2), programas o instituciones que prestan servicios médicos para la prevención del cáncer cervical (tópico 3), programas o estudios para eliminar esta enfermedad (tópico 4, tópico 7), historias de mujeres que han padecido de cáncer cervical (tópico 6) y estudios dirigidos a mujeres latinas o hispanas (tópico 5).

En este estudio, a pesar de que se consideraron tuits en inglés, los tópicos identificados pueden servir de ejemplo para que instituciones de salud y gobiernos puedan difundir información que sea de mayor interés para sus pacientes y para definir estrategias de prevención de este tipo de enfermedades. Además, la Organización Panamericana de la Salud recalca que la mortalidad por cáncer cervical es tres veces más alta en América Latina y el Caribe en comparación con Norteamérica. Este tipo de desigualdad en la salud se mencionó en uno de los tópicos identificados en este trabajo.

6. TRABAJO FUTURO

Tomando en cuenta el alto nivel de muertes entre las mujeres hispanas por este tipo de cáncer, en un inicio se consideró analizar cuáles son los temas que en Twitter se comentan sobre esta enfermedad en Ecuador o Latinoamérica, pero al no haber una cantidad importante de tuits se realizó el análisis de tuits en inglés, según lo indicado en la sección 2. Por tanto, para un trabajo futuro, se puede intentar analizar otras redes sociales donde haya mayor participación, o también la información sobre otro tipo de enfermedades, de acuerdo con las necesidades locales o nacionales. Esto puede dar respuesta a algunas preguntas de investigación como las planteadas a continuación:

- ¿Cuáles son los tópicos de interés para los pacientes que se discuten en una red social determinada?
- ¿Cuáles son los tópicos de interés para los médicos que se discuten en una red social determinada?
- ¿Qué tópicos discutidos en una red social pueden orientar a la creación de políticas públicas?

REFERENCIAS

- Alahmari, N., Alswedani, S., Alzahrani, A., Katib, I., Albeshri, A., & Mehmood, R. (2022). Musawah: A data-driven AI approach and tool to co-create healthcare services with a case study on cancer disease in Saudi Arabia. *Sustainability*, 14(6). DOI: 10.3390/su14063313
- Asghari, M., Sierra-Sosa, D., & Elmaghraby, A. (2019). Trends on health in social media: Analysis using Twitter topic modeling. En *2018 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)* (pp. 558-563). <https://doi.org/10.1109/ISSPIT.2018.8642679>
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3, 993-1022.
- Bogdanowicz, A., & Guan, C. (2022). Dynamic topic modeling of Twitter data during the COVID-19 pandemic. *PLoS ONE*, 17(5). DOI: 10.1371/journal.pone.0268669
- Hinduja, S., Afrin, M., Mistry, S., & Krishna, A. (2022). Machine learning-based proactive social-sensor service for mental health monitoring using twitter data. *International Journal of Information Management Data Insights*, 2(2). DOI: 10.1016/j.jjime.2022.100113

- Hernández, J. A., Ramírez, P. X., Valbuena-García, A. M., Acuña, L., & González-Díaz, J. A. (2021). Factors associated with delays in time to treatment initiation in Colombian women with cervical cancer: A cross-sectional analysis. *Gynecologic Oncology Reports*, 35. <https://doi.org/10.1016/j.gore.2021.100697>
- Jansli, S. M., Hudson, G., Negbenose, E., Erturk, S., Wykes, T., & Jilka, S. (2022). Investigating mental health service user views of stigma on twitter during COVID-19: A mixed-methods study. *Journal of Mental Health*, 31(4), 576-584. doi: 10.1080/09638237.2022.2091763
- Joshi, C., Attar, V. Z., & Kalamkar, S. P. (2022). An unsupervised topic modeling approach for adverse drug reaction extraction and identification from natural language text. En S. Tiwari, M. C. Trivedi, M. L. Kolhe, K. Mishra & B. K. Singh (Eds.), *Advances in data and information sciences. Lecture notes in networks and systems* (vol. 318, pp. 505-514). Springer. https://doi.org/10.1007/978-981-16-5689-7_44
- Kwok, S. W. H., Vadde, S. K., & Wang, G. (2021). Tweet topics and sentiments relating to COVID-19 vaccination among Australian Twitter users: Machine learning analysis. *Journal of Medical Internet Research*, 23(5). <https://doi.org/10.2196/26953>
- Lyu, J. C., Luli, G. K., & Ling, P. M. (2021). Vaping discussion in the COVID-19 pandemic: An observational study using twitter data. *PLoS ONE*, 16(12). <https://doi.org/10.1371/journal.pone.0260290>
- Ma, P., Zeng-Treitler, Q., & Nelson, S. J. (2021). Use of two topic modeling methods to investigate covid vaccine hesitancy. En *International Conferences ICT, Society, and Human Beings 2021; Web Based Communities and Social Media 2021; and e-Health 2021* (pp. 221-226). https://www.ict-conf.org/wp-content/uploads/2021/07/04_202106C030_Ma.pdf
- Musaeu, A., Britt, R. K., Hayes, J., Britt, B. C., Maddox, J., & Sheinidashtegol, P. (2019). Study of twitter communications on cardiovascular disease by state health departments. En J. Miller, E. Stroulia, K. Lee & L. J. Zhang (Eds.), *Web Services - ICWS 2019. Lecture notes in computer science* (vol. 11512). https://doi.org/10.1007/978-3-030-23499-7_12
- Oliveira, F. B., Haque, A., Mougouei, D., Evans, S., Sichman, J. S., & Singh, M. P. (2022). Investigating the emotional response to COVID-19 news on twitter: A topic modelling and emotion classification approach. En *IEEE Access* (vol. 10, pp. 16883-16897). <https://doi.org/10.1109/ACCESS.2022.3150329>

- Organización Mundial de la Salud. (2022). *Cáncer cervicouterino*. https://www.who.int/es/health-topics/cervical-cancer#tab=tab_1
- Röder, M., Both, A., & Hinneburg, A. (2015). Exploring the space of topic coherence measures. En *Proceedings of the 8th ACM International Conference on Web Search and Data Mining (WSDM '15)* (pp. 399-408). <https://doi.org/10.1145/2684822.2685324>
- Sidana, S., Amer-Yahia, S., Clausel, M., Rebai, M., Mai, S. T., & Amini, M. (2018). Health monitoring on social media over time. *IEEE Transactions on Knowledge and Data Engineering*, 30(8), 1467-1480. <https://doi.org/10.1109/TKDE.2018.2795606>
- Syed, S., & Spruit, M. (2017). Full-text or abstract? Examining topic coherence scores using Latent Dirichlet Allocation. En *IEEE International Conference on Data Science and Advanced Analytics (DSAA)* (pp. 165-174). <https://doi.org/10.1109/DSAA.2017.61>

O PAPEL DA CAPACIDADE ABSORTIVA NA INOVAÇÃO. Evidências de práticas desenvolvidas por empresas de base tecnológica inseridas em redes colaborativas

GRACIELE TONIAL
graciele.tonial@unoesc.edu.br
<https://orcid.org/0000-0001-5137-035X>
Unoesc Joaçaba, Brasil

NATHALIA BERGER WERLANG
nathalia.werlang@ufsc.br
<https://orcid.org/0000-0003-0172-6025>
Universidade Federal de Santa Catarina, Brasil

ALESSANDRA CASSOL
alessandracassol.adm@gmail.com
<https://orcid.org/0000-0003-0753-9642>
Universidade Nove de Julho, Brasil

RESUMO. Existe um escasso número de pesquisas que analisam a configuração da capacidade absorptiva em empresas de países emergentes, como é o caso do Brasil. O presente estudo buscou analisar as práticas que as pequenas e médias empresas de base tecnológica brasileiras aplicam para otimizar os fluxos de conhecimento e conseqüentemente, alavancar as inovações organizacionais. Com esse objetivo, a pesquisa utilizou uma abordagem qualitativa para analisar dados de um total de 11 empresas de base tecnológica. Os dados foram coletados por meio de entrevistas semiestruturadas, observação não participante e análise documental. Os principais resultados apontam que as práticas de capacidade de absorção são determinantes da inovação organizacional, a exemplo de: adaptação o conhecimento externo a uma nova tecnologia para criar um novo produto; e o desenvolvimento da capacidade de aplicar o conhecimento interno e externo em patentes e ou produtos de softwares inovadores para o mercado. Também evidenciaram-se diferenças nas práticas aplicadas por empresas de um mesmo setor, e nessa linha, constatou-se que as empresas com maior tempo de atuação no mercado tendem a possuir mais práticas implementadas, o qual permite alavancar seus processos de inovação. Nossos resultados fornecem evidências concretas sobre como alocar os recursos do conhecimento para maximizar seu potencial para alavancar inovações organizacionais.

PALAVRAS CHAVE: capacidade absorptiva, inovação, pesquisa qualitativa, pequenas empresas, empresas de base tecnológica

EL PAPEL DE LA CAPACIDAD DE ABSORCIÓN EN LA INNOVACIÓN: EVIDENCIA DE PRÁCTICAS DESARROLLADAS POR EMPRESAS DE BASE TECNOLÓGICA INSERTADAS EN REDES COLABORATIVAS

RESUMEN. Hay un número escaso de estudios que analicen la configuración de la capacidad de absorción en empresas de países emergentes, como es el caso de Brasil. El presente estudio busca analizar las prácticas que las pequeñas y medianas empresas brasileñas de base tecnológica aplican para optimizar los flujos de conocimiento y, en consecuencia, aprovechar las innovaciones organizacionales. Con este objetivo, la investigación utiliza un enfoque cualitativo para analizar datos de un total de 11 empresas de base tecnológica. Los datos fueron recolectados a través de entrevistas semiestructuradas, observación no participante y análisis de documentos. Los principales resultados indican que las prácticas de capacidad de absorción son determinantes de la innovación organizacional, tales como adaptar el conocimiento externo a una nueva tecnología para crear un nuevo producto; y el desarrollo de la capacidad de aplicar conocimientos internos y externos en patentes y/o productos de *software* innovadores para el mercado. También se evidenciaron diferencias en las prácticas aplicadas por empresas del mismo sector, y en esta línea se encontró que las empresas con mayor tiempo en el mercado tienden a tener más prácticas implementadas, lo que les permite apalancar sus procesos de innovación. Nuestros resultados proporcionan evidencia concreta sobre cómo asignar recursos de conocimiento para maximizar su potencial para aprovechar las innovaciones organizacionales.

PALABRAS CLAVE: capacidad de absorción, innovación, investigación cualitativa, pequeños negocios, empresas TIC

THE ROLE OF ABSORPTION CAPACITY IN INNOVATION: EVIDENCE OF PRACTICES DEVELOPED BY TECHNOLOGY-BASED COMPANIES INSERTED IN COLLABORATIVE NETWORKS

ABSTRACT. There is a small number of studies that analyze the configuration of absorptive capacity in companies from emerging countries, such as Brazil. This study seeks to analyze the practices that small and medium-sized Brazilian technology-based companies apply to optimize knowledge flows and, consequently, take advantage of organizational innovations. With this objective, the research uses a qualitative approach to analyze data from a total of 11 technology-based companies. Data were collected through semi-structured interviews, non-participant observation, and document analysis. The main results indicate that absorptive capacity practices are determinants of organizational innovation, such as: adapting external knowledge to a new technology to create a new product; and the development of the capacity to apply internal and external knowledge in patents and/or innovative software products for the market. There were also differences in the practices applied by companies in the same sector, and in this line it was found that companies with more time in the market tend to have more practices implemented, which allows them to leverage their innovation processes. Our results provide concrete evidence on how to allocate knowledge resources to maximize their potential to take advantage of organizational innovations.

KEYWORDS: absorptive capacity, innovation, qualitative research, small business, ICT firm

1. INTRODUÇÃO

O conhecimento é reconhecido como um dos principais recursos estratégicos (Grant, 1996), principalmente ao considerar o ambiente organizacional dinâmico e complexo. A gestão deste conhecimento é fator determinante para o desenvolvimento de estratégias sustentáveis nas empresas (Takeuchi & Nonaka, 2009). Assim, as organizações possuem como desafio desenvolver habilidades para integrar os níveis de conhecimento tácito e explícito a outros conhecimentos externos e aplicá-los para o desenvolvimento de novas capacidades (Dávila et al., 2019).

Assim, desenvolver e fortalecer os processos da capacidade absorptiva organizacional torna-se uma alternativa estratégica capaz de contribuir para o crescimento e a sustentabilidade de empresas, principalmente as de base tecnológica (Cassol et al., 2017), que se caracterizam como empresas intensivas em conhecimento, sendo reconhecidas pelo seu elevado grau de inovação e inserção em um ambiente dinâmico e competitivo (Da Silva & Da Costa, 2019).

A capacidade de absorção do conhecimento (ACAP) apresenta um escopo de pesquisa direcionado para múltiplas áreas de pesquisa, sendo um construto teórico multidimensional na área da estratégica, pesquisas comprovam que ACA é uma capacidade que gera vantagem competitiva (Apriliyanti & Alon, 2017). Assim, pesquisas que ampliem o entendimento das relações do construto com os diferentes modelos e propostas de análise são necessários (Siachou et al., 2021).

No Brasil, pesquisas apresentadas por Engelmam et al. (2017), Manthey et al. (2017), Cassol et al. (2022) e Werlang et al. (2022) comprovam que ACAP é uma das principais capacidades que facilitam a inovação como também o desempenho organizacional. Neste contexto, destaca-se o pressuposto desta pesquisa, a qual considera que a capacidade de absorção de conhecimento pode impulsionar o processo de inovação das organizações.

Pesquisas realizadas em empresas brasileiras evidenciam a diversidade do impacto da CA nas organizações de diferentes setores ou tamanhos. Dávila et al. (2017) por exemplo, identificaram que empresas do segmento industrial utilizam os recursos da Capacidade Absortiva para impulsionar o Desempenho Inovador. Cassol et al. (2019) analisaram essa relação no ambiente de PMES brasileiras e concluíram que a capacidade absorptiva impacta positivamente a inovatividade e o desempenho inovador, no entanto não identificaram a moderação na relação entre os construtos.

Todavia, entender como as empresas brasileiras intensivas em conhecimento utilizam os recursos intangíveis do conhecimento disponíveis em ambientes externos para promover a inovação ainda permanece como tema pouco explorado (Cajuela & Galina, 2020). Pesquisas que analisam a contribuição da ACAP para a

inovação e competitividade em organizações que atuam em países emergentes ainda são incipientes (Nagano et al. 2014). Assim, estudos recentes apontam para a necessidade de pesquisas que permitam ampliar o entendimento da ACAP no contexto brasileiro (Dávila et al. 2017). Garrido et al. (2017) também sugerem que abordagens metodológicas qualitativas possam ampliar as análises sobre a relação entre inovação e capacidade absorptiva.

A partir deste contexto, delimita-se a questão de pesquisa: Quais as principais práticas da ACAP desenvolvidas pelas empresas de base tecnológica para promover a inovação? Isto posto, propõe-se como objetivo deste estudo: Analisar as práticas que empresas de base tecnológica desenvolvem para desenvolver sua capacidade absorptiva e inovação, a partir de um estudo exploratório e descritivo de abordagem qualitativa, por meio da análise de diferentes casos.

O estudo torna-se relevante já que trata de empresas de tecnologia de um país emergente com capacidades adaptativa, absorptiva e inovativa, características exigidas pelo ambiente no qual essas empresas atuam (Gerschewski et al. 2015). Ao resolver parte das lacunas nos estudos de ACAP e inovação, esta pesquisa fornece um rol de práticas organizacionais que podem contribuir no avanço da teoria organizacional, assim como auxiliar gestores nas tomadas de decisões acerca do desenvolvimento da ACAP e inovação.

O artigo está organizado da seguinte forma. Após esta introdução, apresenta-se o quadro teórico que explica os fundamentos e proposições da pesquisa. A terceira seção aborda os aspectos metodológicos utilizados neste estudo, a próxima seção explica os resultados e análise dos dados, seguido das considerações finais.

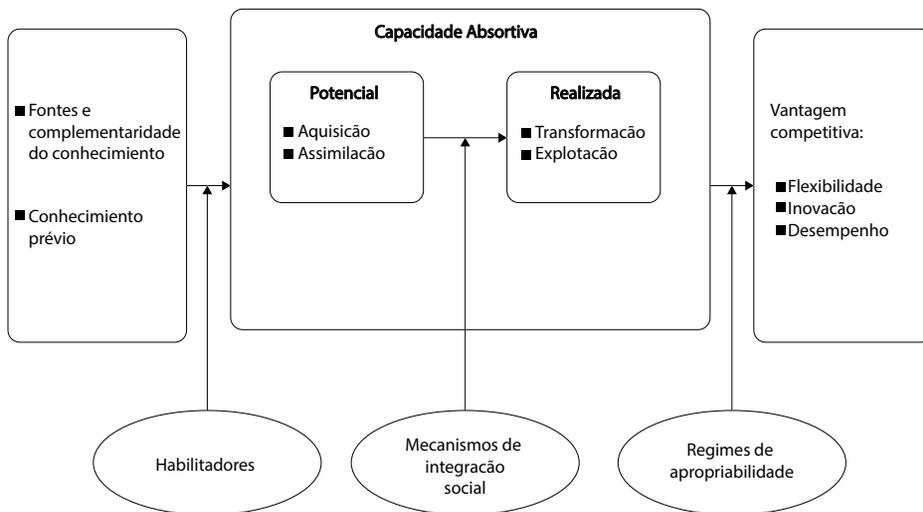
2. MODELO DA CAPACIDADE ABSORTIVA

O conceito da ACAP no contexto organizacional é adaptado por Zahra e George (2002), e analisam a ACAP como o processo pela qual a organização adquire, assimila, transforma e aplica o conhecimento externo, estes construtos é analisado como uma capacidade dinâmica da empresa. Também, estabelecem que a ACAP é “um conjunto de rotinas e processos organizacionais por meio dos quais as empresas adquirem, assimilam, transformam e exploram o conhecimento para produzir uma capacidade dinâmica e sustentável” (Zahra & George, 2002, p. 186).

A proposta dos autores considera um modelo por meio da combinação das quatro capacidades do conhecimento, sendo que as capacidades de aquisição e assimilação do novo conhecimento integram a capacidade absorptiva potencial (PACAP) e a capacidade da transformação e aplicação (*exploitation*) integram a capacidade absorptiva realizada (RACAP), conforme apresenta a figura 1.

Figura 1

Modelo de capacidade absorptiva



Nota. Adaptado de Zahra e George (2002)

Este modelo que distingue a ACAP em dois componentes, conforme Figura 1, demonstra-se útil uma vez que entende-se que as organizações são capazes de compreender bem seus problemas técnicos complexos pela aquisição e assimilação, mas podem não ser capazes de utilizar esse conhecimento para inovar (transformação e aplicação). Os conceitos de ACAP potencial e ACAP realizada envolvem estratégias diferentes, entretanto complementares, estes são apresentado Quadro 1.

Quadro 1

Dimensões da capacidade absorptiva

Componentes	Dimensões/ Capacidades	Definição
Capacidade Absorptiva Potencial	Aquisição	É a capacidade de identificar, valorizar e adquirir conhecimento de fontes externas relevantes para o negócio.
	Assimilação	Capacidade de desenvolver processos e rotinas que permitam que a nova informação ou conhecimento adquirido seja analisado, classificado, interpretado, processado e internalizado.

(continúa)

(continuación)

Capacidade Absortiva Realizada	Transformação	Capacidade de refinar o conhecimento externo adquirido e adequá-lo às rotinas internas, de modo a facilitar a transferência e a combinação de conhecimento prévio com o novo conhecimento adquirido ou assimilado.
	Aplicação	Capacidade de incorporar o conhecimento adquirido a rotinas e processos que criam novas operações, conhecimentos, competências, bens e produtos.

Nota. Tabela elaborada pelas autoras de acordo com de Cohen & Levinthal (1990) e Zahra & George (2002).

Os modelos de Cohen e Levinthal (1990) e Zahra e George (2002) são referências nas pesquisas relacionadas a ACAP, e tornam-se base para os avanços de pesquisas apresentados por Jansen et al. (2005); Lane et al. (2006); Todorova e Durisin (2007); Camisón e Forés, (2010), como novos modelos, e ou instrumentos para mensurar a ACAP ver, (Flatten et al. 2011; Engelman et al. 2017).

As propostas teóricas e empíricas que medem a ACAP, foram propostas por Lane et al. (2006) e Flatten et al. (2011) e sugerem que a ACAP deve ser empiricamente analisada considerando o uso de métricas apropriadas devido a complexidade das dimensões. Na ACAP Potencial, Flatten et al (2011, p. 106) a aquisição é considerada a capacidade que a empresa desenvolve para troca e uso de informações e experiências que surgem do ambiente externo, como também o compartilhamento de informações que acontece para além da organização, destaca-se a importância que o compartilhamento do conhecimento entre diferentes áreas da empresa possa contribuir para o êxito nessa fase.

A capacidade de assimilação é combinação das práticas relacionadas aos processos entre os departamentos que permitam difundir e compartilhar o conhecimento, considerando o fluxo de informação rápido e efetivo, assim a organização dissemina o conhecimento e o incorporar a sua base de conhecimento, (Flatten et al. 2011, p. 106).

Flatten et al. (2011) considera que ambas as capacidades contribuem para promover a renovação contínua do estoque do conhecimento possibilitando e estimulando inovações contínuas, pois quanto maior a capacidade da organização em adquirir e assimilar informações obtidas em fontes externas, maiores as possibilidades para responder a alta concorrência do ambiente externo.

A ACAP Realizada é analisada como a capacidade de transformar o conhecimento por meio da aptidão que os funcionários e ou a empresa estruturam e usam o conhecimento coletado. Assim, Flatten et al. (2011) observam que o aprimoramento das rotinas a partir de instrumentos permitem melhorar o conhecimento existente na organização transformando em vantagem competitiva.

A capacidade de aplicação, é caracterizada como a capacidade que a empresa desenvolve para lançar produtos inovadores no mercado, e refere-se aos resultados das rotinas e aplicação sistemática de processos para a concepção de novos produtos, ou novas formas organizacionais, desenvolvido pela empresa de forma eficaz e inovadora (Flatten et al. 2011, p. 107).

Este estudo utilizou o modelo de Zahra e George (2002) e adaptou as variáveis da escala de Flatten et al. (2011), para elaborar o instrumento de coleta de dados, o que possibilitou identificar as práticas associadas à ACAP nas empresas brasileiras de base tecnológica.

2.1 Capacidade Absortiva e Inovação

Desde os estudos seminais de Cohen e Levinthal (1990), Zahra e George (2002), e Lane et al. (2006) as pesquisas abordam a temática capacidade de absorção do conhecimento, estas são relacionadas com a capacidade da organização impulsionar a sua capacidade inovadora, esta justificativa considera que as inovações são geradas sobretudo a partir da base de conhecimentos já existente na organização e a capacidade em desenvolver e criar mecanismos que possibilitem internalizar o conhecimento externo, gerando melhoria na performance organizacional e inovações (Zahra & George, 2002).

Neste mesmo sentido, Tidd e Bessant (2015) explicam que organizações que melhoram sua capacidade de absorção do conhecimento em treinamento, pesquisa e desenvolvimento, entre outras ações, podem apresentar maior disposição para a inovação e competitividade. Kostopoulos et al. (2011) identificaram que a densidade de fluxos de conhecimento externo aumenta a influência da ACAP no desempenho inovador. A pesquisa de Zahra et al. (2015), ressalta a importância da ACAP junto aos fluxos de conhecimento recebidos do ambiente externo, os quais fortalecem a base de conhecimento da empresa, promovendo a inovação. Também tem sido demonstrado que o desempenho inovador das organizações é influenciado pela ACAP potencial. A pesquisa de Fosfuri e Tribó, (2008) apresenta que o desempenho inovador está relacionado a ACAP potencial, o que difere da pesquisa Cepeda-Carrión et al. (2012) que identifica que o desempenho inovador das organizações é impulsionado pela ACAP realizada.

Estudos recentes e meta-análises de literatura baseados em evidências empíricas apontam para a influência positiva que a capacidade absorptiva exerce sobre a inovação (Gao et al. 2017; Apriliyanti & Alon 2017; Song et al. 2018; Zou et al. 2018). E os achados destas pesquisas propõem que o desenvolvimento de práticas de absorção de conhecimento permite que os membros de uma organização avaliem e adaptem novos conhecimentos externos para gerar novas ideias e potencializar a inovação (Nag & Gioia, 2012; Tseng et al. 2011).

Ainda estudos como Moré, Vargas e Gonçalo (2014) identificaram uma tendência nas pesquisas brasileiras, que tentem analisar as relações entre capacidade absorptiva, gestão do conhecimento e inovação. Neste mesmo sentido, os autores Marlana e Morozini (2017) ressaltam que pesquisas empíricas reconhecem a influência das condições do ambiente externo na inovação, como também no desempenho organizacional, quando relacionado ao entendimento do processo da ACAP em empresas brasileiras dos mais diversos setores.

Um estudo recente desenvolvido por Werlang et al. (2022) em empresas de tecnologia identificou que a criação da capacidade de inovação, depende de fluxos eficazes de informação e do conhecimento em espaços colaborativos, fatores esses que auxiliam no desenvolvimento da ACAP. O estudo também mapeou fatores internos e externos que promovem a ACAP para a inovação, a saber: o conhecimento prévio, as experiências, a liderança, o conhecimento individual, o estilo de gestão, a comunicação interna e P&D e cultura organizacional direcionada a inovação, fatores internos. Já o relacionamento com fornecedores, concorrentes, consumidores e clientes têm o papel relevante no desenvolvimento da ACAP para inovação (Werlang et al. 2022).

Diante do exposto, esta pesquisa considera que o processo da ACAP permite que as empresas explorem novas ideias ou mudem as abordagens atuais e integrem o conhecimento externo ao conhecimento interno da organização para obter melhor Desempenho Inovador.

3. METODOLOGIA

Esta pesquisa adota a abordagem qualitativa, pois de acordo com Merriam (1998) é possível compreender o fenômeno de análise a partir da perspectiva do participante, avaliando o contexto em que ele ocorre, proporcionando analisar sob uma perspectiva integrada. Também se caracteriza como um estudo exploratório e descritivo.

Na etapa exploratória, primeiramente realizou-se uma entrevista semi-estruturada com a coordenadora do setor responsável pelas verticais de atuação da Associação Catarinense de Tecnologia (ACATE), com o objetivo de coletar dados e informações sobre as empresas que caracterizam a população desta pesquisa, a escolha desta associação está atrelada ao fato de que ACATE atua, desde 1986, em prol do desenvolvimento do setor de tecnologia do Estado de Santa Catarina, e se consolidou como uma das principais interlocutoras das empresas catarinenses de tecnologia junto as entidades governamentais e instituições do setor tecnológico, não apenas do Brasil como também em âmbito internacional.

Foram levantadas informações de 165 empresas participantes nestas verticais. Assim, foram iniciados os contatos com os gestores e ou diretores das empresas

selecionadas. Neste processo, foram identificadas 12 verticais, denominadas: conectividade, saúde, segurança, energia, educação, *games* e *apps*, agronegócio, governança e sustentabilidade, manufatura, IOT (indústria e segmentos) e *Fintech* (empresas do segmento financeiro). A criação de verticais de empresas de tecnologia entre as associadas da ACATE tem como objetivo criar grupos de empresas que atuem em mercados semelhantes e complementares, estimulando o associativismo, o compartilhamento de conhecimento e ampliando as relações entre as empresas (ACATE, 2022).

Após contato inicial com as empresas, 11 delas se mostraram interessadas em participar da pesquisa, as quais serão denominadas neste estudo de empresa A1, B2, C3, D4, E5, F6, G7, H8, I9, J10, L11. Dessa forma, foram entrevistados 11 diretores, sócios proprietários e/ou gestores.

O método definido para a pesquisa foi o estudo de casos múltiplos, em concordância com os objetivos do estudo e com pesquisas similares (Hardré & Chen, 2005). A técnica do estudo de caso permite acessar, observar e identificar as informações relevantes neste tipo de pesquisa (Creswell, 2010).

A coleta de dados foi realizada por meio da utilização de um roteiro de entrevista semiestruturada em profundidade, buscas de pesquisa em documentos disponibilizados pelos atores investigados e observação não participante. O roteiro foi desenvolvido com o apoio da pesquisa de Flatten et al (2011), e a partir do modelo teórico proposto por Zahra e George (2002). As categorias de análise foram determinadas por Capacidade absorptiva Potencial: Aquisição e assimilação, e Capacidade Absortiva Realizada: Transformação e Aplicação.

As entrevistas foram transcritas e analisadas por meio da técnica de análise temática que de acordo com Braun e Clarke, (2006), é adequada em estudos qualitativos, pois permite organizar e apresentar os dados de maneira sintética, por meio da identificação, análise e descrições de categorias; e subcategorias, quando necessário. É importante destacar que a unidade de análise são as empresas. Também foi utilizada a técnica de triangulação na análise de documentos eletrônicos, entrevistas e modelo teórico adotado.

4. APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS

Após analisar as entrevistas realizadas, foram descritas as práticas da ACAP identificadas nas organizações objeto deste estudo. Para as análises, foram utilizadas as dimensões propostas por Zahra e George (2002): Aquisição e Assimilação do conhecimento, parte da ACAP Potencial; e Transformação e Aplicação do conhecimento, parte da ACAP Realizada.

No que tange à *aquisição do conhecimento*, percebe-se que é uma das dimensões mais abordadas entre os entrevistados. O desenvolvimento da capacidade de identificar e reconhecer o valor potencial do conhecimento externo disponível no mercado, incentivada pela organização e que ocorre de forma individual, foi citada pelas 11 empresas entrevistadas.

Uma das empresas analisadas (J10), por exemplo, citou a iniciativa de buscar conhecimento externo no mercado, por meio de pessoas que pudessem contribuir para o desenvolvimento das inovações de produtos e processos.

A estratégia da empresa foi buscar o capital humano, o capital intelectual das pessoas, né? Para que estas pudessem contribuir aqui dentro. E também do mercado "As ideias sempre vêm do mercado. A gente tem equipes no cliente, ouvindo cliente né? E aí a gente capta essas ideias, traz para nossa área e a gente começa a processar essas ideias. (Empresa J10)

Neste sentido, literatura previa comenta que os clientes são fontes de conhecimento empresarial. Por meio dos clientes as empresas conseguem adquirir informação e assim responder de forma mais eficaz às demandas do mercado para o curto e longo prazo (Ling-Ching & Wang, 2012).

A empresa F6, explica que a dinâmica de trabalhar com tecnologias, funcionários jovens, e com a criatividade desenvolve uma característica específica das empresas de base tecnológica, que é a busca constante pelo novo, por informações e ferramentas capazes de proporcionar inovações tanto de produto como de processos em seus negócios. Outras pesquisas no Brasil têm evidenciado que motivar as pessoas para buscar e 'trazer' informações externas, aumenta o desempenho inovador das empresas, a motivação dos indivíduos e sua eficiência na gestão e implantação das iniciativas (Dávila et al. 2016).

As sessões de *brainstorming* livre foram mencionadas pela empresa G7, a qual explica que são usadas para incentivar o compartilhamento de ideias, sendo uma prática da organização trocar informações com clientes e colaboradores para buscar o conhecimento externo. O colocado pelos entrevistados coincide com os resultados fornecidos por Kostopoulos et al. (2011), que demonstraram a influência dos fluxos de informação entrantes para alavancar as inovações organizacionais.

No que tange à *assimilação do conhecimento*, observou-se um foco das organizações em implantar processos e práticas orientadas a processar, e internalizar o conhecimento adquirido pelos colaboradores da organização. Embora esta dimensão seja explicada por todas as empresas, observa-se que empresas estabelecidas no mercado por menor tempo, como as empresas E5 e H8, evidenciaram práticas como reuniões periódicas, com o objetivo de alavancar o fluxo de informações entre as diferentes unidades empresariais.

Estes achados estão em linha com Donate e Guadamillas (2011), ao apontar que as empresas precisam fomentar uma cultura aberta, colaborativa e que inspire a confiança. Tudo isso com o intuito de estimular o compartilhamento de ideias e conhecimento entre os colaboradores e também com outras organizações, para assim construir novas práticas e ações que possam responder de forma mais eficaz e rápida às demandas do mercado.

A maioria das empresas analisadas são de base tecnológica, e por possuírem equipes dinâmicas, com colaboradores em nível de doutorado, mestrado e especialistas em suas áreas, também são empresas reconhecidamente inovadoras pelo mercado. Foram observados departamentos de P&D e inovação em sete destas empresas, as quais explicam que desenvolvem a habilidade para vincular o conhecimento a experiências externas do mercado.

Todas as empresas evidenciaram a existência de práticas de compartilhamento de conhecimento, orientadas a favorecer a troca de informações entre pessoas e unidades. As empresas explicam que utilizam ferramentas de base de compartilhamento de conhecimento, porém algumas mais desenvolvidas e aplicáveis. Dentre as 11 empresas entrevistadas, 60 % declaram utilizar ferramentas de metodologia ágil para tomada de decisão, como metodologia SCRUM e SLACK.

Introduzimos uma cultura do Lean Startup, uma mistura de processos ágeis dentro de uma cultura de startup. Ou seja, startup é uma empresa nascente, então é uma pegada muito ágil, muito flexível. Ele simplesmente unifica o conhecimento de todo mundo para um mesmo canal, né? Para chegar no final que é: Essa ideia vale a pena". E que é uma cultura de aprendizado, né?. A gente roda a metodologia Scrum. que possibilitou criarmos um processo e nele nós trouxemos coisas boas de vários processos. (Empresa J10)

O software OKR trabalha com a ideia de (*stretch go*) são metas elásticas. Então sempre são metas de como é que você pode sair da sua zona de conforto. Então isso exige uma busca no mercado sobre melhores práticas, sobre como é que uma empresa faz, o que que ela faz de errado, como é que a gente pode adaptar isso para o nosso negócio, enfim. (Empresa I9)

As empresas associadas a ACATE realizam atividades conjuntas de P&D, o qual é um catalisador de interações. Estas atividades conjuntas, segundo apontado por Fosfuri e Tribo (2008), permitem que cada empresa fortaleça suas capacidades de entender e assimilar melhor o conhecimento externo. Essas interações são muitas vezes suportadas por tecnologia e são catalisadores do compartilhamento de conhecimento, o qual aumenta a contribuição potencial da ACAP para as inovações empresariais (Cepeda-Carrión et al. 2012).

No que tange à *transformação*, constatou-se que existem processos e ou produtos oriundos do conhecimento integrado dos colaboradores, e de clientes

ou *stakeholders* (externos) com potencial de se transformarem em vantagem competitiva para a organização. Nove das empresas confirmam a aplicação de estratégias de reutilização de informações de projetos (lições aprendidas) para tornar mais inovadores os projetos atuais.

O conhecimento recebido dos clientes é fundamental. Nós temos hoje o pessoal do comercial, por exemplo, que fica fazendo entrevistas de valor percebido sobre a nossa solução, as mudanças culturais e o quanto isso impacta dentro da empresa dos clientes. Então a nossa (CS) conversa com os clientes e fica anotando as demandas que o cliente acha necessário, e a gente faz uma reunião para discutir e a gente vai desenvolvendo soluções tecnológicas aí no caso em cima disso. (Empresa I9)

Salter et al. (2014) apontam que esta colaboração com atores externos (clientes e stakeholders) é importante pois cria novos insights ao integrar conhecimentos oriundos de diferentes fontes de informação às quais estes parceiros estão expostos.

Adicionalmente, ficou evidenciado que as 11 empresas utilizam ferramentas para gerenciar o conhecimento dos colaboradores em projetos que atendam as demandas específicas dos mercados, por se tratar em alguns casos de projetos customizados. A mencionada prática acontece de forma personalizada e mais comum.

Se esses *inputs* vão ser impactantes, normalmente a gente vai utilizá-lo, utilizamos nossos arquivos no Google Drive, que a gente organiza e se torna um projeto a gente tem *work shock* de UX, *Used Experience* é uma ferramenta para identificar a experiência do usuário, isto ia ajudar muito a gente na execução de projeto. Este aqui é um resultado (fala mostrando o cartaz). Então, basicamente o que é isto aqui, é a jornada de um usuário que compra um projeto nosso, auxilia entender e modificar nossos processos para ofertar nossos serviços com qualidade. (Empresa D4)

Foi evidenciado que as empresas da ACATE tendem a focar sua estratégia na identificação de novas tecnologias, as quais posteriormente são incorporadas nos seus processos operacionais e nos seus produtos. Este resultado é confirmado por outros estudos em organizações brasileiras de manufatura, as quais focaram em aquisição e desenvolvimento interno de tecnologia para melhorar suas inovações (Goedhuys & Veugelers, 2012).

Em relação à dimensão *aplicação*, esta é definida como a capacidade com que a empresa utiliza o conhecimento para inovar. Aplicação também está referida à capacidade de trabalhar de forma eficaz utilizando novas tecnologias para adaptar o conhecimento adquirido pela empresa. Observou-se que 6 das empresas mais antigas participantes deste estudo, como a A1, B2, C3, D4, J10 e L11 (todas com mais

de 10 anos no mercado), apresentam este processo de forma bastante consistente. Foram identificadas em estas empresas algumas práticas já estabelecidas como o processo de inovação.

A gente acabou montando a área de inovação para atender um novo mercado. E contratamos as melhores pessoas, teoricamente, do mercado, fomentamos o capital intelectual dessa área focada nesses pilares: computação cognitiva, inteligência artificial. Assim conseguimos montar soluções. Por isso que nós já temos, não são protótipos, já são produtos de versão inicial que vão ser demonstrados em primeira mão aqui. Então a gente testa. Se passar por esses crivos, no final a gente avalia a viabilidade dessa ideia como uma oportunidade de negócio. Que vai se tornar um produto inovador. (Empresa J10)

Criamos valor a partir da busca do dado, trazer o dado, aplicar a inteligência ou inserir no processo. A gente criou modelos de negócios que não eram esperados, fazer funcionar de forma diferente, a partir de ferramentas que analisam os dados do cliente, a gente devolve em benefício para o cliente esses dados. (Empresa A1)

É possível constatar aqui que as empresas tiveram a capacidade de criar processos e produtos a partir da aquisição, assimilação e transformação de novos conhecimentos. Gunday et al. (2011) salientam que esta capacidade de inovação da empresa está diretamente relacionada com todas as ações que a organização desenvolve a partir da criação de novos mecanismos, rotinas, procedimentos, promoção do trabalho em equipe, compartilhamento de informações e cultura colaborativa.

Ao final da análise, foi elaborado o Quadro 4, que consolida as principais práticas do processo de ACAP para inovação que foram identificadas a partir dos dados coletados.

Quadro 4*Práticas desenvolvidas para promoção da capacidade absorptiva*

Dimensão	Práticas da ACAP
Capacidade absorptiva potencial	<p>Aquisição</p> <p>Empresa A1: reuniões para trocas de informações externas compartilhando conhecimento, para resolução de problemas.</p> <p>Empresa B2: incentiva os colaboradores a participar de palestras e eventos de inovação como oportunidade de ampliar o conhecimento.</p> <p>Empresa D4: workshops internos para troca de experiências internas entre todos os níveis da organização.</p> <p>Empresa F6: As pessoas procuram informação fora da organização de forma separada e depois eles compartilham essa informação em reuniões internas, selecionando o conhecimento relevante a ser utilizado para o projeto ou tarefa objetivo.</p> <p>Empresa G7: Propõe rotinas e métodos e processos para executar de forma inovadora atividades de gestão da organização.</p> <p>Empresas A1, C3 e L11: promovem eventos com especialistas do setor, colaboradores, clientes e <i>stakeholders</i> com o objetivo de buscar, trocar e acumular informações externas e ampliar a base de conhecimento da empresa;</p> <p>Empresa FI9: investimento em pessoal altamente capacitado, e mão de obra especializada.</p> <p>Empresa J10: incentivam a troca e busca de informações entre os colaboradores.</p>
	<p>Assimilação</p> <p>Empresa A1: Utilização de benchmarking para compreender o desenvolvimento de novas tecnologias de empresas do mesmo setor e de clientes.</p> <p>Empresa B2: Estrutura com fluxos rápidos de troca de informações.</p> <p>Empresas D4, C3, H8, F6 e L11 utilizam softwares e metodologias específicas como <i>Lean startup</i> para difundir o conhecimento em toda a organização.</p> <p>Empresa F6: modelo inovador de gestão organizacional que permite colaboradores compartilharem conhecimento, com rápido fluxo de informação.</p> <p>Empresa J10: Utiliza informações dos clientes para melhoria de produtos e processos, como inovação de produtos.</p> <p>Empresa D4: utiliza base de compartilhamento de conhecimento.</p>
Capacidade absorptiva realizada	<p>Transformação</p> <p>Empresa A1: Adaptação a tecnologias para as necessidades dos clientes.</p> <p>C3: Desenvolvimento de respostas rápidas às mudanças externas ou internas;</p> <p>Empresa I9: busca no mercado melhores práticas e desenvolve soluções tecnológicas inovadoras.</p>

(continúa)

(continuación)

	Transformação	<p>Empresa J10: A empresa foi capaz de transformar as informações em inovação por meio de softwares para atender a demanda específica do cliente.</p> <p>Empresa L11: criou um banco de dados a partir do conhecimento recebido dos clientes e transformou isso em um software com informações capazes de proporcionar melhoria para os negócios de clientes e <i>stakeholders</i>.</p>
Capacidade absorptiva realizada	Aplicação	<p>Empresa A1: lançou produtos inovadores no mercado, a partir da busca do conhecimento externo de pessoas e de mercado.</p> <p>Empresa B2: Aplicou conhecimentos técnicos no desenvolvimento de novos produtos.</p> <p>Empresa C3: desenvolveu e registrou novas tecnologias de forma eficaz para propor novas soluções de inteligência aplicadas ao Trade marketing.</p> <p>Empresa D4: Adaptar o conhecimento externo a uma nova tecnologia para criar um novo produto.</p> <p>Empresa J10 e L11: Desenvolveram a capacidade de aplicar o conhecimento interno e externo em patentes e ou produtos de softwares inovadores para o mercado.</p>

Nossa análise confirmou que as práticas de capacidade absorptiva desenvolvidas pelas organizações podem alavancar as inovações organizacionais, tal como foi proposto por diversos estudos prévios (Zahra & George, 2002; Lane et al., 2006; Cassol et al., 2016). Embora isso, é preciso reconhecer que os estágios de absorção do conhecimento entre as organizações são diferentes, pois algumas são capazes de gerar inovações sustentáveis, e outras não.

Adicionalmente, foi constatado que existem diferentes níveis de absorção do conhecimento relacionados aos níveis de desenvolvimento da organização, como apresentado no quadro 4 e em linha com Flatten et al. (2011), que salienta que o conhecimento organizacional necessita ser compartilhado em um processo dinâmico para que possa conduzir à geração de inovações organizacionais.

No tocante às práticas de inovação, foi observado que as empresas desenvolvem práticas como: inovações de produtos, aprimoramento e melhorias nos softwares com funcionalidades diferenciadas do modelo de negócio atual. Inovações em rotinas, métodos e atividades de gestão das empresas. Desenvolvimento da cultura organizacional voltada para inovação, priorizando práticas de compartilhamento de conhecimento: ambientes colaborativos, e metodologias inovadoras no sistema de informações do gerenciamento organizacional. Portanto, existem oportunidades de melhoria no desempenho inovador, relacionadas a uma melhor aquisição e aplicação do conhecimento tácito, um fato já constatado no estudo de Dávila et al. (2017).

Por fim, e em linha com Cassol et al. (2022) e Werlang et al. (2022), fica evidente com base nos dados analisados, que as organizações com a habilidade de desenvolver a capacidade de absorção do conhecimento apresentam práticas que promovem a inovação, alcançando diferencial competitivo.

5. CONCLUSÕES

Esta pesquisa contribui com a literatura sobre gestão do conhecimento, pois identifica práticas que caracterizam a ACAP nas empresas de base tecnológica no Brasil. Assim, o presente estudo fornece recomendações concretas sobre como alocar os recursos para um melhor aproveitamento do conhecimento externo na geração de valor para as organizações.

Empiricamente, a pesquisa permite analisar quais as práticas de capacidade absorptiva do conhecimento promovem o desenvolvimento sustentável e competitividade das organizações no mercado. Destacam-se algumas práticas, como por exemplo reconhecer o valor das informações externas e estimular a busca por elas, e incorporar o conhecimento nas rotinas organizacionais. Estas práticas conduzem ao fortalecimento das capacidades de aprendizagem, e permitem suportar estratégias para a inovação. As organizações que foram parte deste estudo são cientes que a busca pelo conhecimento interno deve ser valorizada e parte de uma rotina cotidiana.

As evidências aqui apresentadas podem ser influenciadas por diversas ações específicas do setor de tecnologia que tem lugar dentro do ambiente do cluster. Alguns exemplos de estas ações são as estratégias de compartilhamento de conhecimento, estratégias de retenção do conhecimento, promoção de capacitações, incentivos monetários e não monetários, assim como a relação com os clientes, *stakeholders* e universidades. A explicação a isto é que muitas inovações surgem do contato constante entre a organização e outros atores chave que atuam em mercados complementares, sendo por isso fundamental, ao empreender um processo de inovação, saber vincular o conhecimento e experiências de outras organizações.

Os achados estão em linha com o modelo teórico proposto no modelo de Zahra e George (2002). Mais especificamente, nossos resultados confirmam que os constructos de ACAP são multidimensionais. Também foi possível identificar a natureza evolutiva daquelas dimensões, e constatar que as práticas desenvolvidas pelas organizações se fortalecem à medida que as inovações são desenvolvidas e seu valor é reconhecido pelo mercado. Constatou-se também que capacidade de absorção tem diferenças significativas entre empresas do mesmo setor participantes do mesmo cluster. Dessa forma, é possível afirmar que nem todas

as empresas têm a mesma probabilidade de desenvolver inovações ao adquirir, assimilar, transformar e aplicar o conhecimento.

Como toda pesquisa, nosso estudo contribui também com algumas sugestões para estudos futuros. São necessárias pesquisas que avaliem as práticas relacionadas com a ACAP, e seu impacto na inovação, em diferentes redes de cooperação formais e/ou outras formas de relacionamentos intraorganizacionais. Também se considera pertinente desenvolver estudos para compreender a forma na qual se relacionam os componentes da ACAP no nível intraorganizational em outros clusters com propósitos semelhantes ao analisado nesta pesquisa.

REFERÊNCIAS

- Associação Catarinense de Empresas de Tecnologia (2022). *Estatuto Social*. Florianópolis.
- Apriliyanti, I. D., & Alon, I. (2017). Bibliometric analysis of absorptive capacity. *International Business Review*, 26(5), 896-907.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101.
- Cajueta, A. R., & Galina, S. V. R. (2020). Processes in interorganizational relationships to develop absorptive capacity in startups. *Revista de Administração Contemporânea*, 24, 550-566.
- Camisón, C., & Forés, B. (2010). Knowledge absorptive capacity: New insights for its conceptualization and measurement. *Journal of Business Research*, 63(7), 707-715.
- Cassol, A., Marietto, M. L., & Martins, C. B. (2022). Inovação em pequenas e médias empresas: a influência da capacidade de absorção. *Ciências da Administração*, 24(62), 102-121.
- Cassol, A., Gonçalo, C. R., Santos, A., & Ruas, R. L. (2016). A administração estratégica do capital intelectual: um modelo baseado na capacidade absorptiva para potencializar inovação. *Revista Ibero-Americana de Estratégia*, 15(1), 27-43.
- Cassol, A., Zapalai, J., & Cintra, R. F. (2017). Capacidade absorptiva como propulsora da inovação em empresas incubadas de Santa Catarina. *Revista Ciências Administrativas*, 23(1), 9-41.
- Cassol, A., Zanesco, D., Martins, C. B., & Marietto, M. L. (2019). Capacidade absorptiva como moderadora da relação entre inovatividade organizacional e desempenho inovador de pequenas e médias empresas brasileiras. *Interciencia*, 44(1), 15-22.

- Cepeda-Carrion, G., Cegarra-Navarro, J. G. & Jimenez-Jimenez, D. (2012). The effect of absorptive capacity on innovativeness: Context and information systems capability as catalysts. *British Journal of Management*, 23(1), 110–129.
- Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: a new perspective on learning and innovation. *Administrative Science Quarterly*, 35(1), 128-152.
- Creswell, J. W. (2010). Projeto de pesquisa: métodos qualitativo. *Quantitativo e Misto*, 3.
- Dávila, G. A., North, K., & Varvakis, G. (2016). How Brazilian textile enterprises learn to grow. Em *Competitive Strategies for Small and Medium Enterprises* (pp. 241-254). Springer.
- Dávila, G. A., Durst, S., & Varvakis, G. (2017). Knowledge absorptive capacity, innovation, and firm's performance: insights from the south of Brazil. *International Journal of Innovation Management*, 475-508.
- Dávila, G., Varvakis, G., & North, K. (2019). Influence of strategic knowledge management on firm innovativeness and performance. *BBR. Brazilian Business Review*, 16(3), 239-254.
- Da Silva, F. M., & da Costa, P. R. (2019). Desenvolvimento da Capacidade Absortiva em Empresas de Base Tecnológica. *Revista Inovação, Projetos e Tecnologias*, 7(1), 36-51.
- Donate, M.J., & Guadamillas, F. (2011), Organizational factors to support knowledge management and innovation. *Journal of Knowledge Management*, 15(6), 890-914.
- Engelman, R. M., Fracasso, E. M., Schmidt, S., & Zen, A. C. (2017). Intellectual capital, absorptive capacity and product innovation. *Management Decision*, 55(3), 474-490.
- Flatten, T. C., Engelen, A., Zahra, S. A., & Brettel, M. (2011). A measure of absorptive capacity: scale development and validation. *European Management Journal*, 29(2), 98-116.
- Fosfuri, A., & Tribó, J. A. (2008). Exploring the antecedents of potential absorptive capacity and its impact on innovation performance. *Omega*, 36(2), 173-187.
- Gao, S., Yeoh, W., Wong, S. F., & Scheepers, R. (2017). A literature analysis of the use of Absorptive Capacity construct in IS research. *International Journal of Information Management*, 37(2), 36-42.
- Garrido, I. L., Parente, R. C., Gonçalves, C. R., & Vasconcellos, S. L. D. (2017). Remaining innovative: the role of past performance, Absorptive Capacity, and internationalization. *Brazilian Business Review*, 14(6), 559-574.

- Gerschewski, S., Rose, E. L., & Lindsay, V. J. (2015). Understanding the drivers of international performance for born global firms: an integrated perspective. *Journal of World Business, 50*(3), 558-575.
- Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal, 17*(S2), 109-122.
- Goedhuys, M., & Veugelers, R. (2012). Innovation strategies, process and product innovations and growth: firm-level evidence from Brazil. *Structural change and economic dynamics, 23*(4), 516-529.
- Gunday, G., Ulusoy, G., Kilic, K., & Alpkan, L. (2011). Effects of innovation types on firm performance. *International Journal of Production Economics, 133*(2), 662-676.
- Hardré, P. L., & Chen, C. H. (2005). A case study analysis of the role of instructional design in the development of teaching expertise. *Performance Improvement Quarterly, 18*(1), 34-58.
- Jansen, J. J. P., Van Den Bosch, F. A. J., & Volberda, H. W. (2005). Managing potential and realized absorptive capacity: how do organizational antecedents matter? *Academy of Management Journal, 48*(6), 999-1015.
- Kostopoulos, K., Papalexandris, A., Papachroni, M., & Ioannou, G. (2011). Absorptive capacity, innovation, and financial performance. *Journal of Business Research, 64*(12), 1335-1343.
- Lane, P. J., Koka, B. R., & Pathak, S. (2006). The reification of absorptive capacity: a critical review and rejuvenation of the construct. *Academy of Management Review, 31*(4), 833-863.
- Ling-Ching, C. A., & Wang, W. Y. (2012). The causal relationships between aspects of customer capital. *Industrial Management & Data Systems, 112*(6), 848-865.
- Manthey, N. B., Verdinelli, M. A., Rossetto, C. R., & Carvalho, C. E. (2016). *O Impacto das Capacidades de Absorção e de Inovação no Desempenho da Inovação de Produto* [Doctoral dissertation, Dissertação (Mestrado)].
- Marlana, C., & Morozini, J. F. (2017). *Tendências de pesquisa nacionais em capacidade absorptiva: Uma análise bibliométrica e redes sociais em grupos de pesquisa*. Anais do VI Simpósio internacional de gestão de projetos, inovação e Sustentabilidade (SINGEP). São Paulo/SP, 1-18.
- Merriam, S. B. (1998). *Qualitative research and case study applications in education. revised and expanded from case study research in education*. Jossey-Bass Publishers.

- More, R. P. O., Vargas, S. M. L., & Gonçalo, C.R. (2014). Interfaces Da Capacidade Absortiva Numa Perspectiva Organizacional. *Revista Inova Ação*, 3(2), 30-52.
- Nag, R., & Gioia, D. A. (2012). From Common to Uncommon Knowledge: Foundations of Firm-Specific Use of Knowledge as a Resource. *Academy of Management Journal*, 55(2), 421-457.
- Nagano, M. S., J. P. Stefanovitz, J. P., & Vick, T. E. (2014). Innovation management processes, their internal organizational elements and contextual factors: An investigation in Brazil. *Journal of Engineering and Technology Management*, 33, 63-92.
- Salter, A., Ter Wal, A. L., Criscuolo, P., & Alexy, O. (2014). Open for ideation: individual-level openness and idea generation in R&D. *Journal of Product Innovation Management*, 32(4), 488-504.
- Siachou, E., Vrontis, D., & Trichina, E. (2021). Can traditional organizations be digitally transformed by themselves? The moderating role of absorptive capacity and strategic interdependence. *Journal of Business Research*, 124, 408-421.
- Song, Y., Gnyawali, D. R., Srivastava, M. K., & Asgari, E. (2018). In Search of Precision in Absorptive Capacity Research: A Synthesis of the Literature and Consolidation of Findings. *Journal of Management*, 44(6), 2343-2374.
- Takeuchi, H., & Nonaka, I. (2009). *Gestão do conhecimento*. Bookman.
- Tidd, J., & Bessant, J. (2015). *Gestão da inovação*. 5. Bookman.
- Torodova, G., & Durisin, B. (2007). Absorptive capacity: valuing a reconceptualization. *Academy of Management Review*, 32(3), 774-786.
- Tseng, C., Chang Pai, D., & Hung, C. (2011). Knowledge absorptive capacity and innovation performance in KIBS. *Journal of Knowledge Management*, 15(6), 971-983.
- Werlang, N. B., Silva, M., & Candido, A. C. (2022). Capacidade Absortiva para Inovação: análise dos fatores internos e externos em startups catarinenses. *Comunicação & Inovação*, 23(52), 16-33.
- Zahra, S. A., & George, G. (2002). Absorptive capacity: A review, reconceptualization, and extension. *Academy of Management Review*, 27(2), 185-203.
- Zahra, S. A., Larraneta, B., & Galán, J. L. (2015). Absorptive Capacity and Technological Innovation. *Wiley Encyclopedia of Management*, 13, 1-5.
- Zou T., Ertug G., & George G. (2018). The capacity to innovate: a meta-analysis of absorptive capacity. *Innovation: Management, Policy and Practice*, 20(2), 87-121.

CAPITAL HUMANO: A CHAVE DA GESTÃO DO CONHECIMENTO EM ORGANIZAÇÕES PÚBLICAS - UMA ABORDAGEM INTEGRATIVA

DANISSON LUIZ DOS SANTOS REIS

danisson.sergipe@gmail.com

<https://orcid.org/0000-0002-5656-6519>

Universidade Federal de Santa Catarina, Brasil

PAULO ROBERTO MOURA

paulormoura1@hotmail.com

<https://orcid.org/0000-0003-3434-7443>

UNIASSELVI, Brasil

GREGÓRIO VARVAKIS

g.varvakis@ufsc.br

<https://orcid.org/0000-0003-2576-4835>

Universidade Federal de Santa Catarina, Brasil

NERI DOS SANTOS

nerisantos@gmail.com

<https://orcid.org/0000-0002-0356-6750>

Universidade Federal de Santa Catarina, Brasil

RESUMO. O objetivo do presente artigo é investigar como a pesquisa acadêmica aborda a relação entre o capital humano, a gestão do conhecimento e o setor público. Para tal, foi realizada uma revisão integrativa com termos chaves ligados às três temáticas citadas, sendo conduzida com dados importados da base Scopus. Entre os resultados alcançados, foram identificadas seis clusters de abordagem temática que, após a referida análise, possibilitaram a identificação das formas como os recursos humanos contribuem para a gestão do conhecimento nos setores públicos, bem como as consequências para as pessoas decorrentes da implementação desta. Este artigo contribui para o enriquecimento das pesquisas que investigam a importância do fator pessoas sobre a gestão do conhecimento em organizações públicas, promovendo a síntese dos estudos mais recentes sobre o tema.

PALAVRAS CHAVE: gestão do conhecimento, compartilhamento do conhecimento; setor público, pessoas, capital humano

HUMAN CAPITAL: THE KEY TO KNOWLEDGE MANAGEMENT IN PUBLIC ORGANIZATIONS. AN INTEGRATIVE APPROACH

ABSTRACT. This paper investigates how academic research addresses the relationship between human capital, knowledge management, and the public sector. To this end, an integrative review was carried out with key terms related to the three mentioned themes and conducted with data imported from the Scopus database. Among the results achieved, six clusters of thematic approach were identified. After the analysis mentioned above, it was possible to identify how human resources contribute to knowledge management in the public sector and the consequences for people resulting from its implementation. This article contributes to research on the importance of people in knowledge management in public organizations, promoting the synthesis of the most recent studies on the subject.

KEYWORDS: knowledge management, knowledge sharing, public sector, people, human capital

CAPITAL HUMANO: LA CLAVE DE LA GESTIÓN DEL CONOCIMIENTO EN LAS ORGANIZACIONES PÚBLICAS. UN ENFOQUE INTEGRADOR

RESUMEN. Este artículo indaga sobre cómo la investigación académica aborda la relación entre el capital humano, la gestión del conocimiento y el sector público. Para ello, se realizó una revisión integradora con términos clave relacionados con los tres temas mencionados y realizada con datos importados de la base de datos Scopus. Entre los resultados alcanzados, se identificaron seis clústeres de abordaje temático. Luego del análisis, fue posible identificar cómo los recursos humanos contribuyen a la gestión del conocimiento en el sector público y las consecuencias para las personas, derivadas de su implementación. Este artículo contribuye a la investigación sobre la importancia de las personas en la gestión del conocimiento en las organizaciones públicas, promoviendo la síntesis de los estudios más recientes sobre el tema.

PALABRAS CLAVE: gestión del conocimiento, intercambio de conocimientos, sector público, personas, capital humano

1. INTRODUÇÃO

Com a automação e a informatização acarretadas pela introdução da computação e da internet no dia a dia, teve-se a formulação da sociedade 4.0, a era da informação, onde o conhecimento começou a se tornar o novo insumo necessário que move a economia. Atualmente, uma recente revolução provocada pela transformação digital mobiliza o globo em direção à sociedade 5.0, a era da criatividade e da imaginação, onde o conhecimento é o novo “ouro negro”, e geri-lo se torna essencial para todas as atividades humanas (Da Silva et al., 2021; Keidarenk, 2017).

Gestão do Conhecimento (GC) é o processo em que todos os tipos de conhecimentos são identificados, capturados e codificados para que possam ser difundidos, agregando benefícios para qualquer organização (Cooper, 2017). A GC já é reconhecida como fator chave na melhoria da eficiência interna e competitividade nas organizações privadas. Entretanto, o setor público possui um contexto diferente, pois as motivações são distintas dos fins lucrativos, já que a diretiva é atender os cidadãos com rapidez, segurança e eficiência (Baporikar, 2014). Um ponto crucial para a GC no setor público é a descontinuidade da prestação de serviços devido à crescente crise de capital humano em muitas organizações públicas devido à saída, demissão ou aposentadoria, exigindo uma captura mais eficaz do conhecimento para minimizar a sua perda (Pee & Kanhanhalli, 2016; Dewah & Mutula, 2014; Willem & Buelens, 2006). As pessoas desempenham um papel chave no contexto organizacional da GC, pois elas são líderes, desempenham os processos organizacionais e utilizam as tecnologias disponíveis. Da mesma forma, os conhecimentos tácitos que possuem são cruciais para diferenciação e especialização de cada organização (Bell DeTienne et al., 2004).

Desta prerrogativa, surge um questionamento: qual o impacto do componente pessoas para a gestão do conhecimento dentro de organizações do setor público? A fim de responder a esta questão, o presente artigo buscou respostas na literatura acadêmica. Para isso, os autores conduziram uma revisão integrativa com o objetivo de investigar como a pesquisa acadêmica aborda a relação entre o capital humano, a GC e o setor público.

2. REFERENCIAL TEÓRICO

Para que exista uma melhor compreensão sobre gestão do conhecimento, é necessário entender o que é conhecimento. Polanyi (1967, p. 11) afirmava que “nós sabemos mais do que somos capazes de expressar”. Na visão de Stewart (1998, p. 13) “a informação e o conhecimento são as armas nucleares da nossa era”. Probst et al. (2002) definem conhecimento como um conjunto total que inclui cognição e habilidades que os indivíduos utilizam para resolver a gestão do conhecimento.

Gruzman e Siqueira (2007) explicam que “o conhecimento é uma construção social que só ganha sentido quando circula publicamente e se coloca a serviço das comunidades” (p. 12), enquanto Koskien (2013) explana que o “conhecimento é a compreensão humana de um campo especializado de interesse que foi adquirido por meio de estudo e experiência” (p. 137). Por sua vez, Baporikar (2014) afirma que o conhecimento é a nova commodity.

Este artigo usará a definição de conhecimento presente no European Guide to Good Practice in Knowledge Management (2004) que o conceitua como a combinação de dados e informações, somadas as opiniões de especialistas, habilidades e experiência, resultando em um ativo valioso que pode ser utilizado no processo de tomada de decisão, e que pode ser considerado explícito ou tácito, individual ou coletivo.

Nonaka e Takeuchi (1997) definem conhecimento tácito como aquele advindo das experiências, sendo subjetivo e pessoal; já o explícito é aquele que pode ser compartilhado por palavras, dados e informações, sendo mais racional.

A partir da definição proposta no artigo, percebe-se que o conhecimento é um insumo com infinitas possibilidades de outputs, dessa forma, é necessária sua gestão para a sua correta utilização e a maximização da geração de valor. Assim entendido, surge a GC com o objetivo de gerir atividades de criação, captura, conversão e uso do conhecimento, como proposto por Wiig (1997), propondo a planejar, implementar, operar e monitorar as atividades e projetos relacionados ao conhecimento que são demandados para uma gestão bem-sucedida do capital intelectual.

Diante da variedade de conceitos que podem ser encontrados na literatura, o presente artigo utilizará o conceito de Santos e Varvakis (2021) que estabelecem a GC como a gestão das atividades e dos processos organizacionais que promovem o conhecimento organizacional para o aumento da competitividade, por meio do melhor uso e da criação de fontes de conhecimento individuais e coletivas. Apesar da GC ser uma temática relativamente recente (Serenko & Bontis, 2013), é importante compreender que ela envolve temas relevantes para as organizações, tais como: desempenho, competências, liderança e inovação, levando estas a sua eficiência (Martinsons et al., 2017; Dalkir, 2013).

Seja em organizações públicas ou privadas, a GC permite o aumento da competitividade, com a possibilidade de quebrar paradigmas, exigindo assim um novo alinhamento das organizações com seus stakeholders (Nonaka & Takeuchi, 1997). Na compreensão de Moreira et al. (2022), é certo que:

seja no setor público ou privado, na Sociedade do Conhecimento novas configurações organizacionais emergem, apoiadas em novos comportamentos e

paradigmas, para que as organizações e seus pares estejam preparados coletivamente para o ambiente complexo e dinâmico no qual estão inseridos. Neste novo modelo de sociedade é preciso gerenciar o conhecimento de maneira que o impacto de suas ações seja ampliado. (p. 31)

Entende-se por setor público as organizações e instituições, como agências, autoridades locais, paraestatais, forças armadas, secretarias e afins, que sob controle do Estado prestam serviços aos cidadãos. Esse setor não visa o lucro, mas sim o impacto positivo para a sociedade. É caracterizado por sistemas administrativos burocráticos, transparentes e impessoais, bem como procura o combate ao nepotismo, à discriminação e ao preconceito (Dewah & Mutula, 2014).

As organizações do setor público são caracterizadas como intensivas em conhecimento por gerarem uma quantidade substancial de conhecimentos tanto tácitos – modelos mentais, crenças, intuições, palpites, insights, brainstormings, conversas nos intervalos e refeições – quanto explícitos – bancos de dados, regras, regulamentos, memorandos, manuais de procedimentos, registros em fóruns (Dewah & Mutula, 2014).

Cada vez mais, o setor público é pressionado em direção à agilidade e à rapidez nas respostas, exigindo responsabilidade e capacidade de feedbacks rápidos, por haver possibilidade de obsolescência de determinadas estratégias governamentais devido ao uso não eficaz do conhecimento para fazer frente aos desafios da atualidade (Massaro et al., 2015). Em resumo, a necessidade e a importância da GC para o setor público são esmagadoras, portanto, a questão agora não é mais se o setor público precisa de GC, mas sim como eles podem efetivamente colher os benefícios desta (OCDE, 2007).

Por exemplo, a negligência da gestão do conhecimento no setor público pode acarretar inúmeros desafios como a necessidade de adaptação frente a maior quantidade de informações transacionadas entre instituições. À medida que as organizações públicas usam cada vez mais tecnologias para colaborar entre si, emerge a necessidade latente de desenvolver capacidades de compartilhamento, aplicação e criação de conhecimento, principalmente, em atuações em redes, sejam elas nacionais ou internacionais, ou na colaboração para a resolução de questões globais críticas (Pee & Kanhanhalli, 2016).

O capital humano é um dos componentes essenciais para o sucesso das práticas de GC, independentemente do tipo de organização e junto a processos, liderança e tecnologia formam os quatro blocos essenciais para o impacto positivo da GC (Nair & Prakash, 2009). As pessoas são responsáveis por um dos elementos-chaves da GC: o compartilhamento do conhecimento, que é fundamental para a conversão do conhecimento e transferência das expertises para o status de conhecimento

organizacional. Além disso, o capital relacional e a cultura organizacional, presentes entre as pessoas de uma dada organização, são fatores cruciais para o sucesso das estratégias de GC (Pee & Kanhanhalli, 2016; Servin & De Brun, 2005).

3. METODOLOGIA

Em busca de alcançar o objetivo de compreender como a pesquisa acadêmica aborda a relação entre o capital humano, a GC e o setor público, este artigo utilizou uma revisão integrativa.

A escolha da revisão integrativa se efetuou porque ela possibilita traçar uma análise sobre o conhecimento construído sobre o tema abordado no objetivo, o que possibilita um norte para o desenvolvimento de novos conhecimentos (Torraco, 2005; Whittemore & Knafl, 2005). Também pode se justificar o uso da revisão integrativa, porque é um método que permite criar conhecimento a partir das contradições encontradas e identificar gaps outrora não visualizados (Botelho et al., 2011).

Durante a análise do corpus encontrado, também foram utilizadas técnicas de revisão bibliométrica, com a utilização do software Vosviewer, a fim de compreender melhor as informações e dados presentes nos documentos. Assim, possibilitou a identificação da correlação das produções, dos países que estão difundindo o tema, da escala temporal das publicações e quais são os artigos mais relevantes (Araújo, 2006; Santos & Kobashi, 2009).

O string de busca utilizado foi: ((“knowledge management”) AND (“public service” OR “public administration” OR “public sector” OR “public management”) AND (“people” OR “employee” OR “server” OR “worker” OR “human resources” OR “culture”)). Estes termos chaves foram identificados nos títulos, resumos e palavras-chaves. O string selecionado buscou varrer publicações que navegassem pelos três pilares que compõem o alcance do objetivo do artigo: gestão do conhecimento/ serviço público/pessoas.

Para a construção da pesquisa, foi utilizado o repositório Scopus em junho de 2022. A escolha desta base de dados aconteceu devido à sua visão multidisciplinar e integrada contendo fontes relevantes para corroborar com o presente estudo (Neves et al., 2018). Salienta-se que a Scopus possui periódicos nacionais e internacionais com bons índices de quartil e ótimas notas de Impact Factor (FI).

Os critérios de inclusão estabelecidos para revisão integrativa foram:

- Escala temporal dos últimos 10 anos, a fim de compreender as publicações mais recentes sobre o tema e os principais gaps relacionados ao objetivo do presente artigo.

- Permanência de artigos nas línguas inglesa e portuguesa.
- Apenas artigos de periódicos, pela maior possibilidade do uso de double blind review, trazendo isonomia e creditando maior rigorosidade no conhecimento abordado nos artigos.

Após a extração de dados na base Scopus com o string já descrito, foi conduzida a organização dos dados em planilha eletrônica para eliminar redundâncias. Em seguida, foi realizada a leitura exploratória dos títulos, resumos e palavras-chaves, só permanecendo aqueles com aderência ao objetivo do presente artigo.

Posteriormente, foi conduzida uma leitura em profundidade dos artigos que permaneceram, a fim de extrair dados e informações para revisão. Ainda nesta etapa, alguns artigos foram retirados por não contribuírem para o alcance do objetivo final. A Tabela 1 apresenta os números de artigos presentes no corpus após cada uma destas etapas.

Tabela 1

Processo de revisão integrativa

Etapas realizadas	Scopus
Total das buscas	109
Após exclusões das redundâncias	108
Após leitura exploratória	35
Após leitura em profundidade	25

Por fim, é destacado que este artigo pode ser classificado como uma pesquisa qualitativa, pois os autores buscaram interpretar as informações disponíveis no corpus final de forma indutiva e pelas “lentes teóricas” disponíveis, a fim de possibilitar o alcance do objetivo traçado no estudo (Creswell, 2010).

4. RESULTADOS E DISCUSSÕES

A partir do corpus final, com 25 artigos que foram lidos na íntegra, algumas características foram identificadas quanto à origem dos artigos. A Tabela 2 consta os quantitativos de origem por países.

Tabela 2*Origem dos artigos do corpus final*

País de Origem	Quantitativo	Artigos
Índia	05	Tripathi et al. (2021); Lartey et al. (2021); Dey e Mukhopadhyay (2018); Baporikar (2017); Pandey e Dutta (2015).
Malásia	03	Ganapathy et al. (2019); Bakar et al. (2015); Salleh (2013).
Brasil/Alemanha	02	De Angelis (2016; 2013);
Itália	02	Bertolino, et al. (2020); Torre et al. (2020)
África do Sul	02	Ncoyini e Cilliers (2020); May e Fombad (2019).
Paquistão	02	Razzaq et al. (2019); Muqadas et al. (2017).
Colômbia, Estados Unidos, países da África subsaariana, Gana, Palestina, Cingapura, Paquistão, Países Baixos, Bélgica, Grécia e Austrália	01 artigo cada	Fischer e Döring (2021); Castaneda et al. (2016); Dewan e Mutula (2016); Pee e Kanhanhalli (2016) Almudallal et al. (2015); Taskin e Bunnen (2015); Boateng e Agyemang (2015); Tsirikas e Katsaros (2014); Amayah (2013).

Como demonstrado na Tabela 2, o país com o maior número de publicações é a Índia, sendo interessante atentar que dos cinco artigos identificados, dois analisam a NTPC, organismo estatal responsável pela geração de eletricidade (Baporikar, 2017; Pandey & Dutta, 2015); e três focam em órgãos públicos diversos (Tripathi et al., 2021; Lartey et al., 2021; Dey & Mukhopadhyay, 2018).

Um fato interessante é a presença, no corpus final, de apenas um artigo dos Estados Unidos, sendo que, segundo Wang et al. (2018), este país domina o ranking de publicações relacionadas à gestão do conhecimento, enquanto a Índia aparece em décimo sexto lugar. Pode-se teorizar que há menos publicações estadunidenses focadas na relação da GC com o setor público.

No que diz respeito às publicações ao longo do tempo, visível na Tabela 3 e na Figura 1, percebe-se uma distribuição uniforme nos últimos dez anos, com pico de publicação em 2015. Observa-se que o corpus acompanhou a tendência mundial de publicações sobre GC. Conforme estudo de Wang et al. (2018), o acúmulo de número de publicações do tema indica que o campo de pesquisa correlacionado à temática está relativamente maduro desde 2013.

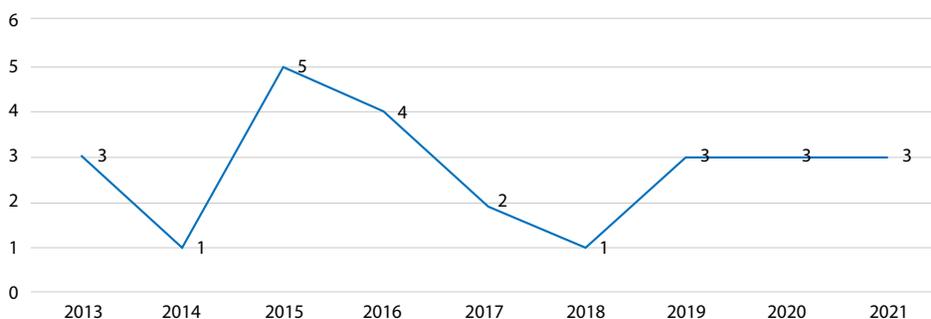
Tabela 3

Ano de publicação dos artigos do corpus final

	2013	2014	2015	2016	2017	2018	2019	2020	2021
Artigos	03	01	05	04	02	01	03	03	03

Figura 1

Ano de publicação dos artigos do corpus



Em análise das palavras-chaves do corpus, destacam-se os termos knowledge management e knowledge sharing, conforme ilustrado nas Figuras 2 e 3. Este resultado também acompanha a tendência global do tema, apontada por Wang et al. (2018), ou seja, inclusive quando realizado um recorte em um dos campos de aplicação da GC, estas palavras-chaves continuam sendo relevantes.

Evidencia-se também que, nos últimos anos, os termos innovation, creativity, organizational learning, social capital theory e employee(s) satisfaction vêm crescendo e demonstram que as pesquisas que correlacionam temas de gestão do conhecimento/serviço público/pessoas estão mais preocupadas em entender como o capital intelectual pode fomentar um setor público mais competitivo e satisfatório para seus colaboradores.

Todavia, os termos barriers, organizational culture e culture estão sendo menos usados. Isso demonstra que, no passado, houve um ápice de publicações que analisavam o impacto da cultura para GC, principalmente no setor público. Atualmente, este impacto já é sapiente da comunidade acadêmica, por isso, as pesquisas estão se direcionando para entender e compreender outros pontos da GC.

Figura 2

Clusters de palavras-chaves dos artigos do corpus final - foco no termo knowledge management

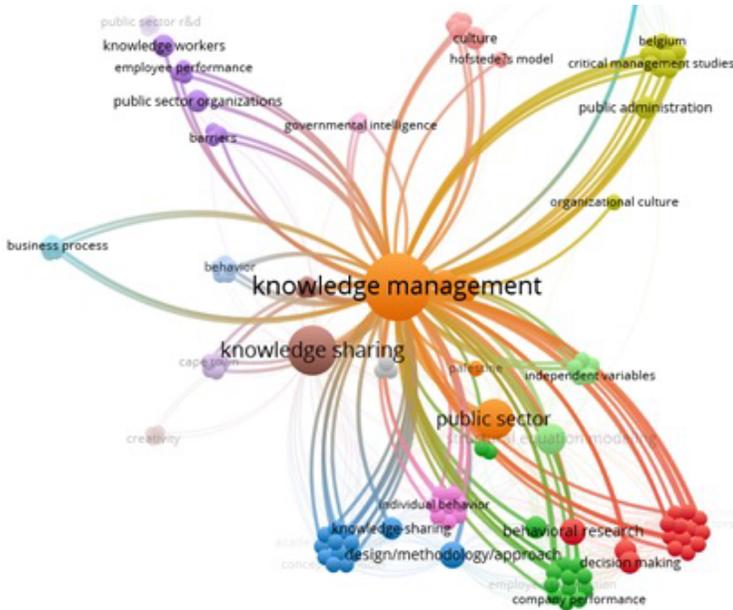
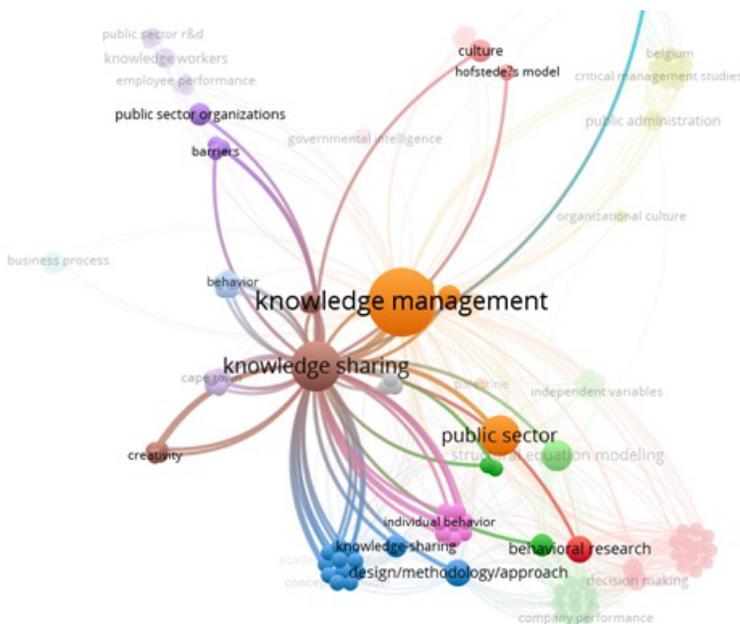


Figura 3

Clusters de palavras chaves dos artigos do corpus final - foco no termo knowledge sharing



Apresenta-se também, nesta análise inicial, os documentos mais citados dentre os que formaram o corpus final. Os cinco artigos mais citados estão descritos na Tabela 4. Destaque para obra de Amayah (2013) que busca, a partir de uma pesquisa quantitativa, determinar quais são as barreiras, os facilitadores e os motivadores do compartilhamento de conhecimento entre funcionários públicos estadunidenses.

Tabela 4

Artigos mais citados do corpus final

Referência do artigo	Quantidade de citações
Amayah (2013)	223
Pee e Kanhanhalli (2016)	66
Muqadas et al. (2017)	56
Razzaq et al. (2019)	50
Castaneda et al. (2016)	16

Ao analisar as abordagens na condução das pesquisas presentes no corpus final, tem-se uma maioria de documentos (16 no total) com pesquisas quantitativas do tipo levantamento. Como a função delas é a generalização de opiniões, tendências ou atitudes, a partir de uma amostra (Creswell, 2010), justifica-se a difusão desta abordagem. Os documentos do corpus visam descrever, comparar ou explicar se determinados comportamentos e atitudes se reproduzem no universo das organizações públicas, a partir da testagem de hipóteses ou análise de correlação de múltiplas variáveis.

Para esta análise de dados, o método de modelagem de equações estruturais é o mais presente. Este método de estatística multivariada é amplamente utilizado nas ciências sociais e é relevante, pois permite que sejam estimados os efeitos diretos e indiretos de variáveis independentes sobre variáveis dependentes (Neves, 2018).

Os demais artigos são oriundos de abordagens qualitativas buscando interpretar como os conceitos e construtos sobre a GC são percebidos, a fim de criar um retrato holístico deste cenário ao fim (Creswell, 2010).

4.1 Clusterização dos artigos

Conforme preconizado pelo método integrativo (Torraco, 2005; Whittemore & Knafl, 2005), após a leitura em profundidade dos artigos do corpus, estes foram clusterizados conforme temas centrais a fim de facilitar a extração de informações, a geração de conhecimento e a identificação dos pontos chaves e gaps. A Tabela 5 consta o resultado deste processo.

Tabela 5*Clusterização dos artigos do corpus final*

Tema central	Artigos
Compartilhamento de Conhecimento	Fischer e Döring (2021); Lartey et al. (2021); May e Fombad (2019); Dey e Mukhopadhyay (2018); Muqadas et al. (2017); Castaneda et al. (2016); Amayah (2013); Salleh (2013).
Recursos Humanos como facilitadores da GC	Torre et al. (2020); Ncoyini e Cilliers (2020); Ganapathy et al. (2019); Dewan e Mutula (2016); Pee e Kanhanhalli (2016); Bakar et al. (2015); Pandey e Dutta (2015).
Consequências positivas da GC para os servidores	Bertolino et al. (2020); Razzaq et al. (2019); Baporikar (2017); Tsirikas e Katsaros (2014);
Liderança	Tripathi et al. (2021); Almudallal et al. (2015).
Cultura organizacional	De Angelis (2016; 2013); Boateng e Agyemang (2015).
Consequências negativas da GC para os servidores	Taskin e Bunnan (2015).

O primeiro cluster é composto pelos documentos que abordam o processo de compartilhamento de conhecimento e como este é crucial entre colaboradores de organizações públicas. Entre os oito artigos deste cluster, dois são estudos qualitativos (May & Fombad, 2019; Muqadas et al., 2017).

May e Fombad (2019) investigam o compartilhamento de conhecimento em uma municipalidade na África do Sul que reconhece a importância deste processo para desencadear melhores benefícios para a população. Porém, carece de estratégias melhores, pois o compartilhamento ainda não acontece a contento.

Já Muqadas et al. (2017) investigam os motivos que impedem o compartilhamento de conhecimento em universidades públicas, incluso quando há incentivos para essa finalidade. Foi constatado que os funcionários acumulam conhecimento, porque desejam manter poder, influência, autoridade ou superioridade, além da ausência de uma cultura que possua como elementos centrais a comunicação eficiente, a socialização, a confiança e o networking.

Os demais artigos deste cluster são quantitativos e buscaram comprovar hipóteses teóricas difundidas na literatura. Dey e Mukhopadhyay (2018), e Castaneda et al. (2016) investigam setores públicos na Índia e na Colômbia respectivamente e conseguem validar a hipótese que a intenção de compartilhar de conhecimento (KSI) afeta diretamente o comportamento de compartilhar de conhecimento (KSB). Os artigos demonstram que o contexto do trabalho e a percepção que os servidores públicos têm do suporte organizacional também alavancam o KSB. Castaneda et al. (2016) ainda postula que a influência do KSI sobre o KSB é resultado da normatização do poder público.

Lartey et al. (2021), e Fischer e Döring (2021) corroboram, em seus estudos, a validação da hipótese da importância do contexto do trabalho para impulsionar o compartilhamento do conhecimento entre colaboradores do poder público, destacando o potencial das redes relacionais formadas no trabalho. Inclusive, Fischer e Döring (2021) validam o efeito positivo do compartilhamento do conhecimento sobre a satisfação com o trabalho. Desta forma, estes autores corroboram construtos teóricos, amplamente disseminados entre acadêmicos e pesquisadores, sobre a gestão do conhecimento.

Entretanto, Amayah (2013) e Salleh (2013) trazem importantes descobrimentos que diferem do preconizado na literatura clássica. Amayah (2013) mostrou que benefícios pessoais, reciprocidade e confiança não são variáveis que afetam positivamente o compartilhamento do conhecimento, sendo que a própria autora enfatiza que essas validações vão de encontro com o estabelecido na literatura. Já Salleh (2013) estabelece que treinamentos e oportunidades de aprendizagem não influenciam positivamente o compartilhamento do conhecimento, outro achado que também diverge do preconizado na literatura clássica sobre o tema apontado pelo autor.

Outro dado divergente é que diferente de Castaneda et al. (2016), que aponta que a natureza normativa do setor público influencia o compartilhamento de conhecimento, Amayah (2013) estabelece que a natureza burocrática do serviço público é o limitador deste processo.

Interessante notar que grande parte dos autores deste cluster reforçam que, independente das condições sociais favoráveis para o compartilhamento do conhecimento dentro de órgãos públicos, existe a necessidade de artefatos tecnológicos que possibilitem esse compartilhamento de forma eficiente e eficaz, contribuindo para o sucesso da gestão do conhecimento (Lartey et al., 2021; Fischer & Döring, 2021; Dey & Mukhopadhyay, 2018; Castaneda et al., 2016; Amayah, 2013).

O segundo cluster obtido é formado por documentos que focam não apenas no processo de compartilhamento do conhecimento, mas em como os recursos humanos podem afetar o processo de GC no serviço público.

Ncoyini e Cilliers (2020), Dewan e Mutula (2016), e Pandey e Dutta (2015) conduzem pesquisas qualitativas em busca de retratar como a GC influencia os colaboradores e vice-versa. Diferente das descobertas de Amayah (2013), os resultados destes autores apresentam a necessidade de incentivos e recompensas, de reciprocidade e de confiança relacional para executar com sucesso o compartilhamento do conhecimento e, por consequência, as instituições possam colher os bons frutos da implementação da GC em seus processos internos e na oferta de serviços para os cidadãos.

Torre et al. (2020), Ganapathy et al. (2019), Pee e Kanhanhalli (2016), e Bakar et al. (2015), usando abordagens quantitativas em diferentes tipos de órgãos públicos pelo mundo, validaram que o capital humano tem influência positiva direta sobre o desempenho da GC. Essa influência pode ser efetuada pelos saberes de cada indivíduo, pela capacidade de relacionamento, pelo compartilhamento ou pela criação de novos conhecimentos, a partir do acúmulo de conhecimento de todos os colaboradores.

O terceiro cluster é formado por artigos cujas pesquisas evidenciam os aspectos positivos da implementação da GC para o colaborador de órgãos públicos. Baporikar (2017), em um duplo estudo de caso envolvendo duas grandes empresas públicas indianas do setor de energia, evidencia que a implementação da GC acarretou melhores oportunidades de progressão de carreira, acesso ao desenvolvimento pessoal e melhor qualidade de vida para os trabalhadores.

Razzaq et al. (2019) usam a modelagem de equações estruturais para validar a influência positiva da GC sobre o comprometimento e o desempenho dos colaboradores em órgãos públicos do Paquistão; enquanto Bertolino et al. (2020) evidenciam as benesses da GC por fomentar a aprendizagem entre os colaboradores, cuja consequência é o aumento das habilidades destes.

Tsirikas e Katsaros (2014) apresentam, em todo o corpus, o único artigo com abordagem quantitativa longitudinal, mensurando a satisfação do trabalhador e a produtividade antes e depois da implementação da GC em um banco público grego. Os autores conseguem validar que a GC tem um impacto positivo sobre a satisfação e a produtividade dos trabalhadores. Esta descoberta também foi registrada nas pesquisas de Lartey et al. (2021) e Torre et al. (2020), reforçando a validade da proposta dos autores gregos.

Salleh (2013), Ncoyini e Cilliers (2020), e Dewan e Mutula (2016) apontam, em seus estudos, a importância do comprometimento e apoio do alto escalão ou lideranças para que o processo de GC tenha resultados positivos e satisfatórios. Assim, o quarto cluster encontrado possui foco exclusivo no papel da liderança na GC em órgãos públicos.

Tripathi et al. (2021) colocando as lideranças como peças chaves para que os funcionários adotem e internalizem as práticas de GC, ou seja, a liderança exerce um papel de empoderamento psicológico. Almudallal et al. (2015) realizam o caminho inverso e validam em seus estudos a hipótese de que a GC tem efeito positivo sobre o desempenho do processo decisório do alto escalão, a partir de um estudo quantitativo na Palestina, usando equações estruturadas.

O penúltimo cluster identificado aborda a influência da cultura organizacional sobre a relação das pessoas com a GC dentro de organizações públicas. Aplicando quatro dos seis pilares do modelo de Hofstede (distância em relação ao poder / coletivo x individual / masculino x feminino / aversão à incerteza) para identificar o perfil cultural presente em instituições públicas do Gana e como este perfil afeta o compartilhamento do conhecimento, Boateng e Agyemang (2015) conseguiram identificar que a cultura presente no poder público daquele país afeta a percepção dos colaboradores sobre a participação no processo decisório, poder e status, delegação de responsabilidades, respeito e justiça, e o compartilhamento do conhecimento.

De Angelis (2016) também utiliza os padrões de perfil cultural determinado por Hofstede (1980) para validar se a cultura local tem influência sobre a GC. Para isso, o autor conduziu uma pesquisa com funcionários públicos da Alemanha e do Brasil e, ao final, conseguiu validar tal influência. Desta forma, coincide com as descobertas de Boateng e Agyemang (2015).

No último documento deste cluster, De Angelis (2013) conduz novamente uma pesquisa com dados oriundos do funcionalismo público brasileiro e alemão, também valida a hipótese de que a cultura organizacional contribui diretamente para a implementação satisfatória da GC. Importante frisar que Ncoyini e Cilliers (2020), Ganapathy et al. (2019), e Pandey e Dutta (2015) também conseguiram ratificar esta hipótese em seus estudos, acrescentando a influência da estrutura organizacional para o sucesso da GC em diversas instituições públicas do mundo.

Por fim, o último cluster é formado apenas por um documento, Taskin e Bunnan (2015), em que os autores registram aspectos negativos da GC para os recursos humanos no poder público. Indo na contramão, o estudo de caso mostra as consequências da implementação da GC em uma instituição pública belga. Entre os efeitos negativos retratados no documento, consigna-se: a deterioração das relações sociais dentro da instituição, a crença na piora das condições de trabalho devido à padronização, a ameaça à riqueza do trabalho dos funcionários, a perda de autonomia e de iniciativa do compartilhamento. O texto descreve diversas estratégias que os funcionários utilizaram como resistência ao processo, tais como boicote, não compartilhamento de conhecimento e falsificação.

4.2 Limitações do corpus analisado

Entre as principais limitações no corpus final desta revisão integrativa, destaca-se inicialmente a ausência de abordagens de pesquisa mais diversas. Observa-se a predominância de pesquisas quantitativas que buscam validar, em ambientes públicos, a relação de variáveis que já foram estabelecidas na literatura de GC em instituições privadas. Quando as abordagens são qualitativas, verifica-se o

uso de estudos de caso focados em apenas uma instituição, impedindo maiores generalizações dos achados. Desta forma, impedem uma maior generalização e maiores aprofundamentos científicos.

Ainda referente às escolhas metodológicas, encontraram-se poucos estudos longitudinais que possibilitem a criação de evidências mais abrangentes e cientificamente comprovadas das consequências da implementação da GC. Nota-se aqui também o número de respostas nos estudos apresentados, cuja taxa de retorno de respondentes não ultrapassa o 30 % em alguns casos, levantando possíveis questionamentos sobre a generalização de determinados resultados.

Por fim, o corpus final apresentou limitações referentes ao país de origem dos artigos, o que impede generalizações. Não se registraram artigos oriundos da Oceania, América Central e México, Leste Europeu, Extremo Oriente e Ásia Central, impedindo traçar uma generalização sobre como a GC afeta os recursos humanos em organizações públicas ao redor do globo.

4.3 Principais contribuições identificadas e oportunidades de pesquisas futuras

A importância do compartilhamento de conhecimento como processo chave para o sucesso da GC, independentemente do tipo de organização, foi trazido à tona nos artigos do primeiro cluster. Os demais clusters identificados contribuiram para validar essa percepção e demonstrar a importância da infraestrutura, do apoio da alta gestão, e da estrutura e cultura organizacionais para o sucesso da GC em órgãos públicos.

Faz-se necessário ampliar esses estudos a fim de entender como todas essas variáveis se relacionam na prática e desenvolver pesquisas quanti e quali que comprovem as relações positivas e negativas existentes, com maior número de dados e informações possíveis.

Importante destacar os trabalhos de Amayah (2013) e Salleh (2013) que vão de encontro ao estabelecido na literatura clássica e demonstraram a não relação positiva entre benefícios pessoais, reciprocidade, confiança e oportunidades de aprendizagem com o compartilhamento do conhecimento. É necessário ampliar o número de estudos ou identificar novos artigos sobre o tema, a fim de compreender os casos em que essa relação positiva não existe e quais fatores sociais influenciam nesta correlação.

Outro ponto chave é a relação positiva da implementação da GC em órgãos públicos e o aumento nos níveis de desempenho, produtividade e satisfação. Identificou-se a necessidade de ampliar o número de pesquisas ou identificar mais artigos que demonstrem este impacto positivo e evidenciem em quais situações essa relação não acontece.

Por outro lado, também é importante identificar artigos, na linha de Taskin e Bunnan (2015), que evidenciem as consequências negativas da implementação da GC na esfera pública, para aumentar o espectro de entendimento sobre as percepções e perspectivas das pessoas envolvidas nestes cenários.

Importante notar o uso do modelo de Hofstede (De Angelis, 2016; Boateng & Agyemang, 2015) para determinar o perfil cultural de instituições, regiões e países. Assim, compreende-se como a cultura local influencia diretamente nos processos de gestão do conhecimento. Aqui encontramos um gap a ser explorado, que permitiria identificar quais são as adaptações necessárias nos frameworks de implementação de GC, no poder público, frente ao contexto cultural de um determinado território.

5. CONCLUSÃO

O presente artigo alcançou o objetivo proposto, por meio de uma revisão integrativa, que demonstrou como a literatura acadêmica investiga a relação entre o capital humano, a GC e o setor público.

A pesquisa conseguiu identificar as seis abordagens mais atuais no que tange à análise da relação supracitada, entendendo o compartilhamento de conhecimento, identificando como os recursos humanos podem facilitar a GC, analisando os impactos das lideranças e da cultura organizacional e compreendendo os efeitos positivos ou negativos da GC sobre o capital humano.

O estudo também conseguiu determinar a tendência crescente de correlacionar a GC no setor público com a inovação e a criatividade, na tentativa de maximizar os efeitos da GC sobre a competitividade, eficiência e eficácia.

Também foi possível ratificar, diante dos resultados, a importância das pessoas para a GC em setores públicos, sendo este um fator chave para implementações de sucesso. Agir sobre a cultura organizacional e incentivar de forma assertiva o compartilhamento de conhecimento mostraram ser componentes fundamentais segundo os autores presentes no corpus de pesquisa.

A investigação também apresentou os impactos da GC sobre os recursos humanos em instituições públicas. Entre eles aparecem uma maior satisfação com o trabalho, um aumento dos índices de produtividade e aprendizagem, bem como maiores probabilidades de progressão de carreira. Por outro lado, foi evidenciada a perda de autonomia e o recrudescimento das relações de trabalho o que reforça a necessidade de pesquisas futuras, preferencialmente longitudinais que consigam evidenciar o impacto da GC em funcionários públicos, a fim de (in)validar os achados do corpus final analisado neste artigo.

Salienta-se novamente a necessidade de pesquisas futuras sobre o impacto da cultura local na adoção da GC, bem como investigações que tentem reproduzir os resultados desta pesquisa, porém com amostras maiores e com maior abrangência geográfica. Desta forma, a análise mostrou um vasto campo de pesquisa a ser explorado sobre as relações entre a GC, o setor público e o contexto organizacional das pessoas envolvidas nesse processo.

Agradecimentos

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

Os autores agradecem a CAPES por apoiar financeiramente este trabalho através da bolsa CAPES/PROEX.

REFERÊNCIAS

- Alavi, M., & Leidner, D. E. (2001). Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS quarterly*, 25(1), 107-136.
- Almudallal, A. W., Bakri, N., Muktar, S. N., & El-Farra, M. M. (2016). Implementing knowledge management in the palestinian public sector institutions: empirical study on the Presidency of the Palestinian Government. *International Review of Management and Marketing*, 6(4), 101-107.
- Amayah, A. T. (2013). Determinants of knowledge sharing in a public sector organization. *Journal of knowledge management*, 17(3), 454-471.
- Araújo, C. (2006). Bibliometria: evolução histórica e questões atuais. *Em questão*, 12(1), 11-3.
- Bakar, A., Rahim, N. Z. A., & Ibrahim, R. (2015, August). *Measurement Model of Relationship Between Knowledge Management Enablers and Knowledge Management Performance in Public Sector* [Conferencia]. International Conference on Knowledge Management in Organizations, Springer, Cham.
- Baporikar, N. (2017). Knowledge Management for Excellence in Indian Public Sector. *International Journal of Social Ecology and Sustainable Development*, 8(1), 49-65.
- Bell DeTienne, K., Dyer, G., Hoopes, C., & Harris, S. (2004). Toward a model of effective knowledge management and directions for future research: Culture, leadership, and CKOs. *Journal of leadership & organizational studies*, 10(4), 26-43.

- Bertolino, A., De Angelis, G., Polini, A., Re, B., & Sergiacomi, A. (2020). Process-oriented knowledge management and learning in public administrations. *Electronic Government, an International Journal*, 16(4), 441-465.
- Boateng, H., & Agyemang, F. G. (2015). The role of culture in knowledge sharing in a public-sector organization in Ghana: Revisiting Hofstede's model. *International Journal of Public Administration*, 38(7), 486-495.
- Botelho, L. L. R., Cunha, C. C. A., & Macedo, M. (2011). O método da revisão integrativa nos estudos organizacionais. *Gestão e Sociedade*, 5(11), 121-136.
- Castaneda, D. I., Rios, M. F., & Duran, W. F. (2016). Determinants of knowledge-sharing intention and knowledge-sharing behavior in a public organization. *Knowledge Management & E-Learning: An International Journal*, 8(2), 372-386.
- Castaneda, D.I., Pardo, C., & Toulson, P. (2015). Validação de instrumento de compartilhamento de conhecimento espanhol. *Revista Eletrônica de Gestão do Conhecimento*, 13(1), 3-12.
- Castaneda, D. I., & Toulson, P. (2013). O valor da mensuração de recursos humanos no capital intelectual e no compartilhamento de conhecimento. *Electronic Journal of Knowledge Management*, 11(3), 226-234.
- Cooper, P. (2017). Data, information, knowledge and wisdom. *Anaesthesia & Intensive Care Medicine*, 18(1), p. 55-56.
- Creswell, J. W. (2010). *Projeto de pesquisa: Métodos qualitativo, quantitativo e misto* (3.ª ed.). Artmed.
- Cummings, J. L., & Teng, B. S. (2003). Transferência de conhecimento RandD: Os principais fatores que afetam o sucesso da transferência de conhecimento. *Journal of Engineering and Technology Management*, 20(1), 39-68.
- Dalkir, K. (2013). *Knowledge management in theory and practice*. Routledge.
- Da Silva, J. C., Charrua-Santos, F. B., Franco, M., Santos, K. F., Garcez, A. P. B., Moreira, J. A., & da Silva, R. M. (2021, November). The Needs of the future society 5.0: A challenge. Em *Proceedings of ICERI 2021 Conference* (Vol. 8, p. 9th).
- De Angelis, C. T. (2016). The impact of national culture and knowledge management on governmental intelligence. *Journal of Modelling in Management*, 11(1), 240-268.
- De Normalisation, C. E., & Normung, E. K. F. (2004, March). European Guide to good Practice in Knowledge Management-Part 5: KM Terminology. Em *CEN Workshop Agreement CWA* (pp. 14924-1).
- Dey, T., & Mukhopadhyay, S. (2018). Linkage between contextual factors, knowledge-sharing mediums, and behaviour: Moderating effect of knowledge-sharing intentions. *Knowledge and Process Management*, 25(1), 31-40.

- Dewah, P., & Mutula, S. M. (2016). Knowledge retention strategies in public sector organizations: Current status in sub-Saharan Africa. *Information Development*, 32(3), 362-376.
- Fischer, C., & Döring, M. (2021). Thank you for sharing! How knowledge sharing and information availability affect public employees' job satisfaction. *International Journal of Public Sector Management*, 35(1), 76-93.
- Ganapathy, S., Mansor, Z., & Ahmad, K. (2019). Investigating Factors Affecting Knowledge Management Practices in Public Sectors. *International Journal of Advanced Computer Science and Applications*, 10(11), 205-212.
- Gruzman, C., & Siqueira, V. H. F. (2007). O papel educacional do Museu de Ciências: desafios e transformações conceituais. *Revista Electrónica de Enseñanza de las Ciencias*, 6(2), 402-423.
- Hofstede, G. (1980). Culture's consequences. Beverly Hills: Sage Publications. *Revista de Gestão Estratégica*, 21, 473-49.
- Koskien, K. U. (2013). Business organizations' knowledge-production processes: an autopoietic approach. *International Journal of Organizational Analysis*, 21(2), 137-153.
- Lartey, P. Y., Kong, Y., Afriyie, S. O., Santosh, R. J., & Bah, F. B. M. (2021). Knowledge management issues in India: A public sector perspective. *International Journal of Public Administration*, 44(3), 215-230.
- Martinsons, M. G., Davison, R. M., & Huang, Q. (2017). Strategic knowledge management failures in small professional service firms in China. *International Journal of Information Management*, 37(4), 327-338.
- Massaro, M., Dumay, J. & Garlatti, A. (2015). Gestão do conhecimento do setor público: uma revisão estruturada da literatura. *Journal of Knowledge Management*, 19(3), 530-558.
- May, B. L., & Fombad, M. (2019). Knowledge sharing at the overstrand municipality in Western Cape Hermanus, South Africa. *Libri*, 69(4), 325-336.
- Moreira, F. K., Freire, P. S., & Souza, J. A. (2022). *Governança Multinível Pública; um novo modelo para a Governança Multinível Pública Brasileira*. Editora CRV.
- Muqadas, F., Rehman, M., & Aslam, U. (2017). Exploring the challenges, trends and issues for knowledge sharing: A study on employees in public sector universities. *VINE Journal of Information and Knowledge Management Systems*, 47(1), 2-15.
- Nair, P., & Prakash, K. (2009). Knowledge Management: facilitator's guide. Em Asian Productivity Organization (Ed.), *Knowledge Management: facilitator's guide*. Tokyo: APO, (pp. 105-209).

- Ncoyini, S. S., & Cilliers, L. (2020). Factors that influence knowledge management systems to improve knowledge transfer in local government: A case study of Buffalo City Metropolitan Municipality, Eastern Cape, South Africa. *SA Journal of Human Resource Management*, 18(1), 1-11.
- Neves, J. A. B. (2018). *Modelo de equações estruturais: uma introdução aplicada*. Disponível em <http://repositorio.enap.gov.br/handle/1/3334>.
- Neves, M. L. C., Rados, G. J. V., & Fialho, F. A. P. (2018). Pessoas, processos e tecnologia na gestão do conhecimento: uma revisão da literatura. *Ciências da Administração*, 20(51), 153-167.
- Nonaka, I., & Takeuchi, H. (1997). *Criação de Conhecimento na Empresa*. Campus.
- Organização para a Cooperação e Desenvolvimento Econômico. (2007). *Orçamento de desempenho nos países da OCDE*. OCDE.
- Polanyi, M. (1967). *The Tacit Dimension*. Roudtledge & Kegan Paul.
- Pandey, S. C., & Dutta, A. (2015). Knowledge infrastructure capabilities and knowledge management: case of an Indian public sector undertaking. *International Journal of Knowledge-Based Development*, 6(1), 50-64.
- Pee, L. G., & Kankanhalli, A. (2016). Interactions among factors influencing knowledge management in public-sector organizations: A resource-based view. *Government Information Quarterly*, 33(1), 188-199.
- Probst, G.; Raub, S. & Romhardt, K. (2002). *Gestão do conhecimento: os elementos construtivos do sucesso*. Bookamn.
- Razzaq, S., Shujahat, M., Hussain, S., Nawaz, F., Wang, M., Ali, M., & Tehseen, S. (2018). Knowledge management, organizational commitment and knowledge-worker performance: The neglected role of knowledge management in the public sector. *Business process management journal*, 25(5), 923-947.
- Salleh, K. (2013). Knowledge sharing and knowledge management modelling in public sector accounting organisation: case evidence from Malaysia. *International Journal of Knowledge-Based Development*, 4(4), 363-381.
- Santos, R. N. M. D. & Kobashi, N. Y. (2009). Bibliometria, cientometria, infometria: conceitos e aplicações. *Tendências da Pesquisa Brasileira em Ciência da Informação*, 2(1), 155- 172.
- Santos, N., & Varvakis, G. (2021). *O que é Gestão do Conhecimento?* [Material de apoio de aula online]. Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento, Universidade Federal de Santa Catarina, Florianópolis, Santa Catarina.

- Serenko, A., & Bontis, N. (2013). Global ranking of knowledge management and intellectual capital academic journals: 2013 update. *Journal of Knowledge Management*, 13(2), 4-15.
- Storey, J., & Barnett, E. (2002). Iniciativas de gestão do conhecimento: aprendendo com falha. *Journal of Knowledge Management*, 4(2), 145-156.
- Stewart, T. A. (1998). *Capital Intelectual. A nova vantagem competitiva das empresas* (10.ª ed.). Campus.
- Taskin, L., & Van Bunnan, G. (2015). Knowledge management through the development of knowledge repositories: towards work degradation. *New Technology, Work and Employment*, 30(2), 158-172.
- Teng, BS (2006). As chaves para o compartilhamento de conhecimento bem-sucedido. *Journal of General Management*, 31(4), 1.
- Torraco, R. J. (2005). Writing integrative literature reviews: Guidelines and examples. *Human resource development review*, 4(3), 356-367.
- Torre, C., Tommasetti, A., & Maione, G. (2020). Technology usage, intellectual capital, firm performance and employee satisfaction: the accountants' idea. *The TQM Journal*, 33(3), 545-567.
- Tripathi, D., Priyadarshi, P., Kumar, P., & Kumar, S. (2020). Does servant leadership affect work role performance via knowledge sharing and psychological empowerment? *VINE Journal of Information and Knowledge Management Systems*, 51(5), 792-812.
- Tsirikas, A. N., & Katsaros, K. K. (2014). Linking knowledge management, job satisfaction and productivity in the Greek public sector. *International Journal of Knowledge Management Studies*, 5(3-4), 244-264.
- Wang, P., Zhu, F. W., Song, H. Y., Hou, J. H., & Zhang, J. L. (2018). Visualizing the academic discipline of knowledge management. *Sustainability*, 10(3), 682.
- Whittemore, R. & Knafl, K. (2005). The integrative review: updated methodology. *Journal of advanced nursing*, 52(5), 546-553.
- Willem, A., & Buelens, M. (2006), Compartilhamento de Conhecimento em Organizações do Setor Público. O efeito das características organizacionais no compartilhamento de conhecimento interdepartamental. *Journal of Public Administration Research and Theory*, 17(4), 581-606.
- Wiig, K. M. (1997). Integrating intellectual capital and knowledge management. *Long range planning*, 30(3), 399-405.

ARTÍCULOS DE INVESTIGACIÓN

REFORZAMIENTO DE HABILIDADES ARITMÉTICAS DE FRACCIONES Y DIVISIONES EN NIÑOS DE PRIMARIA MEDIANTE UN JUEGO SERIO EDUCATIVO*

RENATO CORONADO ÁLVAREZ
renatocoronado99@gmail.com
<https://orcid.org/0000-0002-0309-3690>
Universidad de Lima, Perú

RESUMEN. Si bien la matemática resulta fácil para algunos estudiantes, para otros puede ser una verdadera tortura. Sin embargo, con ella uno puede desde organizarse mejor hasta tener mucho más lógica y razonamiento, además de ser el conocimiento base para muchas profesiones. No obstante, a muchos se les complica y no se sienten motivados a aprenderlas o reforzarlas, principalmente, debido a la metodología tradicional, a la dificultad misma de la matemática o la calidad educativa. Por otro lado, la tecnología y los juegos pueden fomentar el aprendizaje, volviéndose el mejor aliado para potenciar la educación. Más aún, para involucrar a los estudiantes en el pensamiento de orden superior, es importante diseñar contextos de juego y misiones que permitan lidiar con problemas prácticos. Por esta razón, el presente artículo pretende apoyar el proceso de reforzamiento del aprendizaje de fracciones y divisiones a través de un juego serio educativo disponible para plataformas web y móvil, que incorpora una mecánica continua y ejercicios de refuerzo de fracciones y divisiones. Para el diseño del juego e interfaces, se aplicaron las heurísticas propuestas por Nielsen (1994), y para la mecánica, el *framework* MDA, el cual influyó en el diseño de los niveles, junto con las recomendaciones de un *focus group*. Posteriormente, se validó si hubo cambios en el desempeño de los estudiantes mediante comparación de grupos, así como la usabilidad, con la finalidad de verificar que las heurísticas estén correctamente aplicadas; finalmente, se midió la experiencia del juego en un grupo reducido de escolares a partir de tercero de primaria.

PALABRAS CLAVE: aprendizaje matemático, educación matemática, interacción humano-computadora, juegos serios, aprendizaje basado en juegos, cuestionario de experiencia de juego

* Me gustaría expresar mi más sincero y cordial agradecimiento a los profesores Óscar Ramos y Hernán Alejandro Quintana Cruz, quienes aportaron información valiosa para el desarrollo de la presente investigación y también me acompañaron a lo largo de todas sus fases.

REINFORCEMENT OF IN FRACTIONS AND DIVISION ARITHMETIC SKILLS IN PRIMARY SCHOOL CHILDREN THROUGH A SERIOUS EDUCATIONAL GAME

ABSTRACT. While mathematics may be easy for some students, for others it can be a real torture. Through mathematics we can organize ourselves better and have much more logic and reasoning, as it is also the base knowledge for many professions. However, many people find it difficult and do not feel motivated to learn or reinforce it, mainly due to the traditional methodology, the difficulty of mathematics or the quality of education. On the other hand, technology and games can promote learning, becoming the best ally to enhance education. Moreover, in order to engage students in higher-order thinking, it is important to design game contexts and missions that allow them to deal with practical problems. For this reason, this article aims to support the process of reinforcement of learning fractions and divisions through an educational serious game available for web and mobile platforms, which incorporates continuous mechanics and the inclusion of reinforcement exercises of fractions and divisions. For the design of the game and interfaces, the heuristics proposed by Nielsen (1994), were applied and for the mechanics, the MDA framework was applied, which influenced the design of the levels together with the recommendations of a focus-group. Subsequently, we will proceed to validate whether there were changes in the students' performance by comparing groups, validate the usability in order to verify that the heuristics are correctly applied and, finally, measure the experience of the game in a small group of schoolchildren from 3rd grade onwards.

KEYWORDS: mathematical learning, mathematics education, human-computer interaction, serious games, game-based learning, Game Experience Questionnaire

1. INTRODUCCIÓN

La utilidad de la aritmética en la vida diaria es innegable, así como es fundamental en otras disciplinas, pues constituye un conocimiento básico en muchas profesiones, además de ser importante en el desarrollo del razonamiento crítico. Sin embargo, como afirma Mishra (2020), el temor excesivo a la matemática impide que muchos estudiantes tengan éxito en los cursos básicos de esta materia. A menudo, memorizan el concepto, por lo que existe una mayor probabilidad de que cometan errores debido a un conocimiento inadecuado. De esta manera, aprender matemática, ya sea en el nivel escolar o superior, tiene cierta dificultad para muchas personas. La matemática es una de las disciplinas en las que el aprendizaje de conceptos tiene un lugar esencial; sin embargo, se puede observar que muchos estudiantes poseen conceptos erróneos en esta materia. De hecho, es casi imposible en matemática definir cualquier concepto sin utilizar muchos otros, ya que los planes de estudio de esta disciplina tienen un atributo de espiral. Esos errores y conceptos erróneos son parte del proceso de aprendizaje y se pueden corregir con una explicación y práctica regular (Mishra, 2020; Ninaus et al., 2017). En particular, de acuerdo con Benbow & Faulkner (2008), el tema de fracciones suele generar cierta complejidad; es más, el procesamiento y aprendizaje de fracciones es uno de los problemas más desafiantes en la educación matemática, dado que una parte crucial de su comprensión y procesamiento es la representación exitosa de la magnitud de la fracción (que refleja la relación entre denominador y numerador) para realizar operaciones aritméticas con ellas.

Por otro lado, en Latinoamérica, el Perú ocupa los últimos lugares en evaluación en matemática y comunicación, como lo muestran las pruebas internacionales realizadas por el Programa para la Evaluación Internacional de Alumnos de la OCDE (PISA, por sus siglas en inglés) y el Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (LLECE). Además, el Ministerio de Educación del Perú, mediante la Unidad de Medición de la Calidad Educativa, ha llevado a cabo la Evaluación Censal de Estudiantes (ECE), cuyos resultados muestran diferencias muy marcadas entre los logros alcanzados por las instituciones educativas urbanas y rurales. En esta evaluación, se concluyó que el 46,4 % de los alumnos obtuvieron niveles satisfactorios en comprensión lectora y el 34,1 %, en matemática. Estas cifras han mejorado, pero no alcanzan más del 50 %; la mejoría puede deberse a la educación inicial, a la contratación de docentes en forma oportuna y a los programas de acompañamiento (Ibarra et al., 2016; Ministerio de Educación del Perú, 2018).

La acelerada evolución tecnológica que se vive actualmente también ha alcanzado a la tecnología en la educación —sobre todo con la coyuntura actual, en la que se habla de una educación virtual—, lo que permite un aprendizaje más didáctico, ofreciendo la

posibilidad de integrar la tecnología, los juegos y la educación en un solo sistema. Esta situación ha llevado a los ministerios y escuelas a probar el uso de aplicaciones móviles para lograr objetivos de aprendizaje y enseñanza. Además de las computadoras y páginas web, los *smartphones* y las *tablets* son las nuevas tecnologías emergentes que pueden tener un gran efecto en este ámbito, pues se considera que están extendidos por todo el mundo (casi todo el mundo cuenta con uno) y que están siempre con los usuarios (se lo puede usar en cualquier lugar); así, por ejemplo, jugar videojuegos en dispositivos móviles se ha convertido en una tendencia porque se puede hacer en cualquier lugar y en cualquier momento. Este tipo de tecnología aporta portabilidad, conectividad e interactividad social, características que la convierten en una plataforma preferida para el aprendizaje. Como resultado, el uso de dispositivos móviles puede extenderse al aprendizaje móvil, donde el entorno de aprendizaje no está limitado al aula. Al combinar el factor lúdico de los videojuegos, la flexibilidad del aprendizaje móvil y los beneficios de los juegos, se han desarrollado los juegos móviles serios (Alkhateeb, 2019; El Azizi, 2019).

Muchos estudios han revelado el potencial de los juegos digitales para mejorar los resultados del aprendizaje de los estudiantes en los cursos de matemática (Schenke et al., 2014). Por ejemplo, Kebritchi et al. (2010) encontraron que los estudiantes que aprenden con juegos de matemática mostraron mejores logros de aprendizaje que aquellos que aprendieron con instrucción tradicional. También indican la significancia del uso de los juegos educativos móviles entre los alumnos, porque integran aprendizaje con placer y suspenso. De este modo, se puede aumentar sus logros académicos y desarrollar su pensamiento creativo a través de un método innovador. Asimismo, contribuyen a la individualización de la enseñanza, porque permiten que el alumno progrese en su aprendizaje para que se adapte a sus habilidades y lo acelere, sin timidez ni miedo (Bicen & Kocakoyun, 2018). Finalmente, Robertson y Howells (2008) demuestran que los juegos educativos sirven para ejercitar conocimientos y fomentar la creatividad en la creación de un artefacto cultural complejo e interactivo que puede ser usado y disfrutado luego por otros usuarios.

En vista de las situaciones descritas, el presente estudio de investigación propone el desarrollo de un juego serio educativo para el refuerzo del área de aritmética, específicamente del tema de fracciones y divisiones en niños de primaria, que sea capaz de hacer frente a las limitaciones previamente mencionadas. Para validar la efectividad del sistema, se propuso un método de evaluación de desempeño que fue aplicado a participantes de dos grupos: experimental y de control.

Para comenzar, se explicarán los antecedentes más relevantes para esta investigación. En la metodología, se estipulan los pasos para el desarrollo del juego y la experimentación correspondiente. Luego, se presentarán los resultados

obtenidos y su discusión, donde se hace una interpretación de estos. Finalmente, se ofrecen tanto las conclusiones como las recomendaciones para trabajos futuros sobre la base de los resultados.

2. ESTADO DEL ARTE

En esta sección se hace una revisión de investigaciones previas que han servido como aporte para el desarrollo del sistema y que, a su vez, dan los lineamientos y bases de viabilidad al presente estudio. La propuesta de todos los estudios referenciados en esta sección son juegos para el apoyo del aprendizaje.

2.1 Juegos serios para el aprendizaje de matemática en niños

Existen varios trabajos de la literatura que abordan el problema de la enseñanza de matemática a los niños. Wang et al. (2017), por ejemplo, presentan *Speedy World*, un juego de computadora para mejorar la motivación y el rendimiento de los estudiantes de primaria en el aprendizaje del concepto de velocidad, además de abordar el problema de enseñar matemática con el enfoque tradicional centrado en el maestro. El estudio adoptó un diseño cuasiexperimental, en el que 107 alumnos de sexto grado de primaria se dividieron en dos grupos (experimental y de control) y sus resultados fueron significativos: los estudiantes que aprendieron con el juego superaron a los que aprendieron con el enfoque tradicional. Finalmente, se concluye que el rendimiento de aprendizaje y la participación de los estudiantes mejoraron al adoptar el enfoque de juego (Wang et al., 2017).

Para las sumas y restas, que son el paso previo al aprendizaje de divisiones y fracciones, investigadores como Aljojo (2018) buscan mejorar la capacidad matemática para estas operaciones de los alumnos de primaria de 6 a 8 años mediante un juego con tres niveles de dificultad. En el juego se incorporaron técnicas educativas para producir un juego educativo tan divertido como emocionante, y se introdujo un nuevo concepto de aprendizaje mediante la combinación de educación y entretenimiento a través de un laberinto y preguntas según el nivel (ordenar números, sumar y restar; y sumar y restar con imágenes). El juego fue probado con 50 estudiantes y los resultados rechazan la hipótesis nula, por lo que se concluye que existe una fuerte evidencia de que la intervención docente mejora los puntajes a través del uso de una aplicación de aprendizaje basada en juegos (Aljojo, 2018).

Los resultados concuerdan con los de Muñoz Sanabria y Vargas Ordoñez (2019), quienes a través de *EDUMAT*, un juego en línea, buscan mejorar el desempeño de los estudiantes de primaria enfocándose en las divisiones con restas sucesivas (elegidas porque incluyen las cuatro operaciones básicas). En el juego, se hace

uso de características como la dinámica, la mecánica y los componentes del juego (logros, trofeos, insignias, avatares, etcétera); también incluye la visualización del *ranking* de los mejores jugadores y la opción de iniciar la partida; al final, se muestra un resumen de esta: el puntaje y los intentos fallidos (Muñoz Sanabria & Vargas Ordoñez, 2019).

2.2 Juegos serios para el aprendizaje de fracciones

Kickmeier-Rust et al. (2014) plantean un juego para aprender y practicar divisiones llamado *Sonic Divider*, que incluye funciones de evaluación formativa y retroalimentación (muestra una cara sonriendo y un texto motivacional si la tarea se completa correctamente). *Sonic Divider* incluye técnicas como *goal-setting*, *levelling* y *scoring*, entre otras, y está enfocado en niños de 6 a 8 años, diseñado específicamente para practicar fracciones exactas. También existe la posibilidad de que el profesor configure la aplicación para su salón de clase, y en este caso, ofrece una función que permite evaluar e interpretar las habilidades en un sentido formativo. Por último, esta aplicación implementa una herramienta de análisis del salón de clases para el maestro, la cual le muestra los resultados generales, el tiempo total y el tiempo promedio por tarea, de manera individual o colectiva. Los resultados de la investigación fueron alentadores, pues, conforme se practica más, disminuyen los errores (Kickmeier-Rust et al., 2014).

Desde la posición de Alkhateeb (2019), el bajo desempeño en matemática es un gran problema, por lo que plantea un estudio cuasiexperimental en el que diseñó y produjo un juego educativo para ser utilizado en la enseñanza. Realizó la experimentación en una muestra formada por 66 alumnos de cuarto grado de primaria de la ciudad de Zarqa (Jordania), a los que dividió en dos grupos: experimental y de control (el primero aprenderá mediante el juego y el segundo, por el método convencional). Los resultados demuestran que hubo diferencias en el rendimiento entre ambos grupos. El estudio encontró que el uso de los juegos móviles es efectivo para ayudar a brindar soporte a los estudiantes en un salón de clase, y que el juego propuesto, en particular, fue efectivo tanto por su diseño como por su contenido (Alkhateeb, 2019). Los resultados del estudio son similares a los de Vandercruysse et al. (2015), quienes describen *Monkey Tales: The Museum of Anything*, un juego educativo para computadora utilizado para practicar matemática. Este juego pertenece a la categoría de juegos de aventuras, pero también contiene algunos elementos de los juegos de acción (por ejemplo, disparar burbujas con fracción que representan la respuesta correcta). Los jugadores tienen que derrotar al enorme dinosaurio que se ha apoderado del museo resolviendo rápidamente acertijos matemáticos, como comparaciones u operaciones con fracciones (Vandercruysse et al., 2015).

Con una problemática similar, Zhang et al. (2020) señalan que la dificultad de los niños con las fracciones está asociada a su conocimiento de números enteros y la teoría del sesgo de estos números; es un desafío porque los niños no familiarizados con fracciones pueden ver $2/5$ como dos números, y no como una parte y un todo. Por ello, a través del estudio cuasiexperimental, se combina el aprendizaje de fracciones basado en juegos con lecciones escolares (grupo experimental), y se hace una comparación con la instrucción tradicional (grupo de control). En total, se tenían 37 estudiantes en el grupo de control y 39 en el grupo experimental, cuya edad variaba entre 8 y 9 años. Sus resultados fueron positivos, dado que se corroboró el potencial de los juegos serios para apoyar el aprendizaje de conceptos de fracciones, y se concluye que los niños tienden a concentrarse más en el aprendizaje basado en juegos: normalmente solo se puede mantener su atención durante 40 minutos; sin embargo, jugar en clase puede aliviar un poco la fatiga del aprendizaje mientras se disfruta de una experiencia de aprendizaje agradable (Zhang et al., 2020).

Algunos investigadores como Shang et al. (2018) consideran que el aprendizaje de fracciones sigue siendo un desafío, y que es necesario entenderlo en relación con el concepto. Por ello, sobre la base de algunas teorías relacionadas, los autores diseñaron el juego *Run Fraction* para ayudar a los niños a aprender el conocimiento conceptual de las fracciones. En este juego, la línea numérica aparece con la forma de una pared, y la vista del avatar está bloqueada, lo que despierta la curiosidad y alienta al jugador a explorar más. Además, cuenta con un sistema de recompensa y clasificación con el objetivo de mantener la motivación (Shang et al., 2018). Igualmente, se tienen juegos que se enfocan en la práctica de las fracciones como *iFractions*, un juego serio gratuito para la enseñanza de fracciones y creado principalmente para niños de 6 a 10 años. El juego tiene características flexibles de configuración para el usuario: la selección del idioma, la imagen y el sonido pueden contextualizarse según el país. Además, cada nivel es de una dificultad progresiva (Ibarra et al., 2019).

Para finalizar, Ninaus et al. (2017) desarrollan un motor de investigación de números racionales basado en juegos: *Semideus*. Los estudiantes de 11 años jugaron un juego de matemática en tabletas usando el control de inclinación para navegar con un avatar a lo largo de una recta numérica durante un total de 30 minutos. Se obtuvo que el desempeño de los estudiantes en la comparación de números enteros fue muy bueno, aunque algunos enfrentaron dificultades en esta tarea (Ninaus et al., 2017). En línea con estos resultados, se encuentra la investigación de Gaggi et al. (2018), quienes hacen uso de una narrativa y representación de las fracciones mediante figuras a través de *Pizza al lancio*, un juego serio para ayudar a los niños a comprender las fracciones, en particular las equivalentes y complementarias. El juego cuenta la historia de un repartidor que tiene hambre y come una porción de

pizza durante la entrega. Como le asusta su severo jefe, que le grita si los clientes no están contentos, cuando llega al destino llama al jugador y le pide que complete la *pizza*. El jugador debe lanzar la casilla correcta, entre las cuatro disponibles, la que contiene la fracción complementaria a la *pizza* entregada para cumplir con el pedido del cliente. Sus resultados fueron alentadores debido a que se obtuvo 88 % de aprobación entre los estudiantes y estos asociaron lo reforzado en el juego de manera que, después de la experimentación, en la prueba posterior, se les pidió representar gráficamente fracciones y los niños utilizaron *pizzas* y barras para ello (Gaggi et al., 2018).

3. ANTECEDENTES

En esta sección, se presentan los fundamentos para el desarrollo de la investigación. Primero, se definen los conceptos que se van a utilizar a lo largo del presente estudio. Luego, se expone una definición y aplicación de algunos investigadores de MDA: un *framework* para diseño de juegos.

3.1 Juego serio

Según Aparício y Silva (2019), con el avance de las nuevas tecnologías, el uso de juegos digitales ha aumentado, principalmente, con fines recreativos. Sin embargo, no todos los juegos tienen solo un propósito de entretenimiento y diversión; algunos cuentan con otros propósitos como enseñar o entrenar a los jugadores para situaciones reales concretas (es decir, juegos serios). En esta categoría se encuentran los juegos para el aprendizaje, que tienen como objetivo principal enseñar algunos contenidos a los jugadores de una forma divertida.

Los juegos serios o educativos están adquiriendo un creciente interés en la investigación como herramientas para ampliar los enfoques tradicionales de instrucción sobre el aprendizaje escolar, especialmente de las matemáticas; es más, tienen el potencial de proporcionar formas nuevas, atractivas e innovadoras de entrenar tanto a niños como a adultos en esta materia (Ninaus et al., 2017). Por lo tanto, los juegos serios sugieren que la combinación de la enseñanza en el aula y los juegos podría ser la mejor pedagogía matemática.

Finalmente, como afirman Ibarra et al. (2016), los juegos serios son videojuegos creados con propósitos bien definidos en el ámbito pedagógico y para desarrollar habilidades o competencias directamente relacionadas con una profesión o carrera (con conexión académica).

3.2 Framework de diseño: mecánica, dinámica y estética

La estructura de un juego (reglas, metas) se puede definir con la ayuda de MDA (del inglés: *mechanics, dynamics, aesthetics*), pues describe la separación de los componentes del juego y proporciona un lente útil para pensar en los diversos elementos que componen el juego.

El *framework* fue desarrollado y enseñado como parte del Taller de Diseño y Ajuste de Juegos en la Conferencia de Desarrolladores de Juegos (San José, 2001-2004). Tiene un enfoque de carácter formal e iterativo para el diseño y afinación de juegos, y permite razonar explícitamente sobre los objetivos de diseño particulares y ver cómo los cambios afectan cada aspecto. Consta de tres capas de abstracción, por lo que se puede comprender el concepto dinámico de los juegos.

Hunicke et al. (2004) mencionan que MDA es un enfoque formal para entender los juegos, uno que intenta cerrar la brecha entre el diseño y el desarrollo de juegos, la crítica de juegos y la investigación técnica de juegos. También señalan que la diferencia entre los juegos y otros productos de entretenimiento reside en que su consumo es relativamente impredecible. La serie de eventos que ocurren durante el juego y su resultado se desconocen en el momento en que finaliza el producto, es por ello que MDA formaliza el consumo dividiéndolos en componentes y estableciendo sus contrapartes.

3.2.1 Capa mecánica

La mecánica se refiere a las reglas y la lógica del juego (por ejemplo, puntos, nivel, desafíos/misiones, bienes virtuales, tablas de clasificación). Esta capa es la más cercana al diseñador e incluye todos los recursos, reglas y materiales multimedia utilizados. En ella están las diversas acciones, comportamientos y mecanismos de control que se ofrecen al jugador dentro de un contexto de juego. Junto con el contenido (niveles, recursos, etcétera), la mecánica respalda la dinámica general del juego. Además, describe los componentes particulares del juego en cuanto a representación de datos y algoritmos. Asumiendo la perspectiva del diseñador, la mecánica da lugar a un comportamiento dinámico del sistema, que a su vez conduce a experiencias estéticas particulares. En cambio, desde la perspectiva del jugador, la estética marca la pauta, que nace de la dinámica observable y, finalmente, de la mecánica operable. Por consiguiente, al trabajar con juegos, es útil tener en cuenta las perspectivas tanto del diseñador como del jugador, puesto que ayuda a observar cómo incluso los pequeños cambios en una capa pueden afectar a otras (Hunicke et al., 2004).

3.2.2 Capa dinámica

Siempre que el juego está “funcionando”, su mecánica responde a la entrada del jugador y otras mecánicas del juego; esta capa intermedia se llama *dinámica*. Este componente describe el comportamiento en tiempo de ejecución de las mecánicas que actúan sobre las entradas de los jugadores y las salidas de los demás a lo largo del tiempo; se refiere también a los resultados de la mecánica durante el juego (recompensas, estado, logros, competencia, retroalimentación visual/sonora). La dinámica trabaja para crear experiencias estéticas y está centrada en identificar las actividades claves dentro del flujo de trabajo del usuario a medida que interactúan con los componentes de *hardware* y *software* (Carrol, 2013; Hunicke et al., 2004).

3.2.3 Capa estética

Finalmente, los jugadores experimentan una respuesta emocional compleja al interactuar con el juego, la cual es denominada por varios autores la *estética del juego* (Mora-Zamora & Brenes-Villalobos, 2019; Spieler & Kemeny, 2020). La estética es “el aspecto y la sensación” del juego, y puede reforzarse con el arte, el sonido y otros elementos aclaratorios (por ejemplo, narrativa, sensación, descubrimiento). Como afirman Hunicke et al. (2004), hablar de juegos y jugar se complica porque el vocabulario que usamos es relativamente limitado. Por ello, al describir la estética de un juego, es necesario alejarse de palabras como *diversión* y *juego* hacia un vocabulario más directo, que, si bien incluye, no se limita a la siguiente taxonomía:

1. Sensación: el juego como placer de los sentidos
2. Fantasía: juegos como fantasía
3. Narrativa: juego como drama
4. Desafío: juego como una carrera con obstáculos
5. Compañerismo: juego como un *framework* social
6. Descubrimiento: el juego como un territorio inexplorado
7. Expresión: juego como autodescubrimiento
8. Sumisión: juego como pasatiempo

3.3 Heurísticas de usabilidad de Nielsen

Nielsen (1994) da a conocer diez principios generales para el diseño de interacciones, también llamadas *heurísticas*, dado que son reglas generales y no pautas de usabilidad específica, con el fin de identificar posibles problemas de usabilidad. Dado que son principios de interacción humano-computadora, se pueden aplicar tanto en el diseño

de páginas web como en aplicaciones móviles e inclusive juegos. Dichas heurísticas son las siguientes:

1. *Visibilidad del estado del sistema.* Indica que el diseño debe mantener informados a los usuarios sobre lo que está sucediendo, mediante la retroalimentación adecuada o durante un periodo de tiempo razonable. En síntesis, se refiere a estar actualizando y consultando el estado, en este caso, el estado del jugador en el juego y de sus diferentes componentes.
2. *Relación entre el sistema y el mundo real.* El diseño debe ser familiar para los usuarios; por eso, se sugiere utilizar palabras, frases y conceptos familiares en vez de una terminología técnica. Debido a que los principales usuarios del sistema serán niños, este principio es crucial e influyente en cuanto a la educación.
3. *Control y libertad del usuario.* Los usuarios suelen realizar acciones por error. Por ello, es necesario una “salida de emergencia” claramente visible para abandonar la acción sin tener que pasar por un largo e incómodo proceso. De esa manera, se fomenta una sensación de libertad y confianza para mantener el control del sistema y evitar sentirse atascados o frustrados.
4. *Consistencia y estándares.* Es necesario seguir y tener una convención, usar elementos comunes y que guarden relación en las interfaces del sistema. En caso de no mantener la consistencia, puede aumentar la carga cognitiva de los usuarios al obligarlos a aprender algo nuevo.
5. *Prevención de errores.* Un buen diseño evita cuidadosamente la ocurrencia de problemas en primer lugar; sin embargo, los mensajes de error son importantes. Se deben eliminar las condiciones propensas a errores o verificarlas y presentar a los usuarios una confirmación.
6. *Reconocimiento en vez de recordar.* Se sugiere minimizar la carga de memoria del usuario haciendo visibles los elementos, acciones y opciones. No debe ser necesario que el usuario recuerde información al pasar de una interfaz a otra.
7. *Flexibilidad y eficiencia de uso.* La inclusión de accesos directos podría ser relevante para aquellos usuarios ya familiarizados con el sistema, expertos.
8. *Diseño estético y minimalista.* Este principio manifiesta que las interfaces deben dar información precisa y no irrelevante o rara vez usada.
9. *Ayudar a los usuarios a reconocer, diagnosticar y recuperarse de errores.* Los mensajes de error deben mostrarse de manera sencilla, indicando con precisión el problema. En la propuesta, se presentarán con interfaces visuales que ayudarán a notarlos y reconocerlos.

10. *Ayuda y documentación.* Es recomendable que el sistema carezca de explicación adicional. Sin embargo, puede ser de mucha utilidad proporcionar documentación para ayudar a los usuarios (Nielsen, 1994).

3.4 Cuestionario de experiencia de juego

Lankoski et al. (2003) manifiestan que las entrevistas, los *focus groups* y las pruebas en juego ofrecen información relativamente profunda, pero son difíciles de implementar a mayor escala. Por el contrario, los cuestionarios se pueden aplicar fácilmente en grupos muy grandes y, si bien proporcionan menos profundidad de conocimiento que otros enfoques subjetivos, permiten centrarse en aspectos específicos de la experiencia del jugador. Es por ello que Johnson et al. (2018) explican en su trabajo que el *Game Experience Questionnaire* (GEQ) o cuestionario de experiencia de juego consta de 33 ítems y está diseñado para medir la experiencia de los jugadores en siete dimensiones: inmersión, fluidez, competencia, afecto positivo y negativo, tensión y desafío. Los elementos del cuestionario se presentan como declaraciones, que son calificadas por los encuestados para indicar su experiencia mientras juegan. Las calificaciones se realizan en una escala de Likert de 5 puntos: "0: nada", "1: levemente", "2: moderadamente", "3: bastante" y "4: extremadamente" (Johnson et al., 2018). El cuestionario se desarrolló para evaluar la experiencia del juego en múltiples intervalos durante una sesión de juego o una sesión de reproducción.

Como lo hacen notar IJsselsteijn et al. (2008), este cuestionario tiene una estructura modular y consiste en tres aspectos:

- I. *El cuestionario básico.* Es la parte central. Evalúa la experiencia del juego con puntuaciones en siete componentes: inmersión, fluidez, competencia, afecto positivo y negativo, tensión y desafío. Para una medida robusta, son necesarios cinco elementos por componente.
- II. *El módulo de presencia social.* Investiga la participación psicológica y conductual del jugador con otras entidades sociales, ya sean virtuales (es decir, personajes del juego), mediadas (por ejemplo, otros jugadores en línea) o coubicadas. Este módulo solo debe administrarse cuando al menos uno de estos tipos de cojugadores estuvo involucrado en el juego.
- III. *El módulo posterior al juego.* Evalúa cómo se sintieron los jugadores después de dejar de jugar. Este es un módulo relevante para evaluar el juego naturalista (es decir, cuando los jugadores han decidido jugar voluntariamente), pero también puede ser relevante en la investigación experimental.

Además de estos tres módulos, se desarrolló una versión concisa del GEQ dentro del juego, la cual extrae preguntas del módulo principal. Los autores afirman también

que los módulos deben administrarse inmediatamente después de que finalice la sesión de juego, en el orden indicado anteriormente. Las partes I y II sondean los sentimientos y pensamientos de los jugadores mientras juegan; la parte III, es decir, el módulo posterior al juego, evalúa cómo se sintieron los jugadores después de dejar de jugar (IJsselsteijn et al., 2008).

Finalmente, Johnson et al. (2018) sostienen que mediante el cuestionario se puede medir la experiencia del usuario en siete dimensiones, que serán utilizadas como métricas para la experiencia:

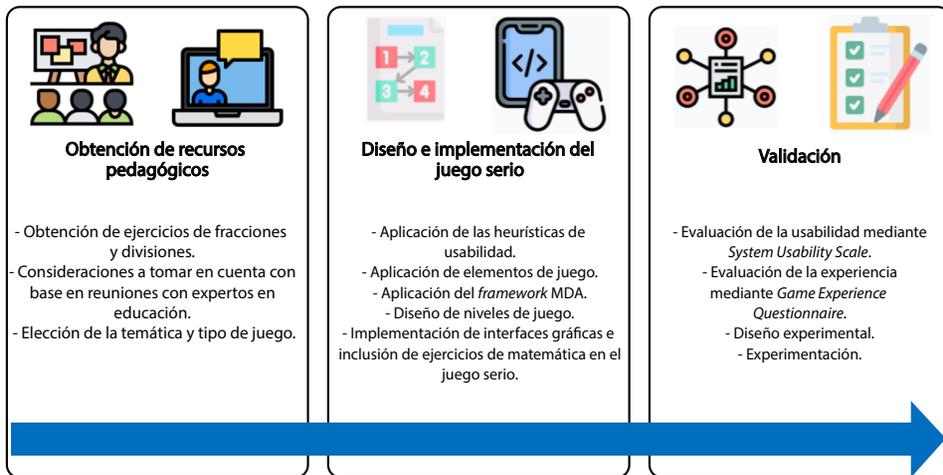
- La inmersión, también llamada *inmersión sensorial e imaginativa*, es evaluada con seis ítems que reflejan aspectos de cuán fuertemente conectados se sienten los jugadores con el juego.
- El flujo se evalúa con cinco ítems, que indican si los jugadores perdieron la noción de su propio esfuerzo o del paso del tiempo durante el juego.
- La competencia se evalúa con cinco ítems relacionados con qué tan bien los jugadores juzgaron su propio desempeño en comparación con los objetivos del juego.
- Afecto positivo (cinco ítems), relacionado con experiencias emocionales positivas.
- Afecto negativo (cuatro ítems), relacionado con experiencias emocionales negativas.
- La tensión, también conocida como *tensión/molestia*, se evalúa con tres ítems relacionados con estas emociones negativas específicas.
- El desafío se evalúa con cinco ítems, que indican el grado en que los jugadores encontraron que el juego era difícil o desafiante.

4. METODOLOGÍA

En esta sección, se presentan los pasos para el diseño, desarrollo, ejecución y validación del juego serio. La propuesta metodológica abarca la obtención de recursos pedagógicos, el diseño de interfaces gráficas aplicando elementos de juego y usabilidad, la implementación del juego serio y su validación respectiva. En la Figura 1 se observa un esquema de la metodología propuesta en el presente estudio.

Figura 1

Diagrama de la propuesta metodológica



4.1 OBTENCIÓN DE RECURSOS PEDAGÓGICOS

En esta etapa, se abordan las premisas necesarias para el desarrollo del juego, como las consideraciones que brinda la literatura revisada y las técnicas para enseñar fracciones extraídas tanto de la literatura como de profesores de matemática entrevistados. Finalmente, se hace una elección de la temática y tipo de juego correctamente sustentada.

4.1.1 Obtención de ejercicios de fracciones y divisiones

Se conversó con los profesores del área de matemática del colegio San José Obrero Marianistas, en la ciudad de Trujillo. Las reuniones se hicieron mediante videoconferencias debido al peligro de una reunión física, producto de la situación pandémica que enfrenta la sociedad actualmente. Se les preguntó cómo enseñan ellos las fracciones y divisiones, cómo las evalúan y cómo hacen didáctica su clase. De la misma manera, se les pidió apoyo con la elaboración de la evaluación que será tomada al final de la experimentación, en la comparación de grupos.

Las consideraciones de los profesores concuerdan con la investigación de Naiser et al. (2003), pues sostienen que no solo los estudiantes luchan con el aprendizaje de las fracciones, sino que también los maestros se sienten frustrados al buscar formas de enseñar este tema de manera efectiva, y es que las fracciones suelen tornarse “extremadamente difíciles de dominar para los niños”. En las reuniones se buscó identificar cuáles de las formas que los maestros implementan para enseñar

conceptos de fracciones funcionan y cuáles no. Los profesores manifestaron que, dependiendo del grado, cambian su forma de enseñar en función del desarrollo del pensamiento matemático de los niños; además, otro aspecto importante de las lecciones es el estilo de enseñanza; por ende, los maestros deben ser conscientes de que los estudiantes aprenden de diferentes maneras. Mencionaron también que construir sobre el conocimiento previo de los estudiantes es otra estrategia que ellos pueden usar para la instrucción, la cual requiere que los profesores conozcan los conceptos erróneos de los estudiantes sobre el contenido. Utilizan varias estrategias para involucrar a los estudiantes en lecciones de fracciones, como los problemas de revisión y práctica, las aplicaciones del mundo real y la construcción de conocimientos previos de los estudiantes.

4.1.2 Consideraciones importantes según expertos en educación

En primer lugar, de acuerdo con El Azizi (2019), se consideró basar el juego serio en un juego popular, como *Mario Bros*. También se tomó en cuenta el diseño de mundos, niveles y recompensas. Se adaptó también la acción de redirigir a responder una pregunta cuando colisiona con los enemigos o pierde. Del trabajo hecho por Aparício y Silva (2019), se admitió la mecánica continua utilizada en *Aithmetic Bird*, de manera que se resuelva un ejercicio de matemática mientras se juega. De Gaggi et al. (2018), se tomó la representación de fracciones en elementos familiares (en el caso de la propuesta, se incluyen elementos del mismo juego como balas y armas) y, finalmente, de Ibarra et al. (2016) se aplicó el inicio del juego con un menú gráfico. Igualmente, se tuvieron en cuenta las maneras de enseñar fracciones en cada sección del juego: la primera permite que el alumno determine gráficamente el numerador y el denominador; la segunda, que ubique una fracción en la recta numérica; la tercera, que realice operaciones básicas con fracciones; y la cuarta, que relacione fracciones con porcentajes.

La mayoría de las investigaciones revisadas considera el sesgo numérico como un problema que surge, como señalan Obersteiner et al. (2013), cuando los escolares comparan los valores numéricos de las fracciones; con frecuencia se ha encontrado que están sesgados por los números naturales involucrados (por ejemplo, creen que $1/4 > 1/3$, porque $4 > 3$); es decir, piensan en las fracciones componencialmente como dos números naturales, en lugar de holísticamente como un número. Esto se vuelve explícito en los problemas de comparación de fracciones, donde los estudiantes suelen estar sesgados por los números naturales que componen las fracciones. Así pues, un problema importante en el aprendizaje de las fracciones parece ser que el conocimiento previo de los estudiantes sobre los números naturales no siempre les ayuda a lidiar con las fracciones. Más bien, ese conocimiento puede sesgarlos y, por lo tanto, ser la fuente de un desempeño deficiente. En síntesis, como cualquier

fracción está representada por dos números naturales, depender únicamente de los valores de estos dos números naturales en lugar de los valores de las fracciones enteras puede conducir a respuestas correctas en algunos casos (como $4/5 > 3/5$, porque $4 > 3$), pero a respuestas incorrectas en otros (como $1/4 < 1/3$, aunque $4 > 3$) (Bailey et al., 2012; Obersteiner et al., 2013). Por ello, además de las características adoptadas en el diseño del juego, la propuesta de solución incluirá ejercicios que puedan evitar o minimizar el sesgo de los números naturales en los niños.

4.1.3 Elección de la temática y tipo de juego

Se hizo un *focus group* con Yanapay Educa Perú, una organización bilingüe sin fines de lucro dedicada al beneficio humanitario, que brinda clases virtuales de reforzamiento a niños de bajos recursos en todo el país, propiciando una ideología de desarrollo integral en los estudiantes. Se hizo el estudio mediante una reunión virtual en la plataforma Zoom, donde se encontraban los coordinadores del equipo y también los niños de diferentes colegios del país. Se descubrió que algunos niños no estaban al nivel de su grado académico: alumnos de quinto grado de primaria o posteriores no aprendieron bien el tema de fracciones ni otros temas esenciales en matemática, lo que indica que la educación virtual no fue muy favorable en los colegios nacionales, sobre todo por la falta de práctica.

Los alumnos mencionaron que les gustaría un juego de aventuras, similar a *Mario Bros.*, donde se pueda pelear contra enemigos y obtener puntaje. También consideraron los niveles de los juegos muy largos. De entre las temáticas tocadas se sugirieron zombis, extraterrestres, vaqueros, astronautas, etcétera.

Finalmente, luego de recopilar información, se hizo un diseño preliminar de los personajes (protagonistas y enemigos), así como de los niveles (se utilizaron recursos libres de la siguiente página web: <https://kenney.nl/assets/jumper-pack>). Se decidió que el personaje principal sería un caballero con una espada, que se encontraría en una ciudad desconocida donde aparecen extraterrestres. En la Figura 2 se visualizan los *sprites* de los personajes.

Figura 2

Personajes del juego serio



4.2 Diseño e implementación del juego serio

Esta fase corresponde a la implementación del juego serio, la cual nos permitirá evaluar el éxito, conocer la retroalimentación y validar la propuesta. El lenguaje de programación utilizado para construir el juego es C#, lenguaje de programación multiparadigma desarrollado y estandarizado por Microsoft. Adicionalmente, en el *back-end* del juego se usó el lenguaje de programación Python. También se utilizaron de manera conjunta las siguientes herramientas:

- *Unity*. Es uno de los motores de creación de juegos más utilizados en la actualidad. Es una herramienta que comprende no solo motores para el renderizado de imágenes, de físicas de 2D/3D, de audio, de animaciones y otros, sino también herramientas de *networking* para multijugador, herramientas de navegación NavMesh para inteligencia artificial o soporte de realidad virtual.
- *Tilemap*. El componente Tilemap es un sistema que almacena y maneja recursos para crear niveles 2D de manera sencilla. Transfiere la información requerida de los mosaicos colocados en él a otros componentes relacionados, como Tilemap Renderer y Tilemap Collider 2D.
- *Flask*. Es un *microframework* de Python para el desarrollo de aplicaciones web bajo el patrón MVC. Esta tecnología será utilizada para construir servicios REST, que permitan obtener y escribir datos de los jugadores. Este componente actuará como un intermediario entre el juego y la base de datos.
- *Firebase*. Es una plataforma en la nube para el desarrollo de juegos, aplicaciones web y móviles. Firebase cuenta con muchísimas funciones y

servicios, los cuales dan la facilidad de desarrollar sistemas funcionales en menor tiempo. En este caso, se utilizarán los servicios de autenticación y base de datos.

4.2.1 Aplicación de heurísticas de usabilidad

Una vez que se tienen identificados los datos, la fase posterior corresponde al diseño y elaboración de las interfaces gráficas, y al diseño de los niveles. En la Tabla 1 se describe, sobre la base de cada heurística, cómo se aplicarán las heurísticas de Nielsen en el sistema propuesto.

Tabla 1

Aplicación de las heurísticas de usabilidad de Nielsen

Heurística	Aplicación
Visibilidad del estado del sistema	Se incluyen pantallas de inicio y fin del juego. También se muestra la línea de vidas del personaje y, cuando es atacado, se tiene una retroalimentación visual, pintándolo de color rojo.
Relación entre el sistema y el mundo real	Se adoptaron propiedades físicas en el juego, de modo que se asemeje lo más posible a una persona corriendo y saltando. De la misma manera, para los mensajes en la pantalla, se usarán frases coloquiales o propias de la educación infantil.
Control y libertad del usuario	Dado que el juego cuenta con múltiples pantallas, cada una tendrá la opción de poder salir y realizar la actividad que el usuario desee, como continuar el juego, ingresar al menú o reiniciar el nivel.
Consistencia y estándares	Se tiene consistencia en los colores, niveles y elementos visuales. En cuanto a los estándares, el juego será exportado para plataformas web y móviles, por lo que se le dará enfoque a la pantalla y a que todos los elementos se puedan visualizar.
Prevención de errores	Si se tiene un error desconocido, el juego no se cerrará, sino que mostrará un modal con un mensaje, indicando que hubo un error desconocido y que es necesario volver a cargar.
Reconocimiento en vez de recordar	Se colocaron visualmente las opciones de interacción importantes en la pantalla. También los minijuegos y la mecánica contienen la misma estructura.
Flexibilidad y eficiencia de uso	El juego contará con la inclusión de ejercicios de matemática. Al momento de perder o al recoger una vida, se tendrá que resolver el ejercicio correctamente para poder obtener la vida o intentarlo una vez más.
Diseño estético y minimalista	Serán incluidos, únicamente, los elementos fundamentales y se evitará saturar las interfaces gráficas con una gran cantidad de elementos, de forma que el usuario no se pierda ni se aburra.

(continúa)

(continuación)

Heurística	Aplicación
Ayudar a los usuarios a reconocer, diagnosticar y recuperarse de errores	Se utilizará al momento de realizar consultas a la base de datos o en aquellas acciones donde puedan existir errores.
Ayuda y documentación	Será implementada en la pantalla de galería de niveles, donde se tiene el tutorial y las acciones que el usuario tiene permitido hacer, de manera breve y concisa.

4.2.2 Aplicación de elementos de juego

Nah et al. (2014) realizaron una revisión de la literatura e identificaron ocho elementos de diseño de juegos que se utilizan ampliamente en contextos educativos y de aprendizaje: puntuaciones, niveles, medallas/trofeos, tablas de posiciones, premios y recompensas, barra de progreso, historia y, finalmente, retroalimentación.

Para la propuesta de solución, se escogieron cinco de estos elementos de juego, que se describen brevemente, indicando cómo y por qué se aplicaron en la propuesta de solución.

- *Puntuaciones.* Funcionan como un indicador de éxito o logro y pueden usarse como recompensas, progreso o posición del jugador. En la propuesta de solución, se utilizará la puntuación al momento de resolver correctamente un ejercicio de matemática, al atacar a un enemigo o al esquivar una bala con la espada del protagonista. La puntuación obtenida es una de las métricas más importantes porque mide el desempeño a lo largo del nivel y se va acumulando por cada jugador para luego clasificar la puntuación total.
- *Niveles/etapas.* Se utilizan en muchos juegos para dar una sensación de progreso. Se tiende a requerir menos esfuerzo en los niveles iniciales, que son más rápidos de superar en comparación con los niveles más avanzados. En la implementación, los niveles sirven para dar una sensación de progreso de la dificultad, donde el estudiante irá desde ejercicios con divisiones hasta la comparación de fracciones. Este elemento es importante, pues el desarrollo de las habilidades del estudiante debe ser incremental.
- *Tablas de clasificación.* El objetivo de este elemento es mantener motivados y entusiasmados a los alumnos por querer visualizar su nombre en la lista. Las tablas de clasificación sirven para crear un entorno competitivo y para mostrar los niveles actuales de los puntajes más altos. En el juego, las tablas de posiciones harán uso del puntaje total (acumulado entre cada intento y cada nivel) para establecer un tablero con los *nicknames* de los cinco mejores jugadores.

- Retroalimentación.* Es un elemento sumamente útil, junto con la frecuencia e intensidad, para la participación del alumno, pues cuanto más frecuente e inmediata sea la retroalimentación, mayor será la afectividad del aprendizaje y el compromiso. La retroalimentación será incluida en el sistema al momento de responder un ejercicio de matemática (la casilla de la respuesta se pinta de verde o rojo según su respuesta), y también se mostrará un texto con un mensaje al momento de finalizar el nivel o perder. La retroalimentación es clave para que el jugador pueda entender cómo se está desempeñando al resolver los ejercicios de matemática propuestos.

4.2.3 Aplicación del framework MDA

Para la aplicación del *framework* MDA, se utilizó la plantilla de evaluación de PECC (Spieler & Girvan, 2020), la cual fue inicialmente propuesta por Spieler (Spieler & Kemeny, 2020) y permite clarificar los diferentes elementos del juego y las capas del *framework* mediante un *storyboard*. En la Tabla 2, se puede apreciar la aplicación del *framework* con base en la plantilla sugerida.

Tabla 2

Storyboard de aplicación del framework MDA

Nombre del juego	<i>Mathesis</i>
Personaje principal	Caballero
Cómo se juega	Se tiene un ambiente apocalíptico, infestado de extraterrestres que salen con el objetivo de disparar al personaje principal. El caballero perdió su caballo y le queda un largo camino a casa en el que tiene que recolectar ítems para regresar a su hogar. El jugador puede moverse a lo largo del nivel e ir luchando con los enemigos y recolectando ítems para obtener más puntaje.
Género/tema	Arcade, minijuegos/ciencia ficción.
Objetivo	El objetivo principal del juego es que el jugador refuerce sus habilidades de fracciones de forma divertida mientras juega. El objetivo en el juego es que el caballero pueda regresar a su hogar y obtener el mayor puntaje posible.
Mecánica	Conforme va recogiendo corazones e ítems, el jugador va resolviendo ejercicios y, según su respuesta, se le otorga o no una recompensa. Cada 500 puntos se le aparecerá un minijuego con el objetivo de obtener más puntaje. Además, a partir del segundo nivel, al final se le aparecerá un gran enemigo con cinco vidas; cuando se lo ataca, surge un ejercicio y la vida disminuirá o se mantendrá según su respuesta. Finalmente, si el jugador pierde, ya sea porque se cayó o le dispararon, tiene la oportunidad de resolver un ejercicio y continuar aumentando su puntuación.

(continúa)

(continuación)

Dinámica	<p>Vida: el jugador comienza cada nivel con tres corazones de vida (el máximo número de corazones posible) y va disminuyendo conforme le disparan.</p> <p>Puntos: el jugador obtiene puntos mientras va atacando enemigos y va recolectando ítems.</p> <p>Recompensas: se le otorga un corazón de vida en caso de que no tenga las vidas completas; de lo contrario, gana 150 puntos.</p>
Estética	<p>Retroalimentación sonora: al disparar, los enemigos emiten un sonido de disparo láser. El jugador emite sonidos al saltar, atacar, obtener ítems y al finalizar el nivel.</p> <p>Retroalimentación visual: se puede notar cómo el jugador se anima mientras se mueve igual que los enemigos. Las trampas y plataformas cuentan también con animaciones y movimiento, al igual que al atacar, saltar y disparar.</p> <p>Desafío: juego como una carrera de obstáculos.</p> <p>Su misión: juego como pasatiempo.</p>

Nota. Elaboración propia basada en la plantilla de Spieler & Kemeny (2020).

4.2.4 Interfaces iniciales en el juego

La interfaz principal contiene la tabla de posiciones. Al momento de ingresar al juego, se ve una tabla con los mejores jugadores y abajo dos botones: uno invitando al jugador a ingresar a jugar y el otro a registrarse en caso de que no tenga un usuario.

Respecto al ingreso, el usuario podrá acceder a la aplicación con su nombre de usuario y su contraseña, previamente registrados en la pantalla de registro. La interfaz de registro cuenta con un campo adicional que permite verificar la contraseña. El registro y el logueo hacen posible que los usuarios puedan almacenar su progreso en la nube, garantizando disponibilidad, integridad y seguridad. Asimismo, el registro en la nube sirve para almacenar información sobre el tiempo de juego, si se equivocó en una pregunta o no y sus puntajes respectivos, con el objetivo de utilizar esa información para la emisión de los resultados del presente estudio.

Luego de acceder con sus credenciales, el jugador visualizará un menú que contiene los niveles disponibles. Cada nivel tiene una temática diferente y no se puede jugar el siguiente nivel si no se completa el anterior.

En la Figura 6, se puede observar las interfaces de logueo y registro. Cabe anotar que las pantallas se muestran en su primera versión; como parte del desarrollo iterativo del juego, este se encuentra sujeto a modificaciones.

Figura 3

Interfaces de inicio, ingreso y niveles disponibles



4.2.5 Diseño de los niveles del juego

En el diseño de los niveles del juego, se introdujo una nueva forma de reforzar las habilidades aritméticas a través de la combinación de educación y entretenimiento en un juego simple, haciendo un esfuerzo por permitir a los alumnos aprender e interactuar de manera efectiva con el mundo que los rodea.

Una vez que se tuvo definida la mecánica, se procedió con el diseño de los niveles y su contenido sobre la base del *framework* MDA. Los niveles del juego serio son el escenario donde ocurren las interacciones entre el usuario y los elementos de juego. Este diseño es de suma importancia, especialmente para los usuarios objetivo, que son niños de primaria, por lo que debe ser simple, fácil de entender y navegar, así como autoexplicativo. En los niveles se incluyen los ejercicios de matemática para el refuerzo en forma de minijuegos.

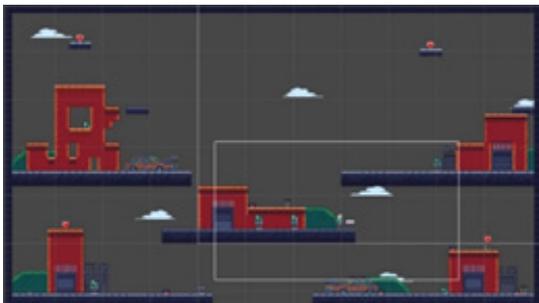
Se utilizó la herramienta Tilemap para el diseño de los niveles. Se escogen los elementos que irán en el nivel y se arrastran en la escena. Tilemap no permite la generación de niveles, por lo que el nivel y sus elementos (terreno, trampas, etcétera) permanecen de manera estática; lo que cambia a lo largo del nivel es la posición de los ítems que se deben coleccionar; además, los minijuegos y los ejercicios de matemática son distintos según el nivel y la dificultad.

El juego se encuentra compuesto de tres niveles:

- *Nivel 1 (fácil)*. Este nivel se encuentra ambientado en un escenario apocalíptico, donde se encuentran repartidos ítems de vida y dos hordas generadoras de enemigos. En la Figura 3 se observa el diseño del primer nivel del juego.

Figura 4

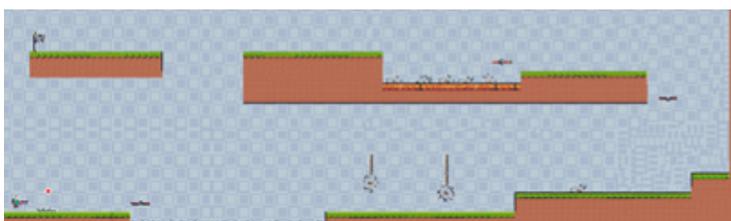
Diseño del primer nivel del juego



- *Nivel 2 (medio)*. Este nivel se encuentra ambientado en un paisaje donde se simula el camino de regreso a casa para el caballero. A lo largo del nivel, además de los enemigos, se incluyeron plataformas móviles y trampas como sierras móviles y espinas. Este nivel busca disminuir el sesgo de los números naturales, de manera que se consideren las fracciones como un número, en lugar de dos números naturales distintos. Por ello, se muestran preguntas relacionadas con las fracciones, como la identificación del numerador y el denominador. En la Figura 5 se observa el diseño del segundo nivel del juego.

Figura 5

Diseño del segundo nivel del juego

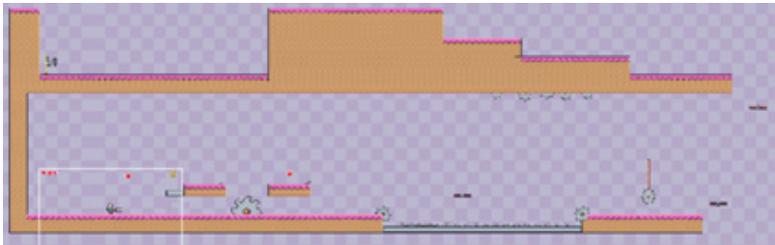


- *Nivel 3 (difícil)*. En la validación de la experiencia, realizada en la primera fase del estudio, se encontró que las niñas no estaban entusiasmadas ni interesadas en atacar a los enemigos tanto como los niños. Por ello, en este nivel se incluyeron pocos enemigos y, en su lugar, se colocaron más ítems a lo largo del nivel, así como trampas y desafíos. Este nivel busca que el estudiante refuerce la comparación de fracciones y la ubicación de los números en la recta numérica; por esta razón, se presentan todos los minijuegos, dándoles

prioridad a los que se encuentran enfocados en el orden y la comparación de las fracciones. En la Figura 6 se observa el diseño del tercer nivel del juego.

Figura 6

Diseño del tercer nivel del juego

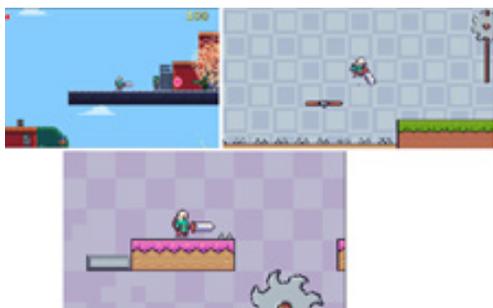


4.2.6 Juego funcional

El juego es sencillo y fácil de entender: el jugador aparece al inicio del nivel con una espada, vidas completas y una puntuación de 0; a lo largo de cada nivel, tiene que ir recolectando ítems, como monedas o corazones de vida, a cambio de puntaje. De la misma manera, el jugador puede defenderse esquivando las balas y atacando a los enemigos. Si ataca a la bala grande, esta irá de regreso a atacar al enemigo y el jugador obtendrá 50 puntos. En cambio, si este ataca a los enemigos, estos automáticamente desaparecen y se le otorga 100 puntos al jugador. El objetivo es obtener el mayor puntaje posible e ir recolectando la mayor cantidad de corazones e ítems a lo largo del nivel. Cada 500 puntos, el juego será pausado y se mostrará un menú emergente con uno de los ejercicios propuestos. A partir del segundo nivel, encontrará un gran enemigo al final: para poder superar el nivel, tiene que atacar al enemigo y dejarlo sin vidas. Su número de vidas es el número de ejercicios de matemática que el jugador tiene que resolver correctamente para pasar de nivel. En la Figura 7 se puede observar el juego en acción a lo largo de los niveles.

Figura 7

El juego en acción



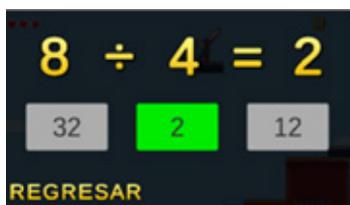
4.2.7 Incorporación de ejercicios de matemática en el juego

A lo largo del juego, mientras el jugador se encuentra peleando contra los enemigos o esquivando las trampas, puede ir recogiendo corazones que le suman una vida cada uno; sin embargo, al recoger el ítem, deberá responder un ejercicio de matemática en forma de minijuego a cambio de un corazón más en su línea de vida o de puntaje. Para todos los minijuegos, primero se obtienen las fracciones de la base de datos, según la dificultad, y se establece una lógica de respuesta. Si el jugador acierta, la casilla de la opción se pinta de color verde y, después de unos segundos, se cierra el menú emergente y se le suma la vida o el puntaje. Caso contrario, la opción se pinta de rojo y, luego de cerrarse, no se le otorga nada. A continuación, se describe cada minijuego y se muestra un ejemplo de estos en las Figuras 8, 9 y 10.

- *Minijuego de nivel fácil.* Incluye preguntas y minijuegos relacionados con las divisiones, con el objetivo de reforzar esta habilidad como paso previo a los ejercicios con fracciones. En este tipo de minijuego, se muestra una división exacta y se le pide al alumno que elija la respuesta correcta; el sistema se encarga de generar respuestas erróneas de manera aleatoria.

Figura 8

Minijuego de nivel fácil



- *Minijuego de nivel medio.* Se muestran interfaces con objetos familiares a lo largo del juego, pidiéndole al jugador que indique la fracción de los objetos que se encuentran o no pintados. También se puede presentar una serie de elementos aleatorios y solicitarle que indique la fracción de los elementos de cierto tipo.

Figura 9

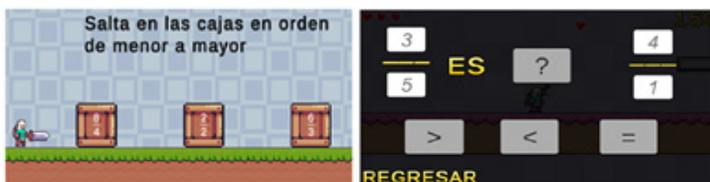
Minijuegos de nivel medio



- *Minijuegos de nivel difícil.* En este nivel, se tienen preguntas enfocadas en la comparación de fracciones y el orden. Se obtienen varias fracciones de la base de datos y se las coloca en desorden, solicitándole al jugador que salte en las cajas en el orden correcto. También se propone un minijuego que permite comparar dos fracciones e indicar si la fracción de la izquierda es mayor, menor o igual que la fracción de la derecha. En este último minijuego, se le sugirió al participante hacer la comparación de fracciones multiplicando el numerador de la primera fracción por el denominador de la segunda, y comparar el resultado con el producto del denominador de la primera fracción y el numerador de la segunda; esta técnica fue mencionada por los profesores en la primera fase de la metodología.

Figura 10

Minijuegos de nivel difícil



4.2.8 Gran enemigo al finalizar el nivel

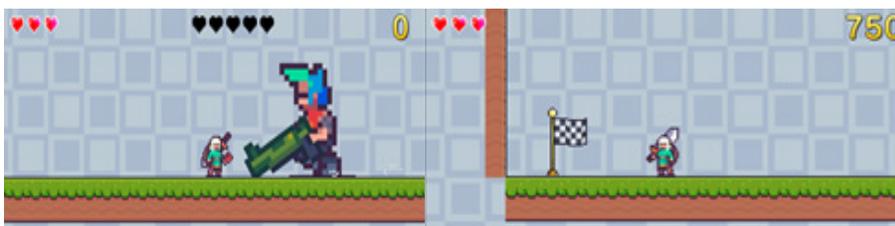
A partir de los niveles 2 y 3, se añadió una mecánica diferente: se presenta un gran enemigo al finalizar el nivel. El objetivo de este gran enemigo es hacer una breve

evaluación, en la que se aplica todo lo reforzado a lo largo del nivel. Este es el paso determinante para la finalización del nivel, pues si no se logra combatir al gran enemigo, el estudiante no podrá visualizar el trofeo y pasar al siguiente nivel.

Cada vez que el caballero ataca al gran enemigo, se le muestra al estudiante el tipo de minijuego correspondiente al nivel, el cual contiene un ejercicio de matemática. El gran enemigo cuenta con cinco vidas y pierde una si la respuesta del estudiante es correcta; caso contrario, no sucede nada. Los ejercicios propuestos en esta última parte serán almacenados en la base de datos de tipo *boss* y servirán para hacer seguimiento al progreso del estudiante a lo largo del juego. En la Figura 11 se observa al gran enemigo y el trofeo que aparece una vez que este es derrotado.

Figura 11

Gran enemigo al finalizar el nivel



4.3 Validación

La fase final de la metodología es la validación de la propuesta de solución, para lo cual previamente se definen las métricas para evaluar la investigación. Se propone evaluar tres criterios: la usabilidad del sistema, la experiencia del jugador después del juego y la efectividad del juego serio para el refuerzo de las habilidades de fracciones y divisiones.

4.3.1 Evaluación de usabilidad

Respecto a la usabilidad, es necesario validar que las heurísticas de usabilidad de Nielsen estén correctamente aplicadas, pues es un proceso crítico para el desarrollo de juegos serios efectivos. Además, como señalan Procci et al. (2012), los juegos serios presentan desafíos únicos porque deben ser utilizables, divertidos y efectivos. Por eso, el diseño del juego se debe evaluar en varias etapas durante el desarrollo para garantizar que se cumplan todos esos objetivos.

La usabilidad, que es uno de los elementos centrales en el proceso de desarrollo del juego, se encuentra profundamente relacionada con la experiencia general del jugador y puede influir en su interacción con el juego. Por este motivo, será medida mediante la *System Usability Scale* o Escala de Usabilidad del Sistema, en español, o simplemente SUS, por sus siglas en inglés, que es una herramienta metodológica muy similar a la escala de Likert, utilizada para medir la usabilidad de un objeto, dispositivo o aplicación. Consta de un cuestionario que mide la usabilidad de forma numérica en una escala del 0 al 100. Está formado por diez preguntas, de las cuales las impares tienen una tonalidad positiva y las pares una tonalidad negativa. Por lo general, se le pide al usuario que evalúe del 1 al 5 cada pregunta (siendo 1 totalmente en desacuerdo y 5 totalmente de acuerdo) sobre cómo percibe aspectos del sistema. Finalmente, tras un estudio realizado en más de 500 webs y aplicaciones, Sauro e Lewis (2016) llegan a la conclusión de que el puntaje promedio es 68, estableciendo un umbral de aceptación. Esto quiere decir que un resultado por debajo de esta cifra indica que hay varios aspectos por corregir.

4.3.2 Evaluación de la experiencia mediante el *Game Experience Questionnaire*

Tal como expresan IJsselsteijn et al. (2008): “Uno de los principales desafíos que enfrenta la comunidad de investigación de juegos es la falta de un conjunto coherente y detallado de métodos y herramientas que permitan medir las experiencias de entretenimiento de una manera sensible, confiable y válida”. Es por ello que el primer paso para la validación de la propuesta es la aplicación del *Game Experience Questionnaire*, cuya finalidad es medir la experiencia y los sentimientos del jugador al momento de jugarlo.

4.3.3 Diseño experimental

Para la evaluación de la efectividad de la propuesta, se hará una comparación de dos grupos de estudiantes. Debido a la situación pandémica que enfrenta la sociedad actualmente, no se pudo probar el sistema de manera presencial; además, se conformaron grupos con una cantidad reducida de participantes.

Para el grupo experimental, se procederá a calcular las estadísticas descriptivas por nivel en cuanto al puntaje. Luego, se calculará un puntaje total por alumno, un porcentaje de respuestas correctas y tiempo promedio por pregunta, así como el tiempo promedio por nivel.

Con el objetivo de determinar si la propuesta es efectiva para el refuerzo de las habilidades aritméticas de fracciones y divisiones en niños de primaria, se llevará a cabo una prueba de hipótesis con un estadístico de prueba establecido por la distribución normal o z , porque las muestras son independientes. Sin embargo,

el presente estudio, además de la limitación de la pandemia, tiene otras como la reducida cantidad de participantes en ambos grupos y el amplio rango de edades de estos, por lo que se sugiere que los resultados no son del todo confiables. Los resultados emitidos en este estudio serán tomados como punto de partida para un análisis preliminar.

Se establecieron las hipótesis nula y alternativa:

- H_0 : los niños que reforzaron sus habilidades aritméticas de fracciones y divisiones tuvieron un menor rendimiento respecto a aquellos que reforzaron sus habilidades de forma tradicional.
- H_a : los niños que reforzaron sus habilidades aritméticas de fracciones y divisiones tuvieron un mayor rendimiento respecto a aquellos que reforzaron sus habilidades de forma tradicional.

El nivel de rendimiento se mide en función de los puntajes obtenidos por los participantes en la prueba posterior. Luego, se procede a calcular el nivel de significancia, es decir, la máxima cantidad de error que se está dispuesto a aceptar para dar como válida la hipótesis. Debido a que es el valor más utilizado para este tipo de prueba, se estableció el nivel de significancia en 0,05 ($\alpha = 0,05$). Seguidamente, se determinó un valor crítico de 1,64, que se obtuvo utilizando la tabla de probabilidad de la distribución normal, donde se busca el nivel de significancia establecido y una aproximación al 95 %. Para concluir, se procede a definir la función pivotal como la fórmula que va a involucrar el análisis de los datos obtenidos de la muestra. Como la muestra es menor que 30 y se trata de una prueba de hipótesis con varianzas desconocidas y diferentes, en estos casos, lo conveniente es utilizar la prueba t de Student para dos muestras independientes. Esta prueba compara las dos medias de una variable de resultado cuantitativo y se basa en el cálculo del estadístico t , que tiene en cuenta la diferencia entre las medias comparadas y su error estándar, según la siguiente fórmula:

$$t_0 = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n_2}}} \quad (1)$$

Donde \bar{x} es la media de los puntajes obtenidos por los participantes del grupo experimental, σ_x es su desviación estándar y n_1 es el tamaño de la muestra de este grupo. En cambio, \bar{y} es la media de los puntajes obtenidos por el grupo de control, σ_y es su desviación estándar y n_2 es el tamaño de su muestra. Se rechazará la

hipótesis nula si el valor del estadístico de prueba pertenece a la región crítica (el valor del estadístico de prueba es mayor que el valor crítico).

4.3.4 Experimentación

- *Participantes.* Como señalan Olsen et al. (2011), en una investigación hecha por Nielsen se estableció que emplear cinco usuarios en las pruebas de usabilidad permite encontrar casi tantos problemas de usabilidad como los que se hallarían empleando muchos más participantes. Tomando como base dicha afirmación, se tuvo un mínimo de cinco niños para validar la propuesta. Las condiciones de reclutamiento, en principio, eran estar en tercer o cuarto grado de primaria. Sin embargo, dada la coyuntura actual y el hecho de que hay niños en grados superiores que no se encuentran al nivel de su grado (como se ha mencionado anteriormente), estas condiciones se modificaron para incluir a niños a partir del tercer grado de primaria y que quisieran reforzar el tema de fracciones y divisiones. Cabe recalcar que la participación en el experimento fue totalmente voluntaria por parte de los niños y que ellos podían retirarse de este en cualquier momento. Los participantes fueron reclutados mediante el trabajo en conjunto con Yanapay Educa Perú, pero también se hizo un anuncio en redes sociales para convocar a más usuarios. En la validación, participaron nueve niños: tres de 10 años, uno de 11 años, uno de 12 años y cuatro de más de 12 años. En su mayoría, los participantes provenían de instituciones públicas y solo un pequeño porcentaje pertenecía a instituciones privadas.
- *Instrumentos.* Se hizo la experimentación mediante la comparación de dos grupos de estudiantes (como se mencionó en la sección anterior); de ahí se sacaron estadísticas, se calculó el estadístico de prueba y se buscó comprobar o rechazar la hipótesis planteada. Para la evaluación de la usabilidad, se utilizó la *System Usability Scale* como instrumento para obtener un puntaje general de la usabilidad del juego. Finalmente, para la validación de la experiencia, se usó el *Game Experience Questionnaire*.
- *Procedimiento.* Con respecto a la realización de pruebas, cuyo objetivo es medir tanto la eficacia de la propuesta de solución como la experiencia del jugador, el primer paso fue la validación mediante una sesión en videoconferencia con ambos grupos, los tutores de Yanapay y los padres de familia, con el objetivo de explicarles el procedimiento del experimento, las instrucciones y, antes de iniciar la reunión, se le pidió al tutor que firmara el Consentimiento Informado de Menores de Edad. Se programaron reuniones tanto individuales como grupales con los estudiantes, de acuerdo con su

disponibilidad, y se les recordaba a sus padres que sus hijos debían usar con frecuencia el sistema. Los niños entraban a las sesiones acompañados de sus padres o tutores. Debido al peligro de una reunión física por la pandemia que enfrenta la sociedad actualmente, toda sesión fue realizada de manera virtual mediante un programa gratuito de videoconferencia llamado Zoom. El juego estuvo disponible para web y aplicación móvil; por ello, en la reunión se dieron las indicaciones ya sea para entrar a la web o para instalar el aplicativo móvil. El grupo experimental reforzó sus habilidades jugando durante dos semanas, mientras que los alumnos del grupo de control no hicieron uso del juego en ese tiempo. Este último grupo realizó el refuerzo de sus habilidades mediante la metodología tradicional por dos semanas también, a través de los ejercicios y asesoría brindada por Yanapay. Una vez culminada esta fase, se tomó una evaluación a ambos grupos y se compararon los puntajes obtenidos. Asimismo, conforme el alumno iba jugando, el juego guardaba data como la pregunta planteada, si el jugador se equivocó o no y el tiempo empleado para resolver el ejercicio y pasar el nivel, la cual servirá para el posterior análisis.

Luego, se procedió a validar la usabilidad mediante la *System Usability Scale* y la experiencia mediante el *Game Experience Questionnaire*. Debido a que los niños que no utilizaron el sistema también deseaban probarlo, se les pidió a ambos grupos interactuar con el sistema sin imponer un límite de tiempo y revisar todas las funciones. El investigador estuvo presente todo el tiempo para aclarar dudas. Una vez concluida la experimentación, se les entregó ambos cuestionarios y se les pidió responder con la mayor sinceridad posible.

5. RESULTADOS

En esta sección, se presentan los resultados más relevantes de la experimentación. Primero, se exponen los resultados de usabilidad de acuerdo con la *System Usability Scale*, y luego, los de la aplicación del *Game Experience Questionnaire*; finalmente, se describe la efectividad del juego serio para el reforzamiento de matemática.

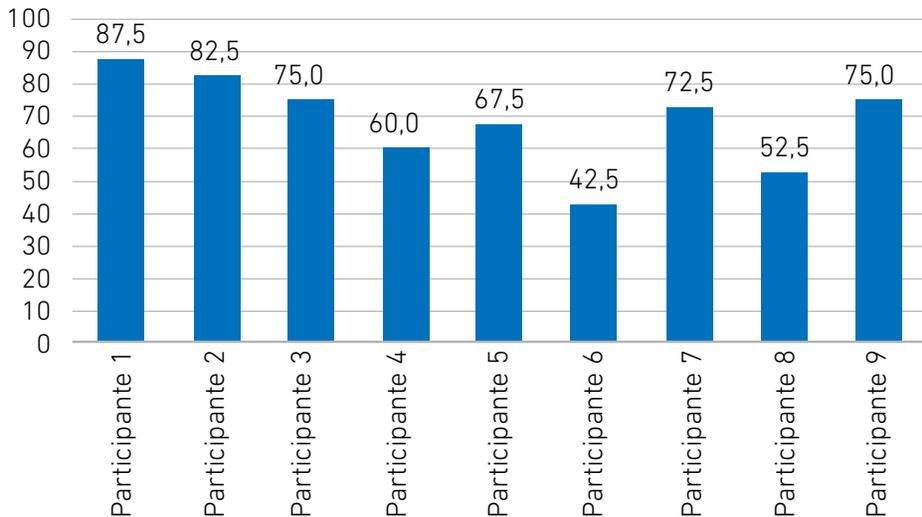
5.1 Usabilidad del juego

Este criterio fue evaluado a través de una adaptación al español de la *System Usability Scale* (SUS), entregada a todos los participantes de la experimentación tras haber interactuado con el juego, con la finalidad de validar que el juego sea usable y tenga correctamente aplicadas las heurísticas. En la Figura 12 se muestran los resultados por participante, donde se puede observar que el menor puntaje obtenido fue de 42,5, mientras que el más alto fue de 87,5, con una desviación estándar de 14,42. El

puntaje de usabilidad promedio fue de 68,33. Si bien este resultado se sitúa sobre el umbral de aceptación, lo pasa con las justas, lo que indica que la usabilidad es aceptable.

Figura 12

Resultados del puntaje SUS por participante



5.2 Experiencia de juego

La Tabla 3 presenta las estadísticas descriptivas correspondientes al *Game Experience Questionnaire* de siete factores. Para la elaboración de las estadísticas, se clasificó con números la escala de Likert (1 = para nada, 2 = levemente, 3 = moderadamente, 4 = bastante, 5 = extremadamente). Luego, se agrupó cada pregunta según la dimensión a la que pertenecía y se extrajo para cada pregunta el mínimo y máximo valor, su media y su desviación estándar. Finalmente, las puntuaciones de los componentes se calculan como el valor medio de sus elementos.

5.2.1 Módulo principal

Como señalan IJsselsteijn et al. (2008), el cuestionario en el juego está desarrollado para evaluar la experiencia del juego en múltiples intervalos durante una sesión de juego o una sesión de reproducción. Esto debería facilitar la validación de indicadores continuos y en tiempo real.

Como se mencionó antes, se intenta medir la experiencia de juego en siete componentes. Para obtener las estadísticas descriptivas de cada componente,

IJsselsteijn et al. (2008) proponen que se debe agrupar cada dimensión según el número de pregunta asignada en el cuestionario y calcular las puntuaciones de los componentes como el valor medio de sus elementos.

Tabla 3

Estadísticas descriptivas para el cuestionario de experiencia de juego

Dimensión	Mínimo	Máximo	Media	Desviación estándar	Preguntas
Competencia	3,20	5,00	4,28	0,81	2, 10, 15, 17 y 21
Inmersión sensorial e imaginativa	1,33	5,00	3,93	1,09	3, 12, 18, 19, 27 y 30
Flujo	1,60	5,00	3,49	1,11	5, 13, 25, 28 y 31
Tensión	1,00	5,00	1,48	0,96	22, 24 y 29
Desafío	1,00	4,80	2,42	1,14	11, 23, 26, 32 y 33
Afecto negativo	1,00	4,50	1,67	1,00	7, 8, 9 y 16
Afecto positivo	1,60	5,00	4,11	0,97	1, 4, 6, 14 y 20

5.2.2 Módulo in-game

Por otro lado, la versión en el juego del GEQ es una versión concisa del cuestionario principal. Tiene una estructura de componentes idéntica y consta de elementos seleccionados de este módulo, los cuales permiten medir la experiencia de juego durante el mismo.

Para este módulo, se aplicó la técnica mencionada en la parte 3.2.1. En la Tabla 4 se muestra la agrupación de preguntas para este módulo.

Tabla 4

Estadísticas descriptivas para el cuestionario de experiencia de juego, módulo in-game

Dimensión	Mínimo	Máximo	Media	Desviación estándar	Preguntas
Competencia	1,00	5,00	3,69	1,17	2 y 9
Inmersión sensorial e imaginativa	1,00	5,00	3,86	1,23	1 y 4
Flujo	2,00	5,00	3,81	1,05	5 y 10
Tensión	1,00	5,00	1,45	0,95	6 y 8
Desafío	1,00	5,00	2,86	1,21	12 y 13
Afecto negativo	1,00	4,50	1,62	0,96	3 y 7
Afecto positivo	2,00	5,00	4,10	0,92	11 y 14

5.3 Efectividad del juego serio educativo para el reforzamiento de matemática

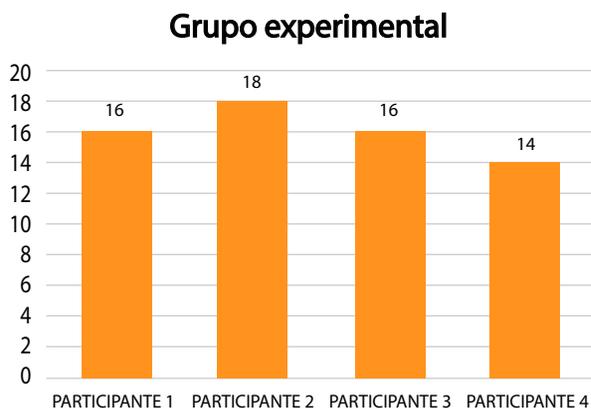
A continuación, en la Figura 13 se presentan los resultados de la prueba posterior, la cual evalúa lo reforzado por los estudiantes durante el experimento. Para ello, cada pregunta tenía un puntaje de 2, y la máxima nota era 20 y la mínima 0. En la figura, se puede apreciar dos gráficos de barras con los puntajes obtenidos por cada participante en ambos grupos. Las barras de color naranja corresponden al grupo experimental, que reforzó sus habilidades aritméticas de fracciones y divisiones mediante el uso del juego serio, mientras que el grupo de control reforzó sus habilidades a través de metodologías tradicionales.

Con estos resultados, se puede hacer una descripción preliminar. En el grupo experimental, se observa que la nota más alta fue 18, mientras que en el grupo de control fue 14. Esto representa una diferencia de 4 puntos a favor del grupo experimental y una variación porcentual de 28,57 % del grupo experimental respecto al de control y de 22,22 % del grupo de control respecto al experimental. Por otro lado, la nota más baja del grupo experimental es 14 y la del grupo de control es 10. Esto representa una diferencia de 4 puntos a favor del grupo experimental y una variación porcentual de 40 % del grupo experimental respecto al grupo de control y de 28,57 % del grupo de control respecto al experimental.

El puntaje promedio de los participantes del grupo experimental es 16, mientras que el de los participantes del grupo de control es de 11,6. En este caso, se encuentra una diferencia de 4,4 puntos a favor del grupo experimental y una variación porcentual de 37,93 % del grupo experimental respecto al grupo de control y de 27,5 % del grupo de control respecto al grupo experimental. Asimismo, considerando todos los puntajes del grupo experimental, se consigue una desviación estándar de 1,63 puntos, mientras que en el grupo de control se obtiene 1,67 puntos, lo que significa una menor dispersión de los puntajes para el grupo experimental.

Figura 13

Resultados de la prueba posterior a la experimentación por participante



Finalmente, se tiene que el promedio de los puntajes del grupo experimental fue de 16, con una desviación estándar de 1,63, mientras que el grupo de control obtuvo 11,6 de promedio, con una desviación estándar de 1,67. Con estos valores, se calculó el estadístico t (ecuación 1) y se obtuvo un valor de 3,98. Este, al superar el valor crítico de 1,6, hace que se rechace la hipótesis nula y se dé validez a la hipótesis alternativa: los niños que reforzaron sus habilidades aritméticas de fracciones y divisiones tuvieron un mayor rendimiento con respecto a aquellos que reforzaron sus habilidades de forma tradicional.

6. DISCUSIÓN DE LOS RESULTADOS

El propósito general de esta investigación es el apoyo del proceso de reforzamiento del aprendizaje de fracciones y divisiones a través de un juego serio educativo, mediante la mejora de la rapidez y habilidad al momento de resolver problemas. Este propósito se lograría luego de la implementación del juego serio con distintos elementos de usabilidad y de juego. Por ello, la sección de discusión establece un análisis y una comparación con la literatura de los resultados obtenidos en la sección anterior para poder interpretarlos y determinar si el sistema cumplió con los objetivos planteados.

En cuanto a la usabilidad del juego serio, de acuerdo con UX Research (2017), la puntuación SUS indica su rendimiento de usabilidad en los aspectos de eficacia, eficiencia y facilidad de uso general. Aunque cada respuesta arroja una puntuación en una escala de 0 a 100, no se debe confundir con un porcentaje o percentil. El puntaje promedio del SUS es 68, lo que significa que un puntaje de 68 lo colocará en el percentil 50. La Tabla 5 muestra la pauta general sobre la interpretación del puntaje SUS, según UX Research.

Tabla 5

Interpretación de la System Usability Scale

Puntaje SUS	Interpretación
> 80,3	Excelente
68-80,3	Bien
68	Aceptable
51-68	Pobre
< 51	Deficiente

Nota. Adaptado de *Measuring and Interpreting System Usability Scale (SUS) - UIUX Trend*, por UX Research, 2017, UIUXTrend (<https://uiuxtrend.com/measuring-system-usability-scale-sus/>).

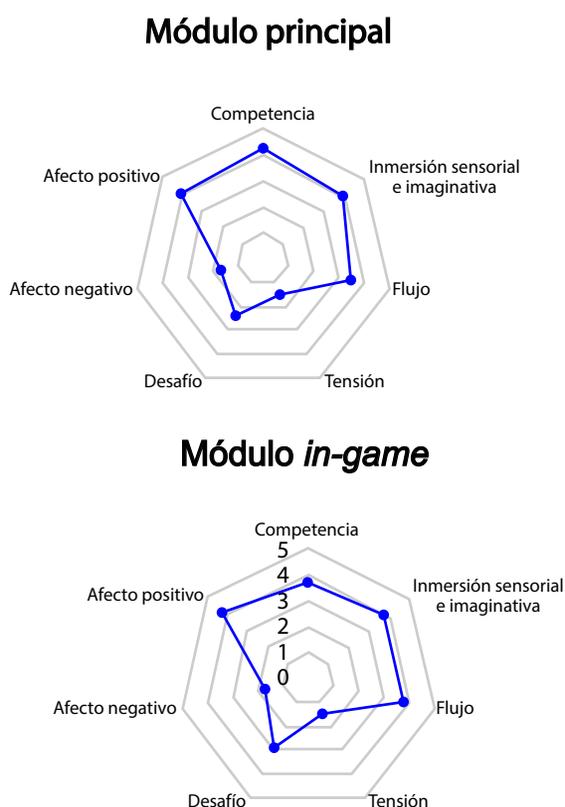
De acuerdo con esta interpretación, el puntaje SUS correspondiente al presente estudio se encuentra situado al medio, lo que indica que se aplicó bien este aspecto. De la misma manera, comparando este puntaje con el estudio de Murthy et al. (2020), el juego serio se encuentra casi 2 puntos por debajo de su propuesta, quienes consiguieron un puntaje de 70,119. En consecuencia, el puntaje obtenido podría señalar un correcto uso tanto de las heurísticas de usabilidad como de los elementos de juego seleccionados. Sin embargo, debido a la reducida cantidad de la muestra, se obtuvo una desviación estándar de 14,42, lo cual podría significar

una mayor dispersión de los datos. Por otra parte, se considera que la usabilidad podría mejorarse mediante la realización de un estudio que incluya otra herramienta para medir la usabilidad y establecer una comparación con la actual, así como con la incorporación de un especialista en diseño de juegos serios educativos y de interfaces en la fase de diseño e implementación.

Respecto a la experiencia de juego, en la Figura 14 se puede visualizar cómo es la experiencia de juego en las siete dimensiones para ambos módulos: principal e *in-game*.

Figura 14

Experiencia de juego por cada dimensión del GEQ, módulos principal e in-game



En la gráfica correspondiente al módulo principal (la primera en la Figura 14), se puede observar que los participantes juzgaron su propio desempeño en comparación con los objetivos del juego. El aspecto más llamativo fue este, correspondiente

a la dimensión de competencia (4,28), cuya media es mayor. Le sigue el afecto positivo, que obtuvo una media de 4,11, la cual sugiere que los jugadores tuvieron una percepción positiva del juego serio. Se esperaba que la competencia fuera la dimensión con menor puntaje, debido a que no se probó el juego físicamente, de manera que los niños pudieran competir entre sí al jugar. Sin embargo, no se tuvo en cuenta que había niños que ya se conocían y llevaban clases juntos, y ellos fueron quienes comenzaron a competir entre sí, respecto a "quién obtiene más puntos". Si bien este comportamiento fue notorio en algunos niños, se podría suponer que la competencia se ve afectada si los niños se conocen. No obstante, se puede decir que los jugadores se sintieron fuertemente conectados con el juego, ya que la media en la dimensión de inmersión sensorial e imaginativa supera la mitad del puntaje (3,93). La media correspondiente a la dimensión de flujo (3,49) indicaría, en parte, que los jugadores perdieron la noción de su propio esfuerzo o del paso del tiempo durante el juego. Por otro lado, no se encontró que el juego sea muy desafiante, pues la media en esta dimensión se encuentra en 2,42, lo cual sugiere que se aumente la dificultad o se la regule. Finalmente, se puede considerar que los participantes sintieron muy poca tensión al momento de probar el prototipo, debido a que su media se encuentra en 1,48; y su percepción no fue negativa (media de 1,67).

En lo que respecta al módulo *in-game*, las dimensiones con mayor puntaje fueron el afecto positivo (media de 4,1), seguido de la inmersión con un puntaje de 3,86, el flujo con 3,81, la competencia con 3,69 y el desafío con 2,86. Estos resultados, al pasar la media (2,5), indican que, mientras el participante jugaba, se sentía conectado con el juego, perdía la noción del tiempo al jugar y que juzgó bien su desempeño; y, finalmente, que los participantes encontraron el juego medianamente desafiante. Por otro lado, las dimensiones que no pasaron la media fueron el afecto negativo con 1,62 y la tensión con 1,45, lo cual es aceptable, debido a que estas dimensiones están relacionadas con una mala experiencia durante el juego y se buscaba el menor puntaje en ambas.

Finalmente, respecto a la efectividad del juego serio para el reforzamiento de matemática, si se analizan únicamente los promedios de los puntajes de los participantes de cada grupo, claramente se evidencia que el grupo experimental tuvo mejor rendimiento que el grupo de control, pues el primero obtuvo 16 puntos y el segundo, 11,6 (4,4 puntos de diferencia). Además, el mayor puntaje del primer grupo es superior al mayor puntaje del segundo (18 y 14, respectivamente); y el menor puntaje del primer grupo sigue siendo superior al menor obtenido en el segundo (14 y 10, respectivamente).

Analizando únicamente estos valores, podría determinarse que el juego serio educativo para el refuerzo de las habilidades aritméticas de fracciones y divisiones

es efectivo y, por lo tanto, solventaría las deficiencias del modelo de refuerzo tradicional. No obstante, la desviación estándar del grupo experimental es levemente menor a la obtenida en el grupo de control, pues estos alcanzaron puntajes de 1,63 y 1,67 respectivamente (una diferencia de 0,04 puntos).

Esto explica por qué la prueba de hipótesis rechaza la hipótesis nula y toma como válida la hipótesis alternativa, dando a entender que los niños que reforzaron sus habilidades aritméticas de fracciones y divisiones tuvieron un mayor rendimiento respecto a aquellos que reforzaron sus habilidades de forma tradicional. Sin embargo, los resultados de esta prueba no son significativos y no determinan la efectividad o no del sistema, ya que se tiene un número bastante reducido de participantes y las muestras no son homogéneas (rango amplio de edades); por tanto, los resultados de esta prueba no son confiables.

7. CONCLUSIONES

Se presentó un juego serio educativo que apoye el proceso de reforzamiento de habilidades aritméticas de fracciones en niños de primaria. Se buscó que la metodología contemple la obtención de los recursos pedagógicos, el diseño, la implementación del juego serio y una etapa experimental, en la cual se contó con dos grupos (de control y experimental) para corroborar la efectividad del sistema. Los participantes fueron sometidos a una evaluación posterior con la finalidad de obtener un puntaje (de 0 a 20) referente a su desempeño en el tema de fracciones y divisiones. Se recolectó el puntaje de 9 niños entre los 8 y 12 años.

Durante el desarrollo se implementó, en su totalidad, el juego serio y el diseño de una mecánica (gracias al *framework* MDA) para el refuerzo de las habilidades aritméticas de fracciones y divisiones. Asimismo, se pudo validar la experiencia de juego en dos momentos (general y durante el juego) a través del cuestionario de experiencia de juego, el cual indica que los participantes tuvieron una experiencia de juego buena y que la competencia y el afecto positivo son las dimensiones que gobiernan, seguidas de la poca tensión y el afecto negativo.

Adicionalmente, se corroboró la usabilidad del sistema, dado que obtuvo una puntuación promedio de 68,33/100 puntos, según el cuestionario SUS, lo cual indica el correcto uso de las heurísticas de usabilidad, así como la adecuada implementación de los elementos de juego. No obstante, debido al limitado grupo de participantes en la experimentación, no se puede afirmar del todo que el uso del juego serio haya reforzado las habilidades aritméticas de fracciones y divisiones en niños respecto a la metodología tradicional. Sin embargo, los resultados preliminares muestran la posible predisposición a que el juego serio educativo desarrollado mejora el proceso

de refuerzo. Finalmente, para poder garantizar que el juego serio desarrollado mejora el proceso de refuerzo de las habilidades aritméticas de fracciones y divisiones, en trabajos futuros se llevará a cabo una experimentación similar, pero con un grupo homogéneo y con un número mayor de participantes, divididos por igual en dos grupos.

REFERENCIAS

- Aljojo, N. (2018). The design and implementation of a mathematics game-base learning application for primary students. *International Journal of Interactive Mobile Technologies*, 12(3), 142-152. <https://doi.org/10.3991/ijim.v12i3.8739>
- Alkhateeb, M. A. (2019). Effect of mobile gaming on mathematical achievement among 4th graders. *International Journal of Emerging Technologies in Learning*, 14(7), 4-7. <https://doi.org/10.3991/ijet.v14i07.10315>
- Aparício, A., & Silva, F. (2019). Arithmetic Bird: a game for training mathematical operations. *EAI Endorsed Transactions on Game-Based Learning*, 5(17), 1-9. <https://doi.org/10.4108/eai.11-7-2019.159526>
- Bailey, D. H., Hoard, M. K., Nugent, L., & Geary, D. C. (2012). Competence with fractions predicts gains in mathematics achievement. *Journal of Experimental Child Psychology*, 113(3), 447-455. <https://doi.org/10.1016/j.jecp.2012.06.004>
- Benbow, C. P., & Faulkner, L. R. (2008). Rejoinder to the critiques of the national mathematics advisory panel final report. *Educational Researcher*, 37(9), 645-648. <https://doi.org/10.3102/0013189x08329195>
- Bicen, H., & Kocakoyun, S. (2018). Perceptions of students for gamification approach: Kahoot as a case study. *International Journal of Emerging Technologies in Learning*, 13(2), 72-93. <https://doi.org/10.3991/ijet.v13i02.7467>
- Carrol, J. (2013, 3 de septiembre). Applying the MDA game design framework to software. *Atomic Spin*. <https://spin.atomicobject.com/2013/09/03/mda-game-design-framework/>
- El Azizi, L. (2019). Design and development of a serious mobile game "MathAdventure". *International Journal of Advanced Research in Computer and Communication Engineering*, 8(9), 16-21. <https://doi.org/10.17148/ijarccce.2019.8903>
- Gaggi, O., Ciraulo, F., & Casagrande, M. (2018). Eating pizza to learn fractions. En *ACM International Conference Proceeding Series* (pp. 220-225). <https://doi.org/10.1145/3284869.3284921>

- Hunicke, R., Leblanc, M., & Zubek, R. (2004). *MDA: A formal approach to game design and game research*.
- Ibarra, M. J., Jiménez, W., Soto, C., Chavez, E., Chiclla, E., Sprock, A. S., & Brandao, L. O. (2019). Game based learning for math learning: iFractions case study. En *Proceedings - 2019 International Conference on Virtual Reality and Visualization (ICVRV)* (pp. 208-211). <https://doi.org/10.1109/ICVRV47840.2019.00050>
- Ibarra, M. J., Soto, W., Ataucusi, P., & Ataucusi, E. (2016). MathFraction: Educational serious game for students motivation for math learning. En *Proceedings - 2016 11th Latin American Conference on Learning Objects and Technology (LACLO)* (pp. 1-9). <https://doi.org/10.1109/LACLO.2016.7751777>
- IJsselsteijn, W. A., De Kort, Y. A. W., & Poels, K. (2008). *The Game Experience Questionnaire*. Technische Universiteit Eindhoven.
- Johnson, D., Gardner, M. J., & Perry, R. (2018). Validation of two game experience scales: The Player Experience of Need Satisfaction (PENS) and Game Experience Questionnaire (GEQ). *International Journal of Human Computer Studies*, 118, 38-46. <https://doi.org/10.1016/j.ijhcs.2018.05.003>
- Kebritchi, M., Hirumi, A., & Bai, H. (2010). The effects of modern mathematics computer games on mathematics achievement and class motivation. *Computers & Education*, 55(2), 427-443. <https://doi.org/10.1016/j.compedu.2010.02.007>
- Kickmeier-Rust, M. D., Hillemann, E. C., & Albert, D. (2014). Gamification and smart feedback: Experiences with a primary school level math app. *International Journal of Game-Based Learning*, 4(3), 35-46. <https://doi.org/10.4018/ijgbl.2014070104>
- Lankoski, P., Björk, S., & Stirling, W. C. (2003). Game research methods. En *Satisficing games and decision making: With applications to engineering and computer science* (p. 9). ETC Press.
- Ministerio de Educación del Perú. (2018). *Perú: ¿cómo vamos en educación? Estadística de la Calidad Educativa*. <http://repositorio.minedu.gob.pe/handle/MINEDU/6104>
- Mishra, L. (2020). Conception and misconception in teaching arithmetic at primary level. *Journal of Critical Reviews*, 7(5), 936-939. <https://doi.org/10.31838/jcr.07.05.192>
- Mora-Zamora, R., & Brenes-Villalobos, E. (2019). Integrated framework for game design. En *Proceedings of the IX Latin American Conference on Human Computer Interaction (CLHC '19)* (Article 23, pp. 1-6). <https://doi.org/10.1145/3358961.3358984>

- Muñoz Sanabria, L. F., & Vargas Ordoñez, L. M. (2019). EDUMAT: herramienta web gamificada para la enseñanza de operaciones elementales. *Campus Virtuales*, 8(2), 9-17.
- Murtiyasa, B., Jannah, I. M., & Rejeki, S. (2020). Designing mathematics learning media based on mobile learning for ten graders of vocational high school. *Universal Journal of Educational Research*, 8(11), 5637-5647. <https://doi.org/10.13189/ujer.2020.081168>
- Nah, F. F.-H., Zeng, Q., Telaprolu, V. R., Ayyappa, A. P., & Eschenbrenner, B. (2014). Gamification of education: A review of literature. En F. F.-H. Nah (Ed.), *HCI in Business. HCIB 2014. Lecture notes in computer science* (vol. 8527, pp. 401-409). Springer.
- Naiser, E. A., Wright, W. E., & Capraro, R. M. (2003). Teaching fractions: Strategies used for teaching fractions to middle grades students. *Journal of Research in Childhood Education*, 18(3), 193-198. <https://doi.org/10.1080/02568540409595034>
- Nielsen, J. (1994, 24 de abril). *10 usability heuristics for user interface design*. Nielsen Norman Group. <https://www.nngroup.com/articles/ten-usability-heuristics/>
- Ninaus, M., Kiili, K., McMullen, J., & Moeller, K. (2017). Assessing fraction knowledge by a digital game. *Computers in Human Behavior*, 70, 197-206. <https://doi.org/10.1016/j.chb.2017.01.004>
- Obersteiner, A., Van Dooren, W., Van Hoof, J., & Verschaffel, L. (2013). The natural number bias and magnitude representation in fraction comparison by expert mathematicians. *Learning and Instruction*, 28, 64-72. <https://doi.org/10.1016/j.learninstruc.2013.05.003>
- Olsen, T., Procci, K., & Bowers, C. (2011). Serious games usability testing: How to ensure proper usability, playability, and effectiveness. En A. Marcus (Ed.), *Design, user experience, and usability. Theory, methods, tools and practice. DUXU 2011. Lecture notes in computer science* (vol. 6770, pp. 625-634). https://doi.org/10.1007/978-3-642-21708-1_70
- Procci, K., Chao, A., Bohnsack, J., Olsen, T., & Bowers, C. (2012). Usability in serious games: A model for small development teams. *Computer Technology and Application*, 3(2012), 315-329.
- Robertson, J., & Howells, C. (2008). Computer game design: Opportunities for successful learning. *Computers & Education*, 50(2), 559-578. <https://doi.org/10.1016/j.compedu.2007.09.020>

- Sauro, J., & Lewis, J. R. (2016). Standardized usability questionnaires. En *Quantifying the user experience* (2.^a ed., pp. 185-248). Elsevier. <https://doi.org/10.1016/b978-0-12-802308-2.00008-4>
- Schenke, K., Rutherford, T., & Farkas, G. (2014). Alignment of game design features and state mathematics standards: Do results reflect intentions? *Computers & Education*, 76, 215-224. <https://doi.org/10.1016/j.compedu.2014.03.019>
- Shang, J., Zhang, L., Hu, R. N., Ma, S. J., Zen, J. L., Yuan, M. Z., & Sun, J. G. (2018). How to support fraction learning with math game "Run Fraction": Theory, design and application. En *Proceedings of International Conference on Computers in Education (Extended Summary Proceedings)* (pp. 20-22). The Ateneo de Manila University.
- Spieler, B., & Girvan, C. (2020). *Das PECC-Framework: Gender-Sensibilität und spielerische Programmierung in der informatischen Grundbildung*. En R. Zender, D. Ifenthaler, T. Leonhardt & C. Schumacher (Eds.), *DELFI 2020 – Die 18. Fachtagung Bildungstechnologien der Gesellschaft für Informatik e. V.* (pp. 247-258). Gesellschaft für Informatik e. V.
- Spieler, B., & Kemeny, F. (2020). Design, complexity, and coding: A framework to evaluate games. En *Proceedings of the 14th International Conference on Game Based Learning (ECGBL)* (pp. 558-566). <https://doi.org/10.34190/GBL.20.156>
- UX Research. (2017). *Measuring and interpreting System Usability Scale (SUS) - UIUX Trend*. UIUXTrend. <https://uiuxtrend.com/measuring-system-usability-scale-sus/>
- Vandercruysse, S., Maertens, M., & Elen, J. (2015). Description of the educational math game "Monkey Tales: The Museum of Anything". En J. Torbeyns, E. Lehtinen & J. Elen (Eds.), *Describing and studying domain-specific serious games* (pp. 24-43). <https://doi.org/10.1007/978-3-319-20276-1>
- Wang, S. Y., Chang, S. C., Hwang, G. J., & Chen, P. Y. (2017). A microworld-based role-playing game development approach to engaging students in interactive, enjoyable, and effective mathematics learning. *Interactive Learning Environments*, 26(3), 411-423. <https://doi.org/10.1080/10494820.2017.1337038>
- Zhang, L., Shang, J., Pelton, T., & Pelton, L. F. (2020). Supporting primary students' learning of fraction conceptual knowledge through digital games. *Journal of Computer Assisted Learning*, 36(4), 540-548. <https://doi.org/10.1111/jcal.12422>

SISTEMA DE MONITOREO DE AMBIENTE PARA EL CONTROL DE LAS CONDICIONES AMBIENTALES DE LAS GRANJAS ACUÍCOLAS DE LA SELVA DEL PERÚ BASADO EN INTERNET DE LAS COSAS*

ADRIAN ERNESTO LINO VILLARAN
20170837@aloe.ulima.edu.pe
<https://orcid.org/0000-0003-1421-6921>
Universidad de Lima, Perú

RESUMEN. La industria alimenticia de la acuicultura está creciendo a un ritmo tan acelerado que está por alcanzar a la pesca en toneladas producidas. Sin embargo, en el Perú, la acuicultura se encuentra poco desarrollada: su participación en el sector es del 1,4 %, y cuenta con productores poco especializados, que residen en zonas alejadas con baja posibilidad de acceso tecnológico. Por ello, en este trabajo se plantea un sistema de monitoreo de ambiente semiautónomo dirigido exclusivamente a la zona oriental del Perú, capaz de medir temperatura (agua y aire), oxígeno disuelto, pH y luminosidad. Con este fin se genera un nodo de medición de parámetros que captura las condiciones ambientales y las envía a un *gateway* a través de la red LoRa. En un despliegue de casi cinco días, con más de 69 000 datos capturados, el sistema demostró que tiene la capacidad de medir las condiciones ambientales de forma precisa y autónoma con tendencias normales para la ubicación y hora de medición. Asimismo, se pudieron identificar ciertas correlaciones entre los datos obtenidos, como la temperatura del ambiente sobre el oxígeno disuelto, que tiene un coeficiente de correlación de 0,81. Por otro lado, el uso de la red LoRa requiere aplicar cierto tipo de control de la integridad de la información, ya que esta no viene por defecto. Finalmente, se concluye que el sistema desarrollado sí permite monitorear las condiciones ambientales en la selva peruana, brindando al productor la capacidad de visualizar su información en tiempo real de forma local y remota.

PALABRAS CLAVE: internet de las cosas, red inalámbrica de sensores, sensores

* Quiero agradecer a mi asesor Pablo Alberto Rojas Jaén, quien me guio en el desarrollo; y a la asociación de productores piscícolas de Campoverde, que puso a mi disposición sus instalaciones para la experimentación.

ENVIRONMENTAL MONITORING SYSTEM FOR THE CONTROL OF ENVIRONMENTAL CONDITIONS OF AQUACULTURE FARMS IN THE PERUVIAN JUNGLE BASED ON THE INTERNET OF THINGS

ABSTRACT. The aquaculture food industry is growing at a fast pace and is about to catch up with fishing in terms of tons produced. However, in Peru aquaculture only represents 1,4 % of the fish industry, with producers who are not very specialized and live in remote areas with lack of technology. For this reason, a semi-autonomous environment monitoring system is being developed exclusively for the eastern zone of Peru, capable of measuring temperature (water and air), dissolved oxygen, pH, and luminosity in real time. For its development, a Parameter Measurement Node is proposed which captures the conditions of the environment and send them to a gateway through the LoRa network. During a 5 days period, more than sixty-nine thousand data points were captured, the system demonstrated that it could measure environmental conditions accurately and autonomously with expected trends for the location and time of measurement. Likewise, it was possible to identify certain correlations between the data, such as ambient temperature and dissolved oxygen, which has a correlation coefficient of 0,81. On the other hand, the use of the LoRa network requires the application of a certain type of information integrity control since it is not provided by default. Finally, it was concluded that the system developed does allow monitoring environmental conditions in the Peruvian rainforest, providing the producer with the ability to visualize his information in real time locally and remotely.

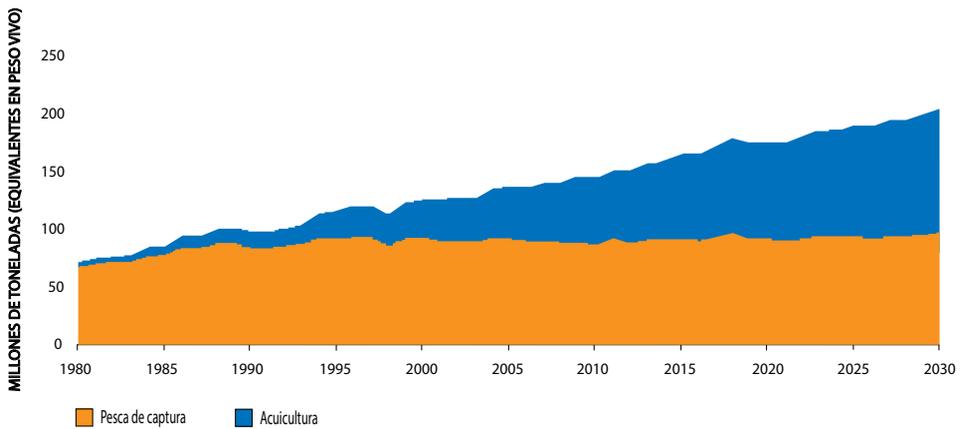
KEYWORDS: Internet of Things, wireless sensor network, sensors

1. INTRODUCCIÓN

Actualmente, la acuicultura es la industria alimenticia con el crecimiento más acelerado: a nivel mundial, ha crecido un 8,9 % desde los años setenta, lo cual demuestra su enorme potencial (Subasinghe, 2005). Asimismo, esta tendencia continúa, como se puede apreciar en la Figura 1 (Organización de las Naciones Unidas para la Alimentación y la Agricultura [FAO], 2020, p. 193).

Figura 1

Producción mundial de pescado



Nota. Reproducido de *El estado mundial de la pesca y la acuicultura* (p. 179), por FAO, 2020.

Por otro lado, a nivel nacional, la producción de pescado representa el 0,30 % del PBI (Banco Central de Reserva del Perú [BCRP], 2020); este sector sufrió una variación de únicamente $-15,8\%$ entre el segundo trimestre del 2019 y el del 2020, por lo que se posiciona como el segundo sector menos afectado en este periodo (Instituto Nacional de Estadística e Informática [INEI], 2020). Asimismo, la FAO (2020) indica que se produjo 7,2 millones de toneladas de pescado, de las cuales 104 000 toneladas provienen de la acuicultura, que tiene una participación del 1,4 % en el sector. Sin embargo, a nivel mundial, la participación acuícola es significativamente mayor: representa el 45,9 % de la producción total de pescado, lo que indica un bajo desarrollo del sector acuícola en el Perú en comparación con la producción global (FAO, 2020, p. 180).

En línea con lo mencionado, su poca representación (INEI, 2020) coloca a los productores pertenecientes a la acuicultura de recursos limitados (47,9 % del sector acuícola) y a la acuicultura de micro y pequeña empresa (50,7 % del sector acuícola)

del país en una difícil posición para acceder a tecnología con el objetivo de hacer crecer su negocio y mejorar sus prácticas de producción, ya que no se encuentran en la capacidad de innovar tecnológicamente debido a su informalidad, inestabilidad financiera y carencia de infraestructura de sus zonas (Ministerio de la Producción, 2018, p. 16).

A la fecha, a nivel internacional, se han llevado a cabo investigaciones sobre la implementación de sistemas basados en internet de las cosas (IoT, por sus siglas en inglés); sin embargo, son muy pocos los trabajos de este tipo para un ambiente de producción en el territorio peruano. Así, no se toma en consideración las características de conectividad, climáticas y geográficas de la zona oriental del país; en cambio, las investigaciones internacionales fueron realizadas bajo condiciones controladas de laboratorio, en donde se tiene disponibilidad de corriente eléctrica y conexión constante a internet (Encinas et al., 2017; Kang et al., 2017; Salim et al., 2016).

Debido a lo mencionado anteriormente, la presente investigación busca enfocarse en los pequeños y microacuicultores de la parte oriental del Perú, y desarrollar un sistema que permita monitorear el ambiente de producción. Asimismo, se busca considerar las características geográficas y climáticas, así como la disponibilidad de conexión a internet para el desarrollo de la investigación.

2. ESTADO DEL ARTE

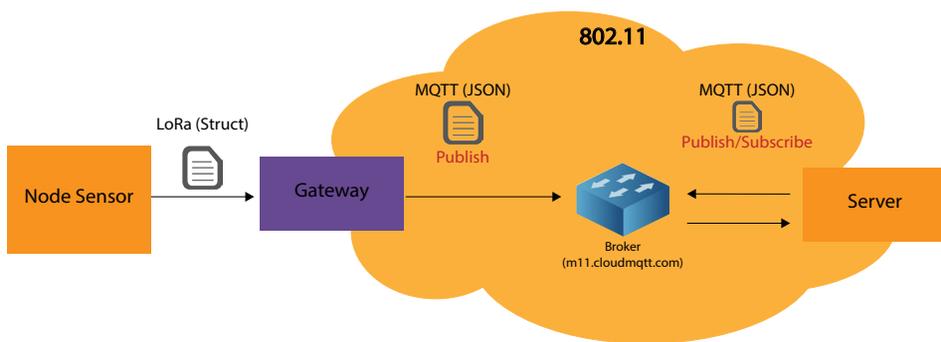
En el territorio nacional existen escasas investigaciones o proyectos sobre sistemas de monitoreo de ambiente orientados a la acuicultura. Una de ellas es la de Ríos Julcapoma y Yauri Rodríguez (2017), quienes realizaron la implementación de un nodo sensor que fue desplegado en la ciudad de Iquitos. Este dispositivo contaba con sensores de temperatura, oxígeno disuelto y pH, los cuales se encontraban conectados a un microcontrolador PIC24 que, a través de la red ZigBee, transmitía los datos recopilados. Este nodo fue construido sobre un armazón de tubos de PVC para su flotabilidad; además, se colocó un panel solar en la parte lateral para la alimentación de energía.

Zhu et al. (2010) y Shi et al. (2018) desarrollaron sistemas bastante similares utilizando una red inalámbrica de sensores (WSN, por sus siglas en inglés) y un *gateway*. Pero se diferencian en que Zhu et al. (2010) usan la red CDMA, mientras que Shi et al. (2018) la red GPRS (véase la Figura 2). Ambos proyectos buscaron no ser dependientes de una red alámbrica de internet para que puedan ser usados en las locaciones más lejanas. Shi et al. (2018) concluyen en su estudio que el uso de una topología de árbol genera una menor pérdida de paquetes en el envío de información,

en comparación con los patrones estrella y malla, así como en la observación se pudo notar un menor uso de energía. Por otro lado, Zhu et al. (2010) lograron implementar redes neuronales de retropropagación que sirvieron para pronosticar de forma correcta y estable el oxígeno disuelto (OD) en el sistema, culminando con un 92,5 % de consistencia y la posibilidad de brindar alertas sobre posibles caídas futuras en los parámetros medidos.

Figura 2

Arquitectura



Nota. Reproducido de "A Wireless Sensor Network-based Monitoring System for Freshwater Fishpond Aquaculture", por B. Shi, V. Sreeram, D. Zhao, S. Duana y J. Jiang, 2018, *Biosystems Engineering*, 172.

Yim et al. (2018), en su investigación, ponen a prueba la respuesta a la interferencia de la red LoRa en un bosque de árboles ubicado en el estado de Indiana, Estados Unidos. Los autores ubicaron un emisor en un extremo del bosque, e hicieron pruebas de recepción de paquetes con un tamaño de 9 bytes a 100 m, 150 m y 200 m de distancia, variando el *spreading factor* (SF), el ancho de banda (BW) y el *coding rate* (CR) de la transmisión. Se realizaron 12 combinaciones de los parámetros anteriormente mencionados y se logró concluir que se debe mantener el SF y el CR lo más alto posible para una mayor fiabilidad en la transmisión.

Encinas et al. (2017) señalan que en la acuicultura es esencial responder rápidamente a las variables cambiantes del entorno para el correcto desarrollo de los peces. Sin embargo, las mediciones suelen ser manuales, a cargo de personal que va directamente a medir la calidad del agua. El problema es una respuesta tardía a las variables del sistema; así, cuando se obtiene un valor alarmante, por lo general es demasiado tarde. Esto se ve reflejado en el estudio de Abbink et al. (2012), quienes alteraron distintos parámetros en un sistema acuícola de recirculación de agua para la especie *Seriola lalandi*. En esta investigación, se realizaron dos experimentos: el

primero analiza el efecto de la variación de la temperatura en los peces; y el segundo, el efecto del pH en la vida de los peces. Con estos experimentos se pudo concluir que la temperatura ideal para la especie es de 26,5 °C, ya que genera un crecimiento del 30 % del peso final luego de 30 días, en comparación con 21 °C. Asimismo, se dedujo que una baja en el pH estresa a los peces y los lleva a la muerte, por lo que su vida tiene una alta dependencia de este parámetro.

El estudio de Grados y Bedon (2020) busca plantear y documentar la arquitectura de un sistema de monitoreo de cultivos de quinua y de un sistema de predicción de la ubicación de peces en el mar. Dicha arquitectura comprende *hardware*, *software*, tecnología y servicios de *cloud*. En el *hardware*, el *gateway* utilizó un Raspberry Pi con el sistema operativo Raspbian, el cual se conectaba mediante XBee a las estaciones. El servicio *cloud* que se utilizó provino de Google Cloud Platform, en donde se realizaba el procesamiento y almacenamiento de la data. A través de este servicio en nube, la aplicación del usuario recupera la data y permite actuar sobre el sistema.

3. ANTECEDENTES

Existe un déficit tecnológico en la acuicultura a nivel nacional, el cual se encuentra enfocado en los productores de menor escala; por eso, es de suma urgencia introducir la innovación tecnológica a las zonas más alejadas del país (Ministerio de la Producción, 2018). Esto requiere que ciencia, tecnología e innovación converjan entre sí para que, en conjunto, se pueda diseñar un sistema con la mayor calidad posible.

3.1 Acuicultura

La acuicultura es la actividad de cultivar especies hidrobiológicas (peces, moluscos, crustáceos y plantas acuáticas) en lugares controlados (FAO, 1988). Esta actividad resulta bastante compleja, pues, en todos los casos, es necesario medir constantemente las condiciones del ambiente para lograr el mejor desarrollo de los animales y plantas. Asimismo, cada uno de los parámetros del ambiente afecta de cierta manera al animal o planta, generando efectos sobre ellos. Por ello, es necesario establecer la importancia de cada uno de estos parámetros para posteriormente determinar un método de medición y control de estos. Uno de los parámetros más importantes es el oxígeno disuelto (DO), ya que su disminución resulta fatal para la mayoría de los animales y bacterias marinas (Kramer, 1987). Le siguen otros como pH, temperatura, sal, nitratos, ácido sulfhídrico, carbonatos, bicarbonatos y amoníaco (Campbell, 1973; Manahan, 2000; Raju & Varma, 2017).

3.2 Internet de las cosas

Dorsemaine et al. (2015) definen internet de las cosas (IoT) como un grupo de infraestructuras basadas en sensores y actuadores que al comunicarse entre sí transportan, almacenan y procesan información. De la misma manera, al diseñar un sistema de IoT, se debe analizar ciertos pilares, que se detallan a continuación.

3.2.1 Fuente de energía

Bormann y Keranen (2014) presentan cuatro tipos de adquisición de energía: energía del ambiente, energía recargable, energía primaria no recargable y siempre conectado. De estas, las tres primeras requieren un manejo del consumo, ya que son limitadas y pueden afectar el tiempo de vida del dispositivo, así como su capacidad de procesamiento. En el caso de los dispositivos ubicados en el exterior, en zonas alejadas y con baja disponibilidad de corriente eléctrica, los paneles solares con baterías son bastante utilizados. Debido a esto, es de suma importancia calcular el requerimiento energético del dispositivo, con el fin de contar con una batería que permita alimentarlo durante las horas necesarias y con un panel solar para alimentar el dispositivo y cargar la batería instalada en un tiempo prudente. Para pequeños dispositivos, se recomienda el uso de un microcontrolador como un Arduino Nano, que tiene un consumo de 19 mA (Arduino, s. f.; Dorsemaine et al., 2015).

3.2.2 Redes de comunicación

Es importante escoger el tipo de red de comunicación, ya que esta puede variar el consumo energético mencionado en el punto anterior. Asimismo, esta elección depende de la funcionalidad que se le da al sistema, si se va a colocar uno o muchos sensores, o el tipo de arquitectura que se va a utilizar. Jawad et al. (2017) presentan las tecnologías LoRa, ZigBee, Bluetooth, wifi y GPRS. Entre estas tecnologías, la red LoRa es la que mejor se ajusta al presente proyecto debido a su amplio alcance (15 km aproximadamente), bajo consumo y flexibilidad.

3.2.3 Sensores y actuadores

Un sensor es un dispositivo que detecta eventos o cambios en el ambiente y provee una señal de salida (Rayes & Salam, 2016). Existe una amplia gama de sensores; Sehrawat y Gill (2019) los dividen en 13 tipos, entre los cuales se encuentran los sensores de proximidad, posicionamiento, presencia, movimiento, velocidad, temperatura, presión, químicos, humedad, calidad de agua, infrarrojo, giroscopio y ópticos.

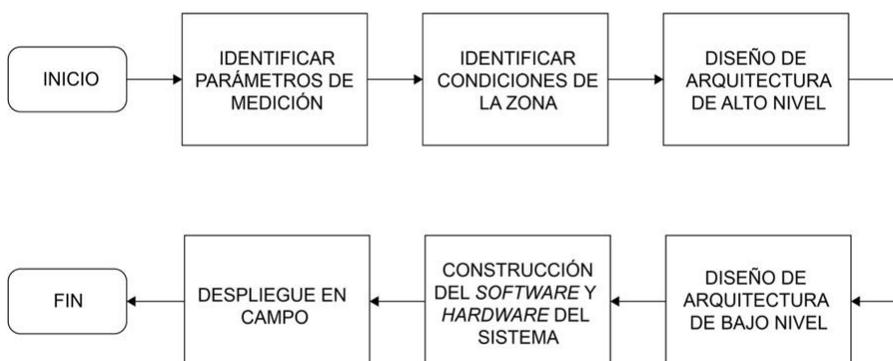
Un actuador es un dispositivo que toma acción sobre el sistema, afectando las condiciones de este (Rayas & Salam, 2016). Estos normalmente toman acción por los resultados que brinda un sensor, regulando las condiciones del sistema.

4. METODOLOGÍA Y EXPERIMENTACIÓN

El objetivo de la presente investigación es desarrollar el *software* y el *hardware* de un sistema de monitoreo y control de ambiente que permita medir las condiciones ambientales de las granjas acuícolas de la zona oriental del Perú. Para ello, en la Figura 3 se detalla la metodología.

Figura 3

Metodología



4.1 Identificación de parámetros de medición

Como se mencionó anteriormente, existen ciertos parámetros que deben mantenerse en rangos óptimos para el correcto desempeño del animal. Sin embargo, algunos de ellos son más importantes que otros, por lo que es esencial clasificarlos. A continuación, se describe en detalle cada uno de los parámetros priorizados.

- *Oxígeno disuelto*. Es uno de los parámetros más importantes en una producción acuícola, pues una baja significativa del oxígeno disuelto puede llevar a los peces a sofocarse y, finalmente, a la muerte (Kramer, 1987). Por otro lado, variaciones no letales pueden ocasionar una mala alimentación y una tasa de conversión alimenticia deficiente (Fondo Nacional de Desarrollo Pesquero, 2004).
- *Temperatura*. Afecta el metabolismo del animal; a mayor temperatura, mayor consumo de oxígeno y alimentación (Fondo Nacional de Desarrollo

Pesquero, 2004). Asimismo, la temperatura del aire afecta directamente a la temperatura del agua.

- *pH*. Una alteración en el pH del agua que sobrepase las condiciones normales de cultivo puede llevar a los peces al estrés y, finalmente, a la muerte (Abbink et al., 2012). Es recomendable medirlo de forma constante, pero no se considera esencial, debido a su poca volatilidad.
- *Luminosidad*. Afecta directamente a la producción de oxígeno natural en el agua a causa de fotosíntesis, por lo que se puede utilizar para relacionarlo con la producción de oxígeno.

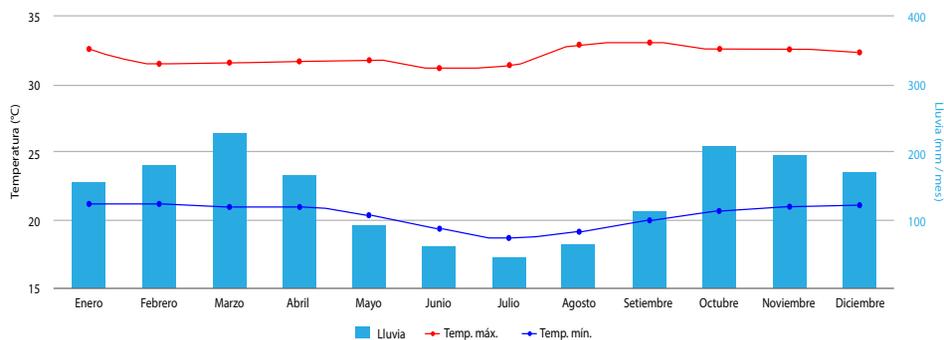
4.2 Identificación de la zona

Para la identificación de las condiciones ambientales de la zona oriental del país, se realizó un viaje a la ciudad de Pucallpa del 30 de julio al 8 de agosto del 2021, en el que se visitaron diez productoras piscícolas ubicadas en el distrito de Campoverde. En estas inspecciones, se encontró que, debido a que el terreno es accidentado y con amplia vegetación, los productores limpian sus zonas y utilizan maquinaria para nivelarlas. A pesar de ello, una pequeña cantidad de productoras piscícolas presentan variaciones en altitud, por lo que construyen sus estanques de forma escalonada con el objetivo de aprovechar la gravedad para transferir el agua entre sus estanques. Los estanques observados cuentan con una profundidad de entre 1,60 m y 1,80 m, con medidas de 20 m × 10 m hasta los 50 m × 30 m, aproximadamente (véase la Figura 4). Por otro lado, en la ciudad de Pucallpa, se producen temperaturas elevadas a lo largo del año, que oscilan entre 19 °C y 33 °C (véase la Figura 5), según el Servicio Nacional de Meteorología e Hidrología del Perú (Senamhi, 2021).

Figura 4

Piscinas construidas de forma escalonada



Figura 5*Clima promedio en Pucallpa*

Nota. Reproducido de *Promedio de temperatura normal para Pucallpa*, por Senamhi, 2021 (<https://www.senamhi.gob.pe/?p=pronostico-detalle-turistico&localidad=0024>).

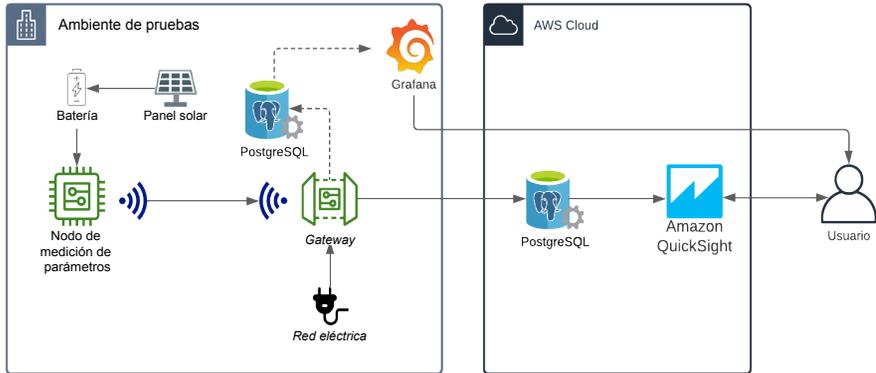
El 70 % de las productoras piscícolas visitadas en el distrito de Campoverde no cuentan con una conexión de red eléctrica constante, por lo que ciertos productores optan por el uso de paneles solares o de generadores a gasolina para manejar sus equipos. Además, la señal de celular en Campoverde no está implementada a lo largo de todo el distrito, por lo que la señal más estable se encuentra alrededor de la carretera. Debido a ello, algunos productores usan internet satelital, que es más costoso y no permite que puedan desplazarse a lo largo de todo su terreno con conectividad constante en sus dispositivos móviles.

4.3 Diseño de arquitectura de alto nivel

Para el diseño de la arquitectura, se planteó el uso de dos dispositivos físicos: un nodo de medición de parámetros y un *gateway*. En el caso del nodo de medición de parámetros, al ser un dispositivo que será ubicado dentro de una piscina, alejado de cualquier tipo de conexión eléctrica, se propone utilizar un panel solar acompañado de una batería que permita su funcionamiento durante las noches. Por otro lado, para el *gateway*, que es un dispositivo que se encontrará colocado en una zona administrativa, se considerará la disponibilidad de red eléctrica. Asimismo, el *gateway* se conectará a internet a través de una señal wifi o cableada, en caso de que se encuentre disponible, la cual puede provenir de distintas fuentes. Para el almacenamiento de los datos sin conexión, se implementará una base de datos y se utilizará el *software* Grafana para visualizar los datos *gateway*, ambos de forma local. Cabe mencionar que se mantendrá la misma estructura de Amazon Web Services (AWS) para la visualización de datos de forma remota. Dichos cambios se pueden apreciar en la Figura 6.

Figura 6

Arquitectura de alto nivel



4.4 Diseño de arquitectura de bajo nivel

4.4.1 Nodo de medición de parámetros

- *Componentes de hardware*

En la construcción del nodo de medición de parámetros, se utiliza como componente principal un Arduino Nano conectado a un módulo LoRa SX1268. Este módulo cumple las funciones de transmisor y receptor mediante la conexión SPI al Arduino Nano. Este módulo transmite a través de la red LoRa a 433 MHz con una potencia máxima de transmisión de 18 ± 1 dBm paquetes de hasta 256 bytes (Ai-Thinker, 2017). Asimismo, el nodo contará con los sensores descritos en la Tabla 1.

Tabla 1

Sensores

Parámetros	Sensor
Oxígeno disuelto	Sensor digital Atlas Scientific
Temperatura del aire	DS18B20
Temperatura del agua	DS18B20
pH	PH-4502C DIY-MORE
Luminosidad	BH1750

Debido a las condiciones extremas de la zona, se decidió utilizar como fuente de energía un panel solar sobredimensionado, ya que en los días nublados o con poca energía solar este sería capaz de compensarlo. En la Tabla 2 se muestra el detalle

de la fuente de energía y en la Figura 7 se observa el diagrama de circuitos de todo el sistema.

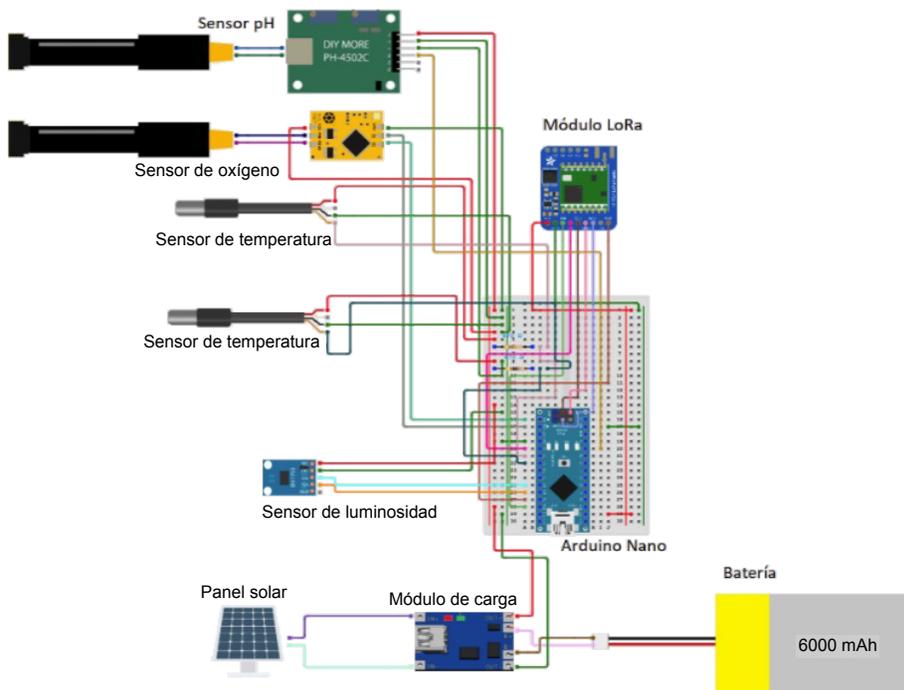
Tabla 2

Fuente de energía

Concepto	Valor
Voltaje máximo de salida	18,36 V
Amperaje máximo de salida	1,63 A
Amperaje mínimo requerido	140 mA
Batería del sistema	6000 mAh

Figura 7

Diagrama de circuitos del nodo de medición de parámetros



En línea con lo mencionado, se utilizaron dos cajas de derivación de 20 cm × 15,5 cm × 8 cm y de 10 cm × 10 cm × 7 cm para el dispositivo y componentes de alimentación de energía, implementando prensaestopas en las conexiones

salientes, de tal forma que el agua no pueda ingresar dentro del dispositivo (véase la Figura 8). Asimismo, se decidió construir una estructura flotante que soporte todos los componentes del nodo de medición de parámetros. De esta forma, los sensores pueden ser ubicados en el centro de la poza y mantienen una profundidad constante en el agua, aunque el nivel de esta cambie. La estructura fue construida con tubos de PVC de 4 pulgadas sellados con pegamento para evitar el ingreso del agua. Asimismo, en el interior de los tubos, se insertaron botellas de plástico para que, en caso de ingreso de agua, esta no pueda hundir al dispositivo. De igual forma, se colocó un tubo de 1 pulgada que soportaba los componentes y el panel solar se instaló en la parte superior. Esta estructura se puede observar en las Figuras 9 y 10.

Figura 8

Nodo de medición de parámetros



Figura 9

Estructura flotante



Figura 10

Estructura flotante (lateral)



- *Software*

El *software* desarrollado cumple la función de interactuar con los sensores conectados al nodo de medición de parámetros. Para ello, se utilizaron las librerías DallasTemperature, OneWire, BH1750, ArduinoJson y RH_RF95, las cuales permitieron al Arduino Nano interactuar con los módulos propuestos (Blanchon, 2021; Burton, 2021; Laws, 2021; Mistry, 2021; Studt et al., 2019). Asimismo, debido a la posibilidad de que los paquetes lleguen incompletos o con data errónea, el código desarrollado genera un *checksum* basado en la codificación de cada carácter del texto enviado, el cual se concatena con el paquete enviado, con el fin de que el *gateway*, al recibir la información, pueda hacer la verificación correspondiente.

Para el desarrollo del código se tomó como referencia el artículo de Raj (2019), publicado en el sitio web CircuitDigest. El código puede ser descargado en el siguiente repositorio: https://github.com/adrianlinov/nmp_inv.

4.4.2 Gateway

- *Hardware*

Para el *gateway* se utilizó un Raspberry Pi 4 Modelo B. Este dispositivo cuenta con un procesador Broadcom BCM2711 de 8 núcleos de 64 bits, 8 GB de RAM DDR4, puerto Ethernet, 4 puertos USB, pines GPIO, así como wifi y Bluetooth integrado. Esto lo vuelve un dispositivo muy capaz para desarrollar el proyecto (The Raspberry Pi Foundation, 2021).

El Raspberry Pi 4 Modelo B, por contar con pines GPIO, es factible de conectarse al módulo SX1268 mencionado anteriormente; esto le da la capacidad

de recibir los mensajes enviados por el Arduino Mega 2560 de forma inalámbrica a través de la red LoRa. Se puede observar el diagrama de circuitos en la Figura 11 y el *gateway* conectado al módulo SX1268 en la Figura 12.

Figura 11

Diagrama de circuitos gateway

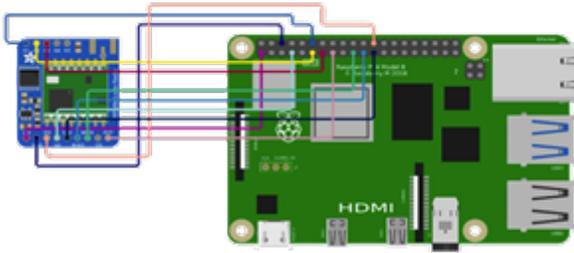


Figura 12

Gateway



- *Software*

Para el desarrollo del *software*, se utilizó un Raspberry Pi 4 con las características de *software* mostradas en la Tabla 3.

Tabla 3*Características del gateway*

Concepto	Descripción
Computador	Raspberry Pi 4 Modelo B
Sistema operativo	Raspbian 10.8
Lenguaje de programación	Python 3.7.3
Base de datos	PostgreSQL

El nodo de medición de parámetros envía al *gateway* información cada 15 segundos, aproximadamente, de forma inalámbrica; esto lo vuelve susceptible a interferencias que ocasionan que la data llegue incompleta o errónea. Debido a ello, se aplicó un método de verificación basado en la suma de caracteres, también llamado *checksum*. Otra funcionalidad desarrollada es la replicación de la base de datos local en la nube (véase la Tabla 4), para la cual se verifica si la fila de información ha sido sincronizada previamente mediante la columna *synced* y, si fuera necesario, se replican las filas nuevas de la base de datos a la nube. Esta funcionalidad depende de una conexión a internet, por lo que se implementó la reiteración constante en caso de que los datos no se hayan subido por una desconexión. Para el desarrollo del código se tomó como referencia el artículo de Raj (2019). El código puede ser descargado del siguiente repositorio: <https://github.com/adrianlinov/gateway-inv>.

Tabla 4*Descripción de tablas*

Tabla	Descripción
Center	Datos del centro de producción
Measurement	Medidas capturadas por el sistema
Pool	Datos de la piscina dentro del centro de producción
Ms_type	Tipo de dato medido por el sistema

Cabe mencionar que, para alojar la base de datos en nube, se usó AWS mediante el servicio RDS en una instancia db.t2.micro con 20 GB de almacenamiento en discos de estado sólido.

5. RESULTADOS

La experimentación se llevó a cabo en una granja piscícola ubicada a 40 km de la ciudad de Pucallpa, en el distrito de Campoverde. El despliegue fue realizado desde

el jueves 21 de octubre del 2021 a las 2:00 p. m. hasta el lunes 25 de octubre del 2021 a las 11:00 a. m. aproximadamente. El dispositivo fue colocado en medio de una piscina con alrededor de 900 peces de la especie amazónica paco, anclado a una estaca en medio de la poza. El despliegue se puede observar en la Figura 13.

Figura 13

Despliegue del nodo de medición de parámetros



Durante el despliegue, el nodo de medición de parámetros envió información al *gateway* de forma constante; sin embargo, sufrió una interrupción por un corte de energía de aproximadamente 27 horas desde la mañana del sábado 23 hasta el domingo 24 de octubre. Este corte de energía hizo que el *gateway* no capture los datos enviados por el nodo de medición de parámetros, por lo que se perdió dicha información. A pesar de ello, el dispositivo registró 169 016 puntos de información, pero en un análisis posterior se pudo notar que se habían almacenado datos duplicados, por lo que se generó un total de 69 257 datos únicos. El detalle se puede observar en la Tabla 5. Además del corte de energía, el lunes 25 de octubre se produjo lluvia y viento, que fueron soportados sin problemas por el nodo de medición de parámetros, ya que este continuó operando y no se observó ingreso de agua dentro de las cajas de los componentes. Los datos obtenidos se encuentran en el siguiente enlace: <https://bit.ly/3L4lGq0>.

Tabla 5*Paquetes recibidos*

Parámetro	Paquetes recibidos con duplicados	Paquetes recibidos sin duplicados
Temperatura del aire	34 191	13 858
Luz del ambiente	34 191	13 858
Oxígeno disuelto	33 524	13 846
pH	33 524	13 846
Temperatura del agua	33 586	13 849

El corte de energía y la falta de internet ocasionaron que el Raspberry Pi perdiera el punto de referencia horaria, almacenando datos con una hora errónea. Por otro lado, si bien se implementó un sistema para sincronizar los datos almacenados de forma local con una base de datos remota alojada en AWS, el dispositivo no pudo ser conectado a internet, por lo que los datos únicamente fueron almacenados de forma local. Esto causó que los *dashboards* de Amazon Quicksight no se encontraran actualizados.

Adicionalmente, se pudo comprobar que una batería de 6000 mAh es suficiente para proveer de energía al dispositivo durante la noche y ser cargada durante el día, manteniendo la continuidad del envío de los datos al *gateway*.

A continuación, se exponen en detalle los datos obtenidos durante el periodo de despliegue. Cabe mencionar que se representará con una línea punteada verde el punto en el tiempo aproximado donde se generó el corte de energía, por lo cual los datos siguientes no están correctamente ubicados en el tiempo.

- *Temperatura del agua y del ambiente*

Figura 14

Mediciones de temperatura del agua y del ambiente

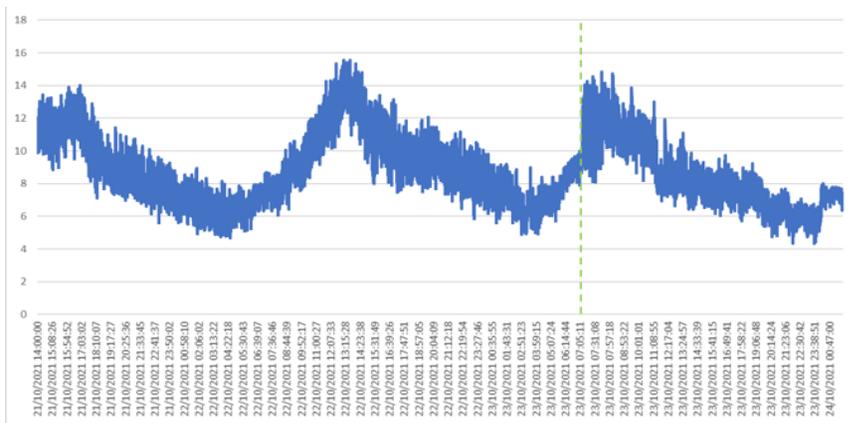


En la Figura 14 se observa que la temperatura del ambiente varía entre los 22 °C y los 36 °C, lo que genera una respuesta con la misma tendencia en la temperatura del agua, la cual varía en menor medida. Asimismo, se puede observar una drástica caída en la temperatura al final del gráfico, que pudo haberse ocasionado por la lluvia del lunes 25 de octubre por la mañana.

- Oxígeno disuelto

Figura 15

Mediciones de oxígeno disuelto

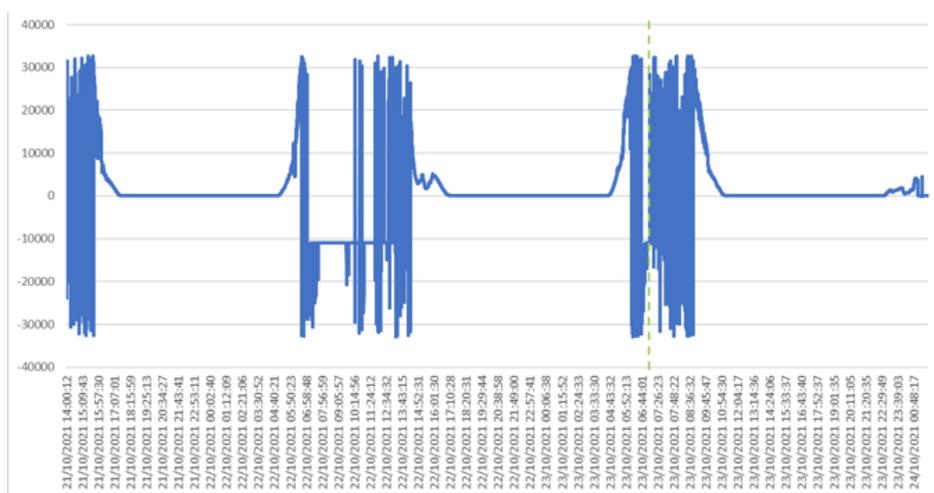


El oxígeno disuelto representado en la Figura 15 mostró el comportamiento esperado: se incrementa durante las horas con más luz y decrece por la tarde y noche, gracias a la fotosíntesis producida por los microorganismos en el agua (Hoganson & Babcock, 1997). Por otro lado, los datos tuvieron una desviación estándar de 2,16, representada con una alta variabilidad en los datos.

- *Luminosidad*

Figura 16

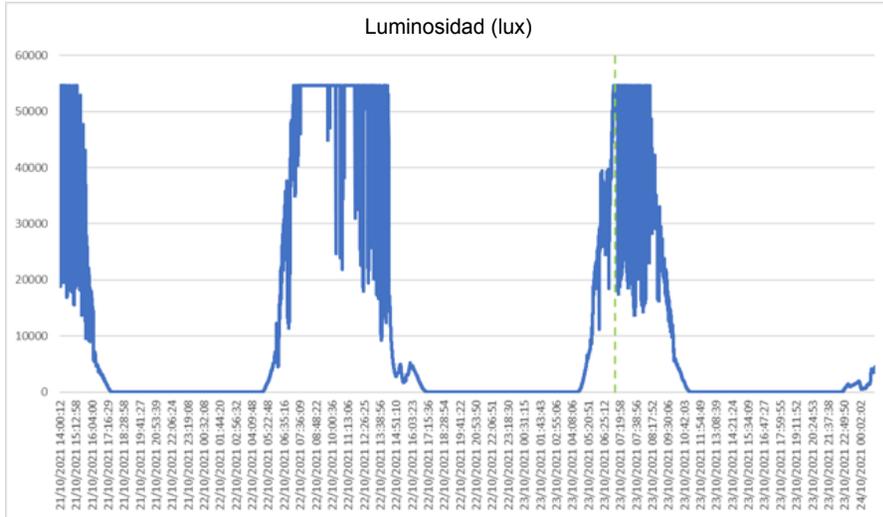
Mediciones de luminosidad



El sensor de luminosidad pudo medir la luz en el ambiente hasta los 32 766 lux; posterior a ello, se registraron datos negativos en la base de datos, los cuales se pueden ver en la Figura 16. A pesar de que el sensor de luminosidad BH1750 puede medir hasta 65 535 lux, los dispositivos Arduino basados en procesadores ATmega almacenan los números enteros o *integers* en un formato de 16 bits, con el bit en posición 16 como el signo que separa positivos y negativos. Esto genera que el rango de valores soportados sea de entre $-32\,768$ a $32\,768$, lo que provoca una pérdida de datos cuando este es superado. Sin embargo, es factible realizar la conversión de los valores almacenados como negativos sumándole 65 536 a dicho valor. Se pueden observar los datos con el rango corregido en la Figura 17.

Figura 17

Mediciones de luminosidad con rango corregido



Por otro lado, los datos presentaron una alta variabilidad en periodos con gran luminosidad, posiblemente debido al efecto de las nubes sobre el dispositivo, aunque no se podría confirmar esta afirmación. Asimismo, si bien el límite superior teórico del sensor es de 65 536 lux, se observa un corte a los 54 612 lux.

- pH

Figura 18

Mediciones de pH

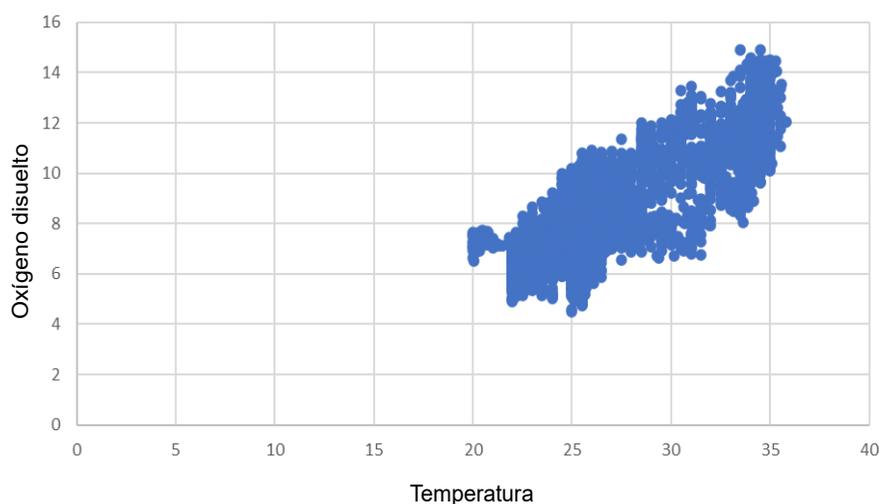


En la Figura 18 se observa que el sensor de pH muestra variaciones durante el primer día de medición. En los días posteriores, se mostraron datos estables hasta el fin del periodo de prueba. Estos datos se encontraron dentro del rango de 5,6 y 7,2 de pH.

Dentro de un ecosistema, es sabido que ciertas condiciones pueden afectar a otras; en este caso, se pudo observar que el oxígeno, la luminosidad y la temperatura aumentan durante el día y disminuyen durante la noche. La Figura 19 muestra una correlación de 0,81 entre la temperatura y el oxígeno disuelto, la cual, como una hipótesis, puede haberse dado a causa del proceso de fotosíntesis.

Figura 19

Relación entre temperatura y oxígeno disuelto



Con respecto a los *dashboards* generados con la data obtenida, se logró realizar los *dashboards* en el *software* Grafana y Amazon Quicksight. Sin embargo, debido a que en la zona no había disponibilidad de internet, la data fue representada únicamente en el *software* Grafana. A continuación, en las Figuras 20 y 21 se muestran algunos de los gráficos utilizados en los *dashboards* de Grafana.

Figura 20

Medidores de pH y oxígeno disuelto



Figura 21

Temperatura del aire y del agua en el tiempo



6. DISCUSIÓN

Los resultados obtenidos en la presente investigación demuestran que el dispositivo desarrollado es capaz de monitorear las condiciones de la piscina en tiempo real en condiciones controladas y no controladas.

Para el despliegue se implementó una base de datos local que permitió al sistema trabajar sin internet de forma ininterrumpida. Por otro lado, se pudo aplicar un mecanismo que verifica si el paquete llega completo, y descarta los que llegan con errores; sin embargo, no reenvía los paquetes perdidos. Debido a ello, se

recomienda en trabajos futuros aplicar un sistema que realice esta verificación, así como contabilizar el número de paquetes perdidos y erróneos en distintas zonas y distancias.

Se han podido encontrar investigaciones, como la de Encinas et al. (2017), en las cuales la data es almacenada de forma local para luego ser replicada en la nube; sin embargo, también se tienen artículos que únicamente utilizan una base de datos en la nube (Shi et al., 2018). Sin embargo, el INEI (2021) señala que en las áreas rurales del territorio peruano únicamente el 34,4 % hace uso de internet, por lo que utilizar una base de datos local y replicar los datos en la nube, como lo hicieron Encinas et al. (2017), es lo más adecuado para la investigación.

Por otro lado, los resultados obtenidos de la temperatura del ambiente, la temperatura del agua y la luminosidad presentan valores que se encuentran relacionados con el comportamiento del clima en las horas y la ubicación seleccionada. Asimismo, para evitar distorsión en el registro de temperatura, Weather Works (2020) recomienda tomar mediciones de temperatura fuera de la luz solar, ya que se desea medir la temperatura del aire, mas no la temperatura a la cual el sol puede calentar el sensor. Debido a ello, en el despliegue, se colocó el sensor de temperatura por debajo del panel solar, ya que este le proporciona sombra y evita el error mencionado.

Durante la experimentación, se observó una tendencia relacionada con los efectos de la generación de oxígeno producida por fotosíntesis, mencionados por Hoganson y Babcock (1997), quienes sostienen que en horas con mayor luminosidad las algas del agua generan mayor cantidad de oxígeno, mientras que en la noche es consumido por los peces en el agua, lo que reduce la saturación de este gas.

Finalmente, el sensor de pH muestra un comportamiento bastante estable; sin embargo, se observaron picos o variaciones de forma agresiva. Estas variaciones pudieron ser causadas por un movimiento o golpe en el sensor, ya sea de algún pez en el agua, o por variaciones de voltaje. Cabe mencionar que estas son algunas hipótesis, y dicha variación pudo haberse generado por alguna otra variable.

7. CONCLUSIONES

El sistema fue capaz de capturar las condiciones ambientales dentro y fuera del agua a través de sensores conectados a un microcontrolador, cuya información fue luego enviada vía la red LoRa a un *gateway*. Este capturaba, almacenaba, presentaba reportes y generaba un respaldo casi en tiempo real en una base de datos PostgreSQL alojada en AWS.

Los sensores utilizados en la investigación son capaces de medir las condiciones ambientales de forma continua; sin embargo, sería conveniente que el sensor de luminosidad tenga un mayor rango para medir los picos correctamente. Por otro lado, se concluye que es necesario contar con una base de datos local que funcione como *buffer* en condiciones de baja e inestable conectividad. Aunque la red LoRa resulta bastante eficiente para la comunicación en zonas rurales, es común que ciertos paquetes lleguen incompletos o con errores en la transmisión, por lo que, a pesar de realizar una operación de *checksum*, se recomienda aplicar un método de comunicación bidireccional que permita reenviar paquetes en caso de error. De igual forma, es esencial contar con la facultad de ver los reportes sin conexión a internet, ya que esta es inestable y puede ocasionar problemas de forma recurrente.

Para trabajos futuros, se recomienda a los investigadores incrementar el periodo de medición a una cosecha (alrededor de cuatro meses). Por otro lado, al contar con un dispositivo capaz de medir las condiciones ambientales de un ambiente de producción acuícola las 24 horas del día, sería interesante analizar el desempeño de modelos de *machine learning* para establecer horarios de alimentación sobre la base de las condiciones ambientales actuales, y no en tablas de alimentación preestablecidas. De esta manera, se podría aprovechar más el alimento y ahorrar recursos productivos.

Finalmente, el sistema desarrollado brinda la facilidad a los productores de mantener un control constante de las condiciones de su producción de forma local y remota, así como almacenar su información para un posterior análisis que maximice sus herramientas para la toma de decisiones.

REFERENCIAS

- Abbink, W., Blanco Garcia, A., Roques, J. A. C., Partridge, G. J., Kloet, K., & Schneider, O. (2012). The effect of temperature and pH on the growth and physiological response of juvenile yellowtail kingfish *Seriola lalandi* in recirculating aquaculture systems. *Aquaculture*, 330-333, 130-135. <https://doi.org/10.1016/j.aquaculture.2011.11.043>
- Ai-Thinker. (2017). *Ra-02 LoRa Product Specification V1.1*. https://docs.ai-thinker.com/_media/lora/docs/c048ps01a1_ra-02_product_specification_v1.1.pdf
- Arduino. (s. f.). *Arduino Nano | Arduino Official Store*. *Arduino.cc*. Recuperado el 10 de mayo del 2021, de <https://store.arduino.cc/usa/arduino-nano>
- Banco Central de Reserva del Perú. (2020). *Series trimestrales*. <https://estadisticas.bcrp.gob.pe/estadisticas/series/trimestrales/pbi-millones-de-soles-constantes-de-2007>

- Blanchon, B. (2021). *JSON library for Arduino and embedded C++*. Simple and efficient. GitHub. <https://github.com/bblanchon/ArduinoJson>
- Bormann, M. E. C., & Keranen, A. (2014). *Terminology for constrained-node networks*. Ietf.org. <https://tools.ietf.org/html/rfc7228#section-4>
- Burton, M. (2021). *Arduino plug and go library for the Maxim (previously Dallas) DS18B20 (and similar) temperature ICs*. GitHub. <https://github.com/milesburton/Arduino-Temperature-Control-Library>
- Campbell, J. W. (1973). Nitrogen excretion. En C. L. Prosser (Ed.), *Comparative animal physiology* (pp. 279-316). W. B. Saunders Company.
- Dorsemaine, B., Gaulier, J.-P., Wary, J.-P., Kheir, N., & Urien, P. (2015). Internet of Things: A definition & taxonomy. En *2015 9th International Conference on next Generation Mobile Applications, Services and Technologies* (pp. 72-77). <https://doi.org/10.1109/ngmast.2015.71>
- Encinas, C., Ruiz, E., Cortez, J., & Espinoza, A. (2017). Design and implementation of a distributed IoT system for the monitoring of water quality in aquaculture. En *2017 Wireless Telecommunications Symposium (WTS)* (pp. 1-7). <https://doi.org/10.1109/wts.2017.7943540>
- Fondo Nacional de Desarrollo Pesquero. (2004). *Manual de cultivo de tilapia*. http://www2.produce.gob.pe/RepositorioAPS/3/jer/ACUISUBMENU4/manual_tilapia.pdf
- Grados, B., & Bedon, H. (2020). Software components of an IoT monitoring platform in Google Cloud Platform: A descriptive research and an architectural proposal. En M. Botto-Tobar, M. Zambrano Vizueté, P. Torres-Carrión, S. Montes León, G. Pizarro Vásquez & B. Durakovic (Eds.), *Applied technologies. ICAT 2019. Communications in computer and information science* (vol. 1193, pp. 153-167). https://doi.org/10.1007/978-3-030-42517-3_12
- Hoganson, C. W., & Babcock, G. T. (1997). A metalloradical mechanism for the generation of oxygen from water in photosynthesis. *Science*, 277(5334), 1953-1956. <https://doi.org/10.1126/science.277.5334.1953>
- Instituto Nacional de Estadística e Informática. (2020). *Producto bruto interno trimestral* [Informe técnico]. https://www.inei.gob.pe/media/MenuRecursivo/boletines/boletin_pbi_trimestral_iit_2020.pdf
- Instituto Nacional de Estadística e Informática. (2021). *Estadísticas de las tecnologías de información y comunicación en hogares* [Informe técnico]. <https://www.inei.gob.pe/media/MenuRecursivo/boletines/02-informe-tecnico-tic-i-trimestre-2021.pdf>

- Jawad, H., Nordin, R., Gharghan, S., Jawad, A., & Ismail, M. (2017). Energy-efficient wireless sensor networks for precision agriculture: A review. *Sensors*, 17(8), 1781. <https://doi.org/10.3390/s17081781>
- Kang, M.-S., Jung, Y.-G., & Jang, D.-H. (2017). A study on the search of optimal aquaculture farm condition based on machine learning. *The Journal of the Institute of Internet Broadcasting and Communication*, 17(2), 135-140. <https://doi.org/10.7236/jiibc.2017.17.2.135>
- Kramer, D. L. (1987). Dissolved oxygen and fish behavior. *Environmental Biology of Fishes*, 18(2), 81-92. <https://doi.org/10.1007/bf00002597>
- Laws, C. (2021, 17 de enero). *An Arduino library for the digital light sensor breakout boards containing the BH1750FVIIC*. GitHub. <https://github.com/claws/BH1750>
- Manahan, S. E. (2000). *Environmental chemistry* (7.ª ed.). Lewis Publishers.
- Ministerio de la Producción. (2018). *Sistema Nacional de Innovación en Pesca y Acuicultura. Fundamentos y propuesta 2017-2022*. <https://nipa.gob.pe/wp-content/uploads/2019/02/PESCA-Y-ACUICULTURA-3-1.pdf>
- Mistry, S. (2021). *An Arduino library for sending and receiving data using LoRa radios*. GitHub. <https://github.com/sandeepmistry/arduino-LoRa>
- Organización de las Naciones Unidas para la Alimentación y la Agricultura. (1988). *Rural aquaculture: Overview and framework for country reviews*. <http://www.fao.org/3/x6941e/x6941e04.htm>
- Organización de las Naciones Unidas para la Alimentación y la Agricultura. (2020). *El estado mundial de la pesca y la acuicultura 2020. La sostenibilidad en acción*. <https://doi.org/10.4060/ca9229es>
- Raj, A. (2019, 7 de mayo). *LoRa with Raspberry Pi – Peer to Peer Communication with Arduino*. CircuitDigest. <https://circuitdigest.com/microcontroller-projects/raspberry-pi-with-lora-peer-to-peer-communication-with-arduino>
- Raju, K. R. S. R., & Varma, G. H. K. (2017). Knowledge based real time monitoring system for aquaculture using IoT. En *2017 IEEE 7th International Advance Computing Conference (IACC)* (pp. 318-321). <https://doi.org/10.1109/iacc.2017.0075>
- Rayes, A., & Salam, S. (2016). The things in IoT: Sensors and actuators. En *Internet of Things from hype to reality* (pp. 57-77). https://doi.org/10.1007/978-3-319-44860-2_3
- Ríos Julcapoma, M., & Yauri Rodríguez, R. (2017). *Internet de las cosas en el monitoreo de la calidad del agua para acuicultura en la Amazonía*. Instituto de investigaciones de la Amazonía Peruana; Universidad Nacional de Ingeniería.

- Salim, T. I., Haiyunnisa, T., & Alam, H. S. (2016). Design and implementation of water quality monitoring for eel fish aquaculture. En *2016 International Symposium on Electronics and Smart Devices (ISESD)* (pp. 208-213). <https://doi.org/10.1109/isesd.2016.7886720>
- Sehrawat, D., & Gill, N. S. (2019). Smart sensors: Analysis of different types of IoT sensors. En *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)* (pp. 523-528). <https://doi.org/10.1109/icoei.2019.8862778>
- Servicio Nacional de Meteorología e Hidrología del Perú. (2021). *Promedio de temperatura normal para Pucallpa*. <https://www.senamhi.gob.pe/?p=pronostico-detalle-turistico&localidad=0024>
- Shi, B., Sreeram, V., Zhao, D., Duan, S., & Jiang, J. (2018). A wireless sensor network-based monitoring system for freshwater fishpond aquaculture. *Biosystems Engineering*, *172*, 57-66. <https://doi.org/10.1016/j.biosystemseng.2018.05.016>
- Studd, J., Pollard T., James, R., Trewitt, G., Dangel, J., Lovato, G., Stoffregen, P., Roberts, S., Sikken, B., Tillotson, M., Butcher, K., Clark, R., & Nystrom, L. (2019). *Library for Dallas/Maxim 1-Wire Chips*. GitHub. <https://github.com/PaulStoffregen/OneWire>
- Subasinghe, R. P. (2005). Epidemiological approach to aquatic animal health management: Opportunities and challenges for developing countries to increase aquatic production through aquaculture. *Preventive Veterinary Medicine*, *67*(2-3), 117-124. <https://doi.org/10.1016/j.prevetmed.2004.11.004>
- The Raspberry Pi Foundation. (2021). *Raspberry Pi 4 Model B specifications – Raspberry Pi*. Raspberry Pi. <https://www.raspberrypi.org/products/raspberry-pi-4-model-b/specifications/>
- Weather Works. (2020, 3 de septiembre). *How to measure temperature correctly*. <https://weatherworksinc.com/news/temperature-measurement>
- Yim, D., Chung, J., Cho, Y., Song, H., Jin, D., Kim, S., Ko, S., Smith, A., & Riegsecker, A. (2018). An experimental LoRa performance evaluation in tree farm. En *2018 IEEE Sensors Applications Symposium (SAS)* (pp. 1-6). <https://doi.org/10.1109/sas.2018.8336764>
- Zhu, X., Li, D., He, D., Wang, J., Ma, D., & Li, F. (2010). A remote wireless system for water quality online monitoring in intensive fish culture. *Computers and Electronics in Agriculture*, *71*, S3-S9. <https://doi.org/10.1016/j.compag.2009.10.004>

PROTOCOLO BÁSICO DE CIBERSEGURIDAD PARA PYMES

OLDA BUSTILLOS ORTEGA

obustillos@uia.ac.cr

<https://orcid.org/0000-0003-2822-3428>

Universidad Internacional de las Américas,
Escuela de Ingeniería, Costa Rica

JAVIER ROJAS SEGURA

jrojass@uia.ac.cr

<https://orcid.org/0000-0002-0488-4056>

Universidad Internacional de las Américas,
Escuela de Ingeniería, Costa Rica

RESUMEN. Las restricciones de circulación impuestas por los gobiernos durante la pandemia del COVID-19 modificaron los patrones de consumo y la forma de hacer negocios, obligando a las pequeñas y medianas empresas a migrar a medios electrónicos para no perecer. Del mismo modo, la delincuencia incrementó su participación en el ciberespacio, explotando la vulnerabilidad de estas empresas por su dependencia de las tecnologías digitales y la falta de recursos para un manejo efectivo de la integridad, confidencialidad y disponibilidad de los datos. Por tanto, es necesario sensibilizar a las pequeñas y medianas empresas sobre la importancia de la ciberseguridad y prepararlas no solo para defenderse contra un ciberataque, sino también para una rápida y oportuna recuperación ante un posible incidente. El objetivo de esta investigación es proponer un protocolo básico de ciberseguridad, como principal estrategia de defensa contra ciberataques, que permita la continuidad del negocio en caso de un ataque cibernético. El protocolo se construye a partir de publicaciones académicas, tesis, normas técnicas y guías prácticas, examinadas en el contexto actual de las amenazas globales. La ciberseguridad es percibida por las pequeñas y medianas empresas como demasiado compleja y onerosa, por lo que se requieren soluciones económicas, efectivas y accesibles.

PALABRAS CLAVE: ciberseguridad, pymes, TIC, protocolo

BASIC CYBERSECURITY PROTOCOL FOR SMES

ABSTRACT. Movement restrictions imposed by governments during the COVID-19 pandemic have changed consumption patterns and the way of doing business, forcing small and medium-sized companies to migrate their businesses to electronic media. Similarly, crime increased its participation in cyberspace, exploiting the vulnerability of these companies due to their dependence on digital technologies and the lack of resources for effective management of data integrity, confidentiality, and availability. Therefore, it is necessary to sensitize small and medium-sized companies about the importance of cybersecurity and prepare them not only to defend themselves against cyberattacks but also for a quick and timely recovery from possible incidents. The objective of this study is to propose a basic cybersecurity protocol as the primary defense strategy against cyberattacks, which allows business continuity in the event of a cyberattack. Academic publications, theses, technical standards, and practical guides examined in the current global threats context were the protocol's basis. Small and medium-sized companies perceive cybersecurity as too complex and onerous; thus, they require economical, effective, and accessible solutions.

KEYWORDS: cybersecurity, SME, ICT, protocol

1. INTRODUCCIÓN

En un ambiente empresarial globalizado y competitivo, como el que existe en la actualidad, las pequeñas y medianas empresas (pymes), la sociedad y las compañías, en general, dependen cada vez más de la tecnología, específicamente de sistemas de información, por lo que deben crear políticas de seguridad como un medio de protección, pues se ha demostrado que tienen una enorme influencia para aumentar los niveles de competitividad (Zuñá Macancela et al., 2019). La preparación para la seguridad cibernética es clave para el sustento y la supervivencia en el entorno digital actual (Benz & Chatterjee, 2020), por lo que en el mundo muchas empresas aumentan sus presupuestos en ciberseguridad para prevenir los ciberataques (Zuñá Macancela et al., 2019). Sobre este tema, el Gobierno de Japón (2021) reconoce que para las pymes es difícil destinar un presupuesto importante a la seguridad tecnológica, por lo que es necesario impulsar medidas de esta naturaleza dirigidas a este sector que sean económicas, efectivas y accesibles.

Las pymes presentan datos alarmantes con respecto a ciberseguridad (Peralta Zuñiga & Aguilar Valarezo, 2021), dado que la gran mayoría de los ataques se dirigen a ellas (Ponsard et al., 2019). Más aún, muchas de estas empresas no denuncian haber sufrido un ciberataque por miedo a afectar su reputación (Maggi Murillo & Gómez Gómez, 2021). Por eso, es importante no solo defenderse contra los ataques cibernéticos, sino también prepararse para una respuesta y recuperación rápida y oportuna ante estos incidentes (World Economic Forum [WEF], 2022).

El 57 % de las pymes indica que, ante un ciberataque, lo más probable es que quiebre o deje de funcionar (ENISA, 2021). A pesar de esto, las pymes no parecen darse cuenta de que la ciberseguridad no es algo que afecte solo a las organizaciones más grandes. De acuerdo con Ponsard et al. (2019), la explicación es que las medidas de seguridad se perciben como demasiado complejas, lentas y que requieren un alto nivel de conocimientos técnicos, así como elevados recursos. Tal como recomienda Ramírez Montealegre (2016), es necesario crear un protocolo orientado a las pymes que sea de fácil aplicación, de bajo costo y altamente efectivo.

El objetivo de esta investigación es proponer un protocolo básico de ciberseguridad para pymes, como herramienta fundamental y principal estrategia de defensa contra ciberataques, que permita asegurar la continuidad del negocio en caso de una contingencia. Lo anterior nos lleva a preguntarnos: ¿cómo pueden prevenir las pymes un incidente de seguridad cibernética?

La metodología utilizada abarca las etapas de investigación, adaptación de la información recopilada, revisión por expertos y, finalmente, la propuesta de un protocolo básico de ciberseguridad.

Como resultado de esta investigación se presenta un protocolo de ciberseguridad con el cual los usuarios y gerentes de las pymes pueden alcanzar una base para proteger sus datos, obteniendo una primera contextualización que les permita reconocer los conceptos de ciberseguridad más relevantes para la comprensión de un primer nivel. Este conocimiento irá madurando con el tiempo (Martínez & Blanco, 2020) hasta crear una cultura de ciberseguridad.

2. REVISIÓN DE LITERATURA

Las pequeñas y medianas empresas (pymes) tienen un papel destacado en el desarrollo económico de los países (Semrau et al., 2016), por lo que son consideradas la columna vertebral de las economías en el mundo (Eggers, 2020). En América Latina conforman el 99 % del parque empresarial y generan cerca de dos tercios del empleo de la región (Fernández & Puig, 2022). En Costa Rica las pymes representan el 97,5 % del parque empresarial, aportan el 35,7 % del PIB y contribuyen con el 33,0 % del empleo formal (Faith et al., 2022). A pesar de ello, estas organizaciones empresariales en Latinoamérica no están exentas de limitaciones y, por ende, de enfrentar diversos desafíos (Bartesaghi & Weck, 2022). La dependencia de las pymes de la tecnología e internet abre la puerta a vulnerabilidades frente al cibercrimen, las cuales ocasionan que la seguridad de la información sea un aspecto crítico para todas las pymes (Bruycker & Darville, 2017).

2.1 Situación actual de las pymes

Existe un reconocimiento generalizado entre los líderes de la mayoría de las industrias de que el papel de la tecnología digital está cambiando rápidamente: de ser un impulsor de la eficiencia marginal a ser un fundamental facilitador de la innovación y de la disrupción (Weinelt, 2016). La necesidad de competitividad e innovación de las pymes las convierte en grandes adoptadores de tecnologías digitales, lo que aumenta su exposición a los ciberataques (Ponsard et al., 2019). Estas empresas se encuentran entre las menos maduras y vulnerables en términos de riesgo de ciberseguridad y resiliencia (Benz & Chatterjee, 2020). Enfrentan muchos de los problemas de ciberseguridad que también tienen las grandes empresas, pero no cuentan con los recursos para abordar los riesgos de manera efectiva (Horn, 2017). No obstante, con la ciberpandemia cobrando impulso, no hay industria u organización que esté a salvo (Benz & Chatterjee, 2020).

Las pymes no pueden permitirse el lujo de retrasar su inversión en ciberseguridad (Benz & Chatterjee, 2020), ya que en la actualidad son el objetivo de la gran mayoría de ataques cibernéticos (Ponsard et al., 2019). Deben hacer un plan de contingencia, invertir en herramientas informáticas, capacitar al personal en ciberseguridad para

que todos hablen y manejen el mismo lenguaje (Zuñá Macancela et al., 2019), y tomar conciencia de la importancia de implementar métodos más rígidos de seguridad informática, no únicamente para contrarrestar los perjuicios que puede generar un fraude informático, sino también para prevenirlos (Peralta Zuñiga & Aguilar Valarezo, 2021).

2.2 Seguridad cibernética

La seguridad de la información se ha convertido en una tendencia a nivel mundial, debido al lugar significativo y relevante que ocupa la información para toda compañía y al creciente aumento de amenazas en los últimos años. Actualmente, la seguridad no solo se refiere a la protección de los equipos con antivirus; con el paso del tiempo y el avance de la tecnología, han surgido nuevas técnicas de ataques para vulnerar sistemas informáticos (Morales et al., 2020).

La preparación para la seguridad cibernética está emergiendo como una competencia crítica para la supervivencia y el crecimiento de las organizaciones (Benz & Chatterjee, 2020), especialmente de las pymes. Ponsard et al. (2019) indican que, según la Alianza Nacional de Seguridad Cibernética (NCSA, por sus siglas en inglés), más de la mitad de las pymes hackeadas no pueden recuperarse y van a la quiebra dentro de los seis meses posteriores al ataque.

Los incidentes más frecuentes ocurren por infección de *malware* en los dispositivos (Maggi Murillo & Gómez Gómez, 2021). El término *malware*, acuñado por Yisrael Radaí en 1990, se utiliza para describir la amplia gama de códigos maliciosos (Messmer, 2008). Existen diversas variantes de *malware*, por ejemplo, los virus, las bombas lógicas, los troyanos y los *worms* o gusanos. La diferencia entre ellas consiste en la forma como actúan (Gutiérrez-Cárdenas & Orihuela, 2016).

2.3 Amenazas

La amenaza de los ciberataques es global. Actualmente, existe *malware* capaz de causar un daño significativo a la infraestructura crítica de un país (Cherepanov & Lipovsky, 2017); por ejemplo, en el 2016, Ucrania sufrió el primer ataque de *malware* diseñado específicamente para atacar las redes eléctricas (Kovac, 2022). Los ataques cibernéticos no solo se dirigen a empresas grandes; hoy en día, la mayoría de las organizaciones se encuentran expuestas a este tipo de amenazas (Maggi Murillo & Gómez Gómez, 2021). Incluso el ciberespacio refleja tensiones geopolíticas, competencias entre los Estados y cuestiones de seguridad nacional (Gobierno de Japón, 2021); ejemplo de ello es que casi el 60 % de los ataques de RDP (*remote desktop protocol*) vistos durante el primer trimestre del 2022 provinieron de Rusia (Kovac, 2022). En Costa Rica, en la primera mitad del 2022, sistemas informáticos

El segundo *malware* es un documento malicioso de Microsoft Word, disfrazado de factura, formulario, documento legal u otra información aparentemente importante que descarga más *malware* de internet.

El tercero en importancia es un documento especialmente diseñado que explota la vulnerabilidad de Microsoft Equation Editor, un componente de Microsoft Office. Cuando el usuario abre el documento malicioso, se activa descargando *malware* adicional en la computadora.

El cuarto es un archivo JavaScript malicioso que se coloca en sitios web comprometidos pero legítimos, con el objetivo de lograr un compromiso de los visitantes.

Debido al alto número de ataques reportados en Latinoamérica en los últimos años y la cantidad creciente de robos y estafas realizadas como consecuencia de estos ataques cada año, el *phishing* es la ciberamenaza más grande en la actualidad (Moncada Vargas, 2020), y consiste en obtener información de la víctima de forma fraudulenta, mediante el envío de correos o mensajes que buscan persuadir al usuario para que acceda a sitios maliciosos o falsos, e ingrese su información y así obtener acceso (Maggi Murillo & Gómez Gómez, 2021). Existen otros métodos como, por ejemplo, el *pharming*, que es el redireccionamiento de un usuario que se encuentra en una página legítima a una página *phishing* a través de enlaces directos implantados en la página legítima (Abu-Nimeh et al., 2007). Las métricas más importantes para detectar páginas *phishing* derivan de experiencias humanas (Mao et al., 2018); sin embargo, como señala Moncada Vargas (2020), para identificar estas páginas maliciosas de manera dinámica y automática, se utiliza la inteligencia artificial y el *machine learning*.

2.4 Soluciones

Actualmente, se vive la era de la transformación digital, donde los negocios y sobre todo la información han migrado a medios digitales, lo cual implica un mayor reto para las organizaciones en la protección de sus datos (Morales et al., 2020). Las pymes y la sociedad en general empiezan a tomar conciencia de la enorme importancia de poseer adecuados sistemas de seguridad de la información, así como de la correcta gestión de datos (Zuñá Macancela et al., 2019). El salvaguardar la información es primordial; por ese motivo, se emplea *software anti-malware*, sistemas informáticos de control de acceso e instrucciones y mecanismos de respaldo de datos (Maggi Murillo & Gómez Gómez, 2021). Al proteger estos activos, se asegura su confidencialidad, integridad y disponibilidad (Marchand-Niño & Ventocilla, 2020).

Las organizaciones mantienen un flujo constante de información con su entorno y a través de él puede entrar en riesgo el propio negocio por diversas amenazas,

tanto internas (fuga de información) como externas (suplantación, estafas, *malware*). Estas amenazas siempre están evolucionando, por lo que el *firewall* es el principal mecanismo de defensa y protección en el flujo de la información (Cortés Aldana, 2016). La estrategia de cuidar los datos mediante el uso de un *firewall* perimetral asegura la integridad, confidencialidad y disponibilidad, que son los tres pilares de la seguridad de la información (Morales et al., 2020). Los *firewalls* son dispositivos que buscan proteger la información, por lo que son una de las herramientas principales de seguridad informática (Cortés Aldana, 2016).

Sumado a esto, la ingeniería social continúa siendo el método de propagación de ataques informáticos más utilizado por los creadores de *malware*, quienes aprovechan las ventajas de cualquier medio de comunicación para engañar a los usuarios y lograr que estos terminen cayendo en una trampa que suele apuntar a un fin económico (Borghello, 2009). En estas circunstancias, es necesario un respaldo mantenido en un servidor independiente (lao, 2021). La recuperación de la información obliga a buscar prácticas como los respaldos de los datos.

Es bien sabido que las herramientas tecnológicas no pueden garantizar por sí solas la seguridad de los sistemas; por el contrario, se requiere apoyo de los colaboradores dentro de la organización (Ponsard et al., 2019). Los usuarios de las pymes deben seguir actualizando sus conocimientos en temas de ciberseguridad a través de capacitaciones en temáticas relacionadas con seguridad de la información (Maggi Murillo & Gómez Gómez, 2021).

3. METODOLOGÍA

La gestión de riesgos puede resultar compleja para las pymes (Ramírez & González, 2020), por lo que en esta investigación cualitativa se propone un protocolo básico de ciberseguridad fundamentado en las etapas metodológicas de investigación, adaptación, revisión por expertos y propuesta.

3.1 Investigación

En esta etapa se procedió a la recopilación de información relacionada con la ciberseguridad de las pymes, consultando bases de datos tales como Google Scholar, ProQuest Digital Dissertation and Theses e IEEE Xplore. Como resultado de esta etapa, se examinaron diversos trabajos de investigación (Benz & Chatterjee, 2020), tesis (Orellana, 2020), normas técnicas (ISO/IEC 27000, 2018), guías (Bruycker & Darville, 2017), así como buenas prácticas y tendencias en ciberseguridad (WEF, 2022).

Por otro lado, se analizó el contexto actual de las amenazas y vulnerabilidades que sufren las empresas a nivel global (ESET, 2022), y se tomó la información más reciente relacionada con amenazas en seguridad relevantes para este artículo.

3.2 Adaptación de la información recopilada

Una vez identificada la información de interés para la creación del protocolo, se procedió a la traducción en los casos en los que fue requerido; se seleccionó el contenido de artículos y documentos, así como la adaptación de la redacción, con vistas al objetivo de la investigación.

3.3 Revisión por expertos

La información recopilada y adaptada fue sometida al conocimiento de expertos locales en la materia. Así, se obtuvieron sugerencias, cambios y consejos para la construcción del protocolo. En esta etapa metodológica, se determinó que el protocolo estuviese integrado por los siguientes procesos: definición del perfil del usuario, políticas de privacidad, antivirus, *firewall*, políticas de acceso, infraestructura, respaldo o *backup* y sistemas de administración gerencial.

3.4 Propuesta de un protocolo

Producto de las anteriores etapas metodológicas, se construyó el protocolo, adaptando el conocimiento existente al contexto y requerimientos de las pymes, para ser implementado de una manera simple, práctica y efectiva. De este modo, las pymes pueden tener un nivel de protección razonable para resguardar la información y minimizar el riesgo.

4. PROPUESTA DE UN PROTOCOLO BÁSICO DE CIBERSEGURIDAD

La norma ISO/IEC 27001 ayuda a las pymes a estructurar la capacitación en seguridad cibernética de acuerdo con las mejores prácticas internacionales, así como a definir responsabilidades en caso de incumplimiento (BSI, 2022). Tal como lo expresan Marchand-Niño y Ventocilla (2020), una de las dificultades de estos estándares y regulaciones es el número y diversidad de controles que contienen; por ejemplo, la familia ISO/IEC establece 114 controles, mientras que el Instituto Nacional de Normas y Tecnología (NIST, por sus siglas en inglés) del Departamento de Comercio de los Estados Unidos define 444 controles de seguridad. De acuerdo con Ramírez Montealegre (2016), la implementación de este tipo de normas implica un gran esfuerzo y costo para las pymes, por lo que es necesario establecer un protocolo básico de fácil implementación y evaluación. En la Tabla 1 se propone el protocolo básico de ciberseguridad para pymes.

Tabla 1*Protocolo básico de ciberseguridad para pymes*

Nombre	Descripción	Fuentes
Definición del perfil de usuario	Definir los lineamientos para gestionar los privilegios de usuario para cada una de las unidades administrativas, departamentos o dependencias, de acuerdo con la estructura organizacional y el manual de funciones, teniendo claras cuáles son las actividades y la interacción de cada usuario con los sistemas informáticos en la organización para limitar su acceso. En caso de que la pyme no tenga dicha información, esta se debe construir según las funciones de los usuarios de cada sistema.	(Bruycker & Darville, 2017), (Australian Cyber Security Centre [ACSC], 2021), (Cybersecurity and Infrastructure Security Agency [CISA], 2021), (NCSA, 2018), (Marchand-Niño & Ventocilla, 2020), (Martínez & Blanco, 2020), (Cabezas Juárez, 2020)
Políticas de privacidad	En caso de que la pyme utilice el comercio electrónico, para que este sea fiable para el cliente potencial, se deben cumplir los requisitos legales y reglamentarios de privacidad, manejo de datos y seguridad, a los que obliga la legislación de cada país. Si se realizan compras en línea, se deben configurar los ajustes de privacidad para reforzar la seguridad y limitar la cantidad de datos compartidos con sus proveedores. En este punto también se deben definir políticas de uso de sitios, evitando que los colaboradores accedan a sitios riesgosos.	(Bruycker & Darville, 2017), (NCSA, 2018), (Navarro Uriol, 2020)
Antivirus	Toda pyme, sin importar su tamaño, debe definir los estándares y lineamientos básicos para el uso de antivirus en los equipos de cómputo. Este <i>software</i> debe estar actualizado y tener las respectivas licencias de uso.	(Bruycker & Darville, 2017), (ACSC, 2021), (CISA, 2021), (ENISA, 2021), (Gutiérrez-Cárdenas & Orihuela, 2016), (Marchand-Niño & Ventocilla, 2020), (De la Rosa, 2019), (Martínez & Blanco, 2020)

(continúa)

(continuación)

Nombre	Descripción	Fuentes
Firewall	Un <i>firewall</i> perimetral ayuda a monitorear, detectar y bloquear la mayor parte de las amenazas que se producen diariamente. Se ubica habitualmente en el punto de conexión de la red interna de la pyme con la red exterior (internet). Su función es realizar un filtrado, permitir y negar el paso a "intrusos" que no cumplan con las políticas que se configuran en el equipo, a fin de proteger la red interna de intentos de acceso no autorizados. La implementación de un <i>firewall</i> es técnicamente sencilla; desafortunadamente, no puede ofrecer protección una vez que el agresor lo traspasa o permanece en el entorno.	(Morales et al., 2020), (ENISA, 2021), (NCSA, 2018), (Bruycker & Darville, 2017), (Marchand-Niño & Ventocilla, 2020), (De la Rosa, 2019), (Martínez & Blanco, 2020), (Cuenca, 2016), (CISA, 2021)
Políticas de acceso	Es importante determinar los permisos de acceso apropiados para cada usuario y grupos de usuarios. Es necesario definir una lista de los usuarios a quienes se les permite el acceso a ciertos sitios y también a los que se les niega.	(Bruycker & Darville, 2017), (ACSC, 2021), (CISA, 2021), (NCSA, 2018), (Marchand-Niño & Ventocilla, 2020), (De la Rosa, 2019), (Ramírez & González, 2020)
Infraestructura	Se refiere a la selección de la arquitectura de operación, es decir, si se utilizará un servidor físico o en la nube, o bien una combinación, esto es, un servidor local con réplica en entorno <i>cloud</i> . Se deben establecer criterios de almacenamiento, control de accesos y perfiles de usuario, entre otros. En caso de que un tercero brinde el servicio, es recomendable tener el alcance claramente definido en un contrato.	(Martínez & Blanco, 2020), (Bruycker & Darville, 2017), (Palafox-Pascual, 2019), (NCSA, 2018), (Navarro Uriol, 2020), (Gobierno de Japón, 2021)
Respaldo de la información sensible	Se sugieren respaldos de los datos sensibles de la organización; para ello, es importante definir qué tipo de datos se deben resguardar, la frecuencia del respaldo, el tipo de dispositivos utilizados y el lugar para resguardar los respaldos, entre otros.	(Cabezas Juárez, 2020), (Martínez & Blanco, 2020), (Bruycker & Darville, 2017), (lao, 2021), (Palafox-Pascual, 2019), (ACSC, 2021), (CISA, 2021), (ENISA, 2021), (Marchand-Niño & Ventocilla, 2020), (De la Rosa, 2019)

(continúa)

(continuación)

Nombre	Descripción	Fuentes
Definición de la metodología para el desarrollo y administración de los sistemas	Los sistemas de información gerencial pueden ser comprados a un tercero o desarrollados internamente. En el primer caso, es importante detallar las expectativas y el alcance en un acuerdo o contrato. En cambio, las pymes que tienen sistemas de información, sea que lo desarrollen internamente o que lo compren a la medida, deberán contar con una metodología y estándares claros para el desarrollo y administración de estos sistemas. La capacitación y mejora continua permitirá accionar anticipadamente para llevar un control de qué información se debe respaldar. Las bases de datos y códigos fuente son lo mínimo recomendable.	(Anchundia Betancourt, 2017), (Bruycker & Darville, 2017), (NCSA, 2018)

El capital humano es el factor fundamental para la transformación cultural (Díaz, 2021). Se requiere una estrecha cooperación entre los colaboradores, la gestión formativa y la gerencia de todas las áreas y funciones de la organización (Vergara-Romero et al., 2021). Para la adecuada implementación de este protocolo, se sugiere complementar con las siguientes dos dimensiones:

4.1 Involucrar a la dirección de la pyme

Tal como lo establece la norma ISO/IEC 27000, una de las bases fundamentales sobre las cuales hay que iniciar un proyecto de este tipo es el apoyo claro y decidido de la dirección de la organización, ya que el cambio de cultura y concientización que implica el proceso hace necesario el impulso constante de las autoridades. Es necesario sensibilizar a los directivos y avanzar en las iniciativas de las empresas para fortalecer la ciberseguridad en línea con la digitalización (Gobierno de Japón, 2021).

4.2 Formación del personal

La principal amenaza que afecta la seguridad de la información de una pyme es el desconocimiento del concepto mismo (Inoguchi & Macha, 2017). De la Rosa (2019) explica que la ciberseguridad dentro de una empresa sigue el modelo de una cadena, que se rompe por su eslabón más débil; en el caso de la ciberseguridad, sabemos

que ese eslabón son los colaboradores. Es todo un reto conseguir que el colaborador esté plenamente identificado con acciones para la protección de los datos de la organización, pero se puede lograr a través de la sensibilización, la capacitación y la divulgación y mejora continua.

4.2.1 Sensibilización

Ramírez Montealegre (2016) recomienda tener programas de concientización en ciberseguridad en las pymes, dándoles un enfoque práctico y tangible que ayude a identificar situaciones de riesgo y pueda aportar en las decisiones y acciones. De la Rosa (2019) pide sensibilizar al personal sobre los riesgos cibernéticos, para que todos y cada uno de los colaboradores se autoperciban como parte fundamental de la ciberseguridad de la empresa, creando una cultura consciente del riesgo. Casi cualquier colaborador puede proporcionar una puerta abierta a la red de una organización, y los ciberdelincuentes lo saben (NCSA, 2018); por ello, es indispensable que los colaboradores también lo entiendan.

4.2.2 Capacitación

Navarro Uriol (2020) indica que cada colaborador debe proteger su puesto de trabajo, por lo que se le debe capacitar sobre el manejo seguro del correo electrónico, sitios web, dispositivos móviles, teletrabajo, redes sociales, unidades USB, etcétera. Debido a los ataques relacionados con errores humanos, como *phishing* u otras técnicas de ingeniería social (Palafox-Pascual, 2019), es primordial capacitar también sobre ingeniería social, para que los colaboradores conozcan los tipos de engaños, tretas y artimañas (Borghello, 2009) de los ciberdelincuentes, los cuales se orientan a que el usuario comprometa al sistema y revele información valiosa a través de acciones que van desde un clic hasta atender un llamado telefónico, los cuales pueden derivar en la pérdida de información confidencial.

4.2.3 Divulgación y mejora continua

Martínez y Blanco (2020) recomiendan generar una estrategia de difusión del programa de gestión de la seguridad que se va a implementar, ya que se requiere que todos los que integran la organización participen para producir un nivel de seguridad. Las personas son una gran debilidad en la ciberseguridad, pero cuando se involucran y se capacitan correctamente, pueden convertirse en la primera línea de defensa contra los atacantes (Ponsard et al., 2019); por ello, es necesario que las pymes comuniquen y refuercen constantemente a sus colaboradores sobre estos temas para crear una cultura de ciberseguridad mediante la formación continua y la actualización de las herramientas tecnológicas utilizadas.

5. CONCLUSIONES

La ciberseguridad no ha sido una alta prioridad para la mayoría de las pymes (Benz & Chatterjee, 2020), aun cuando la mayoría de ellas parecen tener un nivel de conciencia sobre la importancia de la ciberseguridad. Al observar las estadísticas de ataques, continúa habiendo fallas. Una primera explicación es que las medidas de seguridad se perciben como demasiado complejas, lentas y que requieren un alto nivel de conocimientos técnicos sobre los sistemas de informática. Otra razón es la dificultad para pasar de la concientización inicial a la emergencia de una cultura de ciberseguridad interna, debido a la falta de recursos, tales como dinero, tiempo y experiencia (Ponsard et al., 2019).

Las escasas medidas de seguridad de la información, la poca capacitación del personal de la empresa y las deficientes políticas de seguridad informática en las pymes han tenido un impacto negativo en el desarrollo de sus actividades comerciales (Zuñá Macancela et al., 2019). La gerencia, al momento de gestionar sus riesgos, también tiene que enfocarse en el aspecto informático, pues debe ser consciente del riesgo tecnológico, las amenazas cibernéticas y las deficiencias de los sistemas informáticos, ya que estos elementos también ponen en peligro el cumplimiento de los objetivos organizacionales y, por ende, la continuidad del negocio (Peralta Zuñiga & Aguilar Valarezo, 2021).

Las pymes se colocarían en la ruta para evolucionar de una etapa de concientización del riesgo a la construcción de una cultura de ciberseguridad si hicieran una adecuada utilización de los recursos escasos para salvaguardar la confidencialidad, integridad y disponibilidad del activo más valioso: la información. La transición de las pymes no es un proceso rápido, pero la implementación de este protocolo básico es un paso firme para construir una cultura de ciberseguridad.

REFERENCIAS

- Abu-Nimeh, S., Nappa, D., Wang, X., & Nair, S. (2007). A comparison of machine learning techniques for phishing detection. En *Proceedings of the Anti-Phishing Working Groups 2nd Annual eCrime Researchers Summit* (pp. 60-69). <https://doi.org/10.1145/1299015.1299021>
- Anchundia Betancourt, C. E. (2017). Ciberseguridad en los sistemas de información de las universidades. *Dominio de las Ciencias*, 3(3), 200-217. <https://dialnet.unirioja.es/servlet/articulo?codigo=6102849>
- Australian Cyber Security Centre. (2021). *Personal cyber security: Next steps guide*. <https://www.cyber.gov.au/acsc/view-all-content/guidance/personal-cyber-security-next-steps-guide>

- Bartesaghi, I., & Weck, W. (Eds.). (2022). *Los efectos de la digitalización, inteligencia artificial, big data e industria 4.0 en el trabajo de las pymes en Latinoamérica*. Konrad-Adenauer-Stiftung e. V.; Universidad Católica del Uruguay. <https://www.kas.de/en/web/regionalprogramm-adela/single-title/-/content/losefectos-de-la-digitalizacion-inteligencia-artificial-big-data-e-industria-4-0-en-eltrabajo-de-l>
- Benz, M., & Chatterjee, D. (2020). Calculated risk? A cybersecurity evaluation tool for SMEs. *Business Horizons*, 63(4), 531-540. <https://doi.org/10.1016/j.bushor.2020.03.010>
- Borghello, C. (2009). *El arma infalible: la ingeniería social*. ESET. https://www.eset-la.com/pdf/prensa/informe/arma_infalible_ingenieria_social.pdf
- Bruycker, M., & Darville, C. (2017). *Cyber security guide for SME*. Centre for Cyber Security Belgium. <https://ccb.belgium.be/en/document/guide-sme>
- BSI. (2022). Cybersecurity confidence for the SME. *BSI Blog*. <https://www.bsigroup.com/enGB/blog/Small-Business-Blog/cybersecurity-confidence-for-the-sme/>
- Cabezas Juárez, I. C. (2020). *Implementación de un framework de ciberseguridad compuesto por normas y controles para proteger la información de las pequeñas y medianas empresas en Lima* [Tesis de grado, Universidad de San Martín de Porres]. Registro Nacional de Trabajos de Investigación. <https://renati.sunedu.gob.pe/handle/sunedu/2847596>
- Cherepanov, A., & Lipovsky, R. (2017, 12 de junio). *Industroyer: Biggest threat to industrial control systems since Stuxnet*. WeLiveSecurity. <https://www.welivesecurity.com/2017/06/12/industroyer-biggest-threat-industrialcontrol-systems-since-stuxnet/>
- Cortés Aldana, D. G. (2016). *Firewalls de nueva generación: la seguridad informática vanguardista* [Tesis de grado, Universidad Piloto de Colombia]. Re-Pilo. <http://repository.unipiloto.edu.co/handle/20.500.12277/2719>
- Cuenca, J. (2016). *Firewall o cortafuegos*. Universidad Nacional de Loja. https://www.researchgate.net/publication/295256426_FIREWALL_O_CORTAFUEGOS
- Cybersecurity and Infrastructure Security Agency. (2021). *Cyber essentials starter kit*. https://www.cisa.gov/sites/default/files/publications/Cyber%20Essentials%20Starter%20Kit_03.12.2021_508_0.pdf
- De la Rosa, J. (2019). *Ciberseguridad para pymes* [Trabajo de fin de grado, Universidad de Valladolid]. Universidad de Valladolid, Repositorio Documental. <https://uvadoc.uva.es/handle/10324/38735>

- Díaz, M. (2021). *Estado de la ciberseguridad en la logística de América Latina y el Caribe*. Comisión Económica para América Latina y el Caribe, serie Desarrollo Productivo n.º 228. <https://repositorio.cepal.org/handle/11362/47240>
- Eggers, F. (2020). Masters of disasters? Challenges and opportunities for SMEs in times of crisis. *Journal of Business Research*, 116, 199-208. <https://doi.org/10.1016/j.jbusres.2020.05.025>
- ENISA. (2021). *Guía de ciberseguridad para pymes*. The European Union Agency for Cybersecurity. https://www.enisa.europa.eu/publications/report-files/smes-leaflettranslations/enisa-cybersecurity-guide-for-smes_es.pdf
- ESET. (2022, 2 de junio). *Threat Report T1 2022*. https://www.welivesecurity.com/wp-content/uploads/2022/06/eset_threat_report_t12022.pdf
- Faith, M., Leiva, J. C., & Mora, R. (2022). Las pymes en Costa Rica. En I. Bartesaghi & W. Weck (Eds.), *Los efectos de la digitalización, inteligencia artificial, big data e industria 4.0 en el trabajo de las pymes en Latinoamérica*. Konrad-Adenauer-Stiftung e. V.; Universidad Católica del Uruguay. <https://www.kas.de/en/web/regionalprogramm-adela/single-title/-/content/losefectos-de-la-digitalizacion-inteligencia-artificial-big-data-e-industria-4-0-en-eltrabajo-de-l>
- Fernández, M. C., & Puig, P. (2022). *Los desafíos del comercio electrónico para las PyME: principales claves en el proceso de digitalización*. Banco Interamericano de Desarrollo. <https://publications.iadb.org/es/los-desafios-del-comercio-electronico-para-las-pyme-principales-claves-en-el-proceso-de>
- Gobierno de Japón. (2021). *Cybersecurity for all*. <https://www.nisc.go.jp/pdf/policy/kihons/cs-senryaku2021-en-booklet.pdf>
- Gutiérrez-Cárdenas, J. M., & Orihuela, L. L. (2016). Filogenia de *malware* orientada al análisis de librerías. *Interfases*, 9, 67-86. <https://doi.org/10.26439/interfases2016.n009.1241>
- Horn, A. (2017, 11 de diciembre). Cybersecurity should be a top concern for middle-market companies. *SmallBizDaily*. <https://www.smallbizdaily.com/cybersecurity-middlemarket-companies/>
- Iao, K. (2021, 23 de abril). What is a remote desktop protocol attack? *Paubox*. <https://www.paubox.com/blog/what-is-remote-desktop-protocol-attack/>
- Inoguchi, A., & Macha, E. L. (2017). *Gestión de la ciberseguridad y prevención de los ataques cibernéticos en las pymes del Perú, 2016* [Tesis de grado, Universidad San Ignacio de Loyola]. <https://repositorio.usil.edu.pe/handle/usil/2810>

- Kovac, R. (2022, 2 de junio). *Foreword ESET Threat Report T1 2022*. https://www.welivesecurity.com/wpcontent/uploads/2022/06/eset_threat_report_t12022.pdf
- Lara, J. F. (2022, 22 de agosto). Costa Rica recibió 513 millones de intentos de ciberataques en primer semestre. *La Nación*. <https://www.nacion.com/el-pais/servicios/costarica-recibio-513-millones-de-intentos-de/YFLBY3DU55GRDJGL76NGX4V6UQ/story/>
- Maggi Murillo, G., & Gómez Gómez, O. S. (2021). Estudio preliminar sobre conocimiento de ciberseguridad en usuarios de PYMEs: caso de estudio en Riobamba. *Perspectivas*, 3(2), 45-53. <https://doi.org/10.47187/perspectivas.vol3iss2.pp45-53.2021>
- Mao, J., Bian, J., Tian, W., Zhu, S., Wei, T., Li, A., & Liang, Z. (2018). Detecting phishing websites via aggregation analysis of page layouts. *Procedia Computer Science*, 129, 224-230. <https://doi.org/10.1016/j.procs.2018.03.053>
- Marchand-Niño, W.-R., & Ventocilla, E. J. V. (2020). Modelo Balanced Scorecard para los controles críticos de seguridad informática según el Center for Internet Security (CIS). *Interfases*, 13, 57-76. <https://doi.org/10.26439/interfases2020.n013.4876>
- Martínez, J. A., & Blanco, L. X. (2020). *Recomendaciones de buenas prácticas de ciberseguridad en Pymes para la generación de soluciones de detección de intrusos usando Snort* [Tesis de grado, Universidad Autónoma de Bucaramanga]. <https://repository.unab.edu.co/handle/20.500.12749/13911>
- Messmer, E. (2008, 29 de junio). *Tech Talk: Where'd it come from, anyway?* PCWorld. <https://web.archive.org/web/20121016035507/https://www.pcworld.com/article/147698/tech.html>
- Moncada Vargas, A. E. (2020). Comparación de técnicas de *machine learning* para detección de sitios web de *phishing*. *Interfases*, 13, 77-103. <https://doi.org/10.26439/interfases2020.n013.4886>
- Morales, F., Toapanta, S., & Toasa, R. M. (2020). Implementación de un sistema de seguridad perimetral como estrategia de seguridad de la información. *Revista Ibérica de Sistemas e Tecnologías de Informação*, E27, 553-565. https://www.researchgate.net/publication/339956501_Implementacion_de_un_sistema_de_seguridad_perimetral_como_estrategia_de_seguridad_de_la_informacion
- National Cyber Security Alliance. (2018). *The cybersecurity awareness toolkit*. <https://staysafeonline.org/wp-content/uploads/2018/09/SMB-Toolkit-FINAL.pdf>

- Navarro Uriol, C. (2020). *Estrategias de ciberseguridad: el caso de la pequeña y mediana empresa* [Trabajo de fin de grado, Universidad de Zaragoza]. Zaguan. Repositorio Institucional de Documentos. <https://zaguan.unizar.es/record/101988/files/TAZ-TFG2020-1242.pdf?version=1>
- Orellana, F. D. (2020). *Cybersecurity incident response capabilities in the Ecuadorian small business sector: A qualitative study* [Tesis de doctorado, Northcentral University, School of Business and Technology Management]. ProQuest Dissertations & Theses Global. <http://www.proquest.com/pqdtglobal/docview/2466034020/abstract/6BDCDD913D1D469EPQ/1>
- Organización Internacional para Estandarización & International Electrotechnical Commission (2018). *ISO/IEC 27000*. <https://www.normasiso.net/wpcontent/uploads/2016/10/iso-27000.pdf>
- Palafox-Pascual, L. (2019). *NUTRIA: "Una metodología de ciberseguridad para pymes en entornos industriales"* [Tesis de maestría, Universidad Internacional de La Rioja]. Re-Unir. Repositorio Digital. <https://reunir.unir.net/handle/123456789/9422>
- Peralta Zuñiga, M. L., & Aguilar Valarezo, D. N. (2021). La ciberseguridad y su concepción en las PYMES de Cuenca, Ecuador. *Contabilidad y Auditoría*, 53, 99-126. <https://ojs.econ.uba.ar//index.php/Contyaudit/article/view/2061>
- Ponsard, C., Grandclaudon, J., & Bal, S. (2019). Survey and lessons learned on raising SME awareness about cybersecurity. En *Proceedings of the 5th International Conference on Information Systems Security and Privacy (ICISSP)* (pp. 558-563). <https://doi.org/10.5220/0007574305580563>
- Ramírez, C., & González, J. C. (2020). *Guía de controles y buenas prácticas de ciberseguridad para mipymes* [Trabajo de grado, Tecnológico de Antioquia Institución Universitaria]. Repositorio Digital TDEA. <https://dspace.tdea.edu.co/handle/tdea/1394>
- Ramírez Montealegre, B. (2016). *Medición de madurez de ciberseguridad en pymes colombianas* [Tesis de maestría, Universidad Nacional de Colombia]. Repositorio Institucional. Biblioteca Digital. <https://repositorio.unal.edu.co/handle/unal/57956>
- Semrau, T., Ambos, T., & Kraus, S. (2016). Entrepreneurial orientation and SME performance across societal cultures: An international study. *Journal of Business Research*, 69(5), 1928-1932. <https://doi.org/10.1016/j.jbusres.2015.10.082>

Vergara-Romero, A., Márquez Sánchez, F., Sorhegui-Ortega, R., & Olalla-Hernández, A. (2021). Capital humano: actor central para la sostenibilidad organizacional. *Revista Venezolana de Gerencia*, 26(93), 297-307.

Weinelt, B. (2016). *Digital transformation of industries*. World Economic Forum. <https://www.weforum.org/reports/digital-transformation-of-industries/>

World Economic Forum. (2022). *Global Cybersecurity Outlook 2022*. <https://www.weforum.org/reports/global-cybersecurity-outlook-2022/>

Zuñiga Macancela, E. R., Arce Ramírez, Á. A., Romero Berrones, W. J., & Soledispa Baque, C. J. (2019). Análisis de la seguridad de la información en las PYMES de la ciudad de Milagro. *Universidad y Sociedad*, 11(4), 487-492. http://scielo.sld.cu/scielo.php?script=sci_abstract&pid=S221836202019000400487&lng=es&nrm=iso&tlng=en

Recibido: 26 de junio del 2022 / Aceptado: 3 de octubre del 2022
doi: <https://doi.org/10.26439/interfases2022.n016.5953>

APLICACIÓN DE *MACHINE LEARNING* PARA CAMPAÑAS DE *MARKETING* EN LA BANCA COMERCIAL

GANÍMEDES T. ROSALES REYES

grosalesr@unmsm.edu.pe

<https://orcid.org/0000-0001-6415-6109>

Universidad Nacional Mayor de San Marcos, Perú

XAVIER ALBERTO GUTIERREZ CORAL

xavier.gutierrez@unmsm.edu.pe

<https://orcid.org/0000-0002-0250-0864>

Universidad Nacional Mayor de San Marcos, Perú

AUGUSTO ENRIQUE HAYASHIDA MARCHINARES

augusto.hayashida@unmsm.edu.pe

<https://orcid.org/0000-0002-4278-5008>

Universidad Nacional Mayor de San Marcos, Perú

RESUMEN. Los bancos usan el *telemarketing* para contactar directamente con los clientes potenciales para sus productos. Este canal de venta es complejo, pues requiere de grandes bases de datos de posibles prospectos y está sujeto a restricciones de tiempo y personal. Este artículo tiene tres objetivos: comparar cinco modelos de predicción basados en algoritmos de aprendizaje automático para encontrar el que ofrezca la mejor precisión predictiva; desplegar un piloto de este modelo; y recomendar una hoja de ruta para la futura arquitectura que lo soporte. Se encontró que el algoritmo seleccionado mejora considerablemente la eficacia de la identificación de clientes que aceptan el producto, que pasó de 11 % a 94 %, por lo que su implementación puede contribuir a la competitividad de estas organizaciones.

PALABRAS CLAVE: banca, *marketing*, depósitos a plazo fijo, aprendizaje automático, algoritmos de clasificación

MACHINE LEARNING APPLICATION FOR CAMPAIGNS MARKETING IN COMMERCIAL BANKING

ABSTRACT. Banks use telemarketing to contact potential customers for their products directly. This sales channel is complex, requiring large databases of possible prospects, and is subject to time and personnel restrictions. This article has three objectives: to compare five prediction models based on machine learning algorithms to find the one that offers the best predictive accuracy, deploy a pilot of this model, and recommend a roadmap for the future architecture that supports it. The comparison results show that the selected algorithm considerably improves the identification of customers who accept the product, which went from 11 % to 94 %, so its implementation can contribute to the competitiveness of these organizations.

KEYWORDS: banking, marketing, fixed-term deposits, machine learning, classification algorithms

1. INTRODUCCIÓN

Una parte importante del negocio bancario son las operaciones pasivas, como la apertura de cuentas dirigidas a los clientes minoristas, en sus diferentes formas, ya sean cuentas de ahorros, cuentas corrientes o cuentas a plazo fijo. Sin embargo, uno de los desafíos que enfrenta la banca comercial en este ámbito es el hecho de tener que contactar a una cantidad importante de clientes, aun con los escasos recursos, tanto de tiempo como de materiales, con los que se cuenta. Por ello, para lograrlo los bancos han adoptado la ejecución de campañas comerciales en modalidad masiva o en modalidad directa, es decir, dirigidas a segmentos específicos de clientes (Dutta et al., 2021).

No obstante, es sabido que los resultados de este tipo de campañas no tienen ratios de éxito altos, por lo que es necesario el uso de estrategias cada vez más elaboradas a fin de mejorar su efectividad. En este esfuerzo, desde hace unos años, muchos bancos han ido incorporando de manera progresiva herramientas basadas en *machine learning* y minería de datos con el objetivo de incrementar el nivel de éxito de sus campañas, identificando para ello los principales factores que pueden conducir a él (Moro et al., 2011).

La presente investigación tiene tres objetivos. El primero de ellos es mostrar la evaluación de cuatro algoritmos entre los más comunes en el campo de los sistemas cognitivos a fin de demostrar las ventajas de este tipo de herramientas para la gestión comercial en la banca personal; los algoritmos evaluados fueron (i) *decision tree*, (ii) KNN, (iii) ANN (RPROP), (iv) *naive Bayes* y (v) XGBoost. Para dicha evaluación, se tomaron los datos de una campaña comercial llevada a cabo por un banco europeo con una muestra de 41 188 registros, y se utilizó la estrategia de particiones, es decir, se subdividió en dos grupos ese conjunto de datos, de los cuales el 70 % servirá para que el algoritmo evaluado aprenda y el otro 30 % para que dicho algoritmo sea probado. El segundo objetivo es mostrar como demostración una aplicación web que implementa el algoritmo que mejor desempeño tuvo. Por último, el tercer objetivo consiste en proponer una arquitectura empresarial futura con visión integral que permita construir una solución integral al proceso de campañas de *marketing*.

2. ESTADO DEL ARTE

A fin de lograr los tres objetivos planteados, se describe a continuación el estado del arte para los algoritmos de *machine learning*, así como para el concepto de arquitectura empresarial.

Respecto a los algoritmos de *machine learning*, se puede afirmar que los bancos están usando diferentes técnicas en sus implementaciones para descubrir y luego

utilizar patrones a partir de sus bases de datos. Entre los principales se encuentran los siguientes:

2.1 *Decision tree*

Los árboles de decisión se han convertido en uno de los modelos más potentes y populares en la ciencia de datos, como ciencia y tecnología de exploración de grandes y complejos conjuntos de datos, donde ayuda a descubrir patrones útiles. El objetivo del árbol de decisión es clasificar una instancia en un conjunto predefinido de clases en función de los valores de sus atributos (Rokach, 2016). Asimismo, Zhang et al. (2019) indican que, a diferencia de la regresión logística, el árbol de decisión se construye de arriba hacia abajo en una secuencia; esto es, todo el conjunto de datos se divide en particiones más pequeñas hasta que no se pueden realizar más particiones.

2.2 *K-nearest neighbors (KNN)*

Es un modelo simple y eficaz que no requiere parámetros. El proceso de clasificación de KNN consiste en realizar el cálculo de la similitud entre un objeto objetivo y los k vecinos más cercanos y similares en el conjunto de muestra de entrenamiento. La distancia de similitud de KNN normalmente se mide por la distancia euclidiana (Zhang et al., 2019).

Figura 1

Fórmula para el cálculo de la distancia que utiliza el método KNN

$$d(x, x_i) = \sqrt{\sum_{i=1}^n (x - x_i)^2}$$

Nota. Donde x es el objetivo y x_i es el i -ésimo vecino similar más cercano. Luego, al estar x más cerca de sus vecinos, el destino se asignará a la clase más común entre sus k vecinos más cercanos.

2.3 *Redes neuronales ANN (RPROP)*

La ANN está inspirada biológicamente en el cerebro humano. Las neuronas están interconectadas en el cerebro humano de manera similar a como los nodos están interconectados en la red neuronal artificial (Asha & Kumar, 2021). De acuerdo con Riedmiller y Braun (1993), es un algoritmo de aprendizaje para redes de alimentación directa multicapa, que fue propuesto para superar las desventajas inherentes del descenso-gradiente puro. RPROP realiza una adaptación local de las actualizaciones

ponderadas según el comportamiento de la función de error. A diferencia de otros modelos adaptativos, el efecto del proceso de adaptación de RPROP no se ve empañado por la influencia imprevisible del tamaño de la derivada, sino que solo depende del comportamiento temporal de su signo. A fin de mostrar la lógica del algoritmo, en la Figura 1 se presenta un fragmento de pseudocódigo que muestra el núcleo del proceso de adaptación y aprendizaje de RPROP.

Figura 2

Núcleo del proceso de adaptación de RPROP

```

For all weights and biases{
  if  $\left(\frac{\partial E}{\partial w_{ij}}(t-1) * \frac{\partial E}{\partial w_{ij}}(t) > 0\right)$  then {
     $\Delta_{ij}(t) = \text{minimum}(\Delta_{ij}(t-1) * \eta^+, \Delta_{\max})$ 
     $\Delta w_{ij}(t) = - \text{sign} \frac{\partial E}{\partial w_{ij}}(t) * \Delta_{ij}(t)$ 
     $w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}(t)$ 
  }
  else if  $\left(\frac{\partial E}{\partial w_{ij}}(t-1) * \frac{\partial E}{\partial w_{ij}}(t) < 0\right)$  then {
     $\Delta_{ij}(t) = \text{maximum}(\Delta_{ij}(t-1) * \eta^-, \Delta_{\min})$ 
     $w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}(t-1)$ 
     $\frac{\partial E}{\partial w_{ij}}(t) = 0$ 
  }
  else if  $\left(\frac{\partial E}{\partial w_{ij}}(t-1) * \frac{\partial E}{\partial w_{ij}}(t) = 0\right)$  then {
     $\Delta w_{ij}(t) = - \text{sign} \frac{\partial E}{\partial w_{ij}}(t) * \Delta_{ij}(t)$ 
     $w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}(t)$ 
  }
}
    
```

Nota. Reproducido de "A Direct Adaptive Method for Faster Backpropagation Learning: The RPROP Algorithm" (p. 588), por M. Riedmiller y H. Braun, 1993, *IEEE International Conference on Neural Networks*.

Se supone que el operador mínimo (máximo) debe entregar el mínimo (máximo) de dos números. El operador de signo devuelve +1 si el argumento es positivo; -1 si el argumento es negativo; y 0 en caso contrario.

2.4 Naive Bayes

Este modelo de clasificación permite estimar la probabilidad de una hipótesis a partir de la data experimental. Conforme se suma más data a la muestra, se va ajustando la probabilidad resultante. Su base matemática es el teorema de Bayes. En el modelo se asume que el valor de cada característica de la base de datos es independiente de las demás características, dada la variable de clase, y que cada una contribuye independientemente a la predicción final, sin establecerse correlaciones entre ellas; de allí la calificación de *naive* o ingenuo (Berrar, 2018).

Dado un vector $x = (x_1, x_2, \dots, x_n)$ que representa una instancia por ser clasificada, formada por n variables independientes, el modelo considera la asignación de probabilidades para cada clase k (Murty & Devi, 2011).

$$p(C_k | x_1, x_2, \dots, x_n)$$

Según el teorema de Bayes, se puede descomponer la probabilidad condicional de la siguiente manera:

$$p(C_k | x) = (p(C_k) p(x | C_k)) / (p(x))$$

$p(C_k)$ es la probabilidad *a priori* de la hipótesis.

$p(x | C_k)$ es probabilidad de los datos dada la hipótesis o *likelihood*.

$p(x)$ es la probabilidad de observar los datos, independientemente de la hipótesis especificada o evidencia.

2.5 XGBoost

Es un algoritmo de aprendizaje automático basado en un árbol de decisiones y utiliza un marco de potenciación de gradientes (Data Science Team, 2019). Las redes neuronales artificiales superan en su mayoría a otros marcos o algoritmos cuando predicen problemas con texto, imágenes y otros datos no estructurados. XGBoost (*Extreme Gradient Boosting*) es una implementación avanzada del algoritmo de refuerzo de gradiente, y ha demostrado ser un algoritmo de ML muy eficaz, ampliamente utilizado en competiciones de aprendizaje automático y *hackathons*. Tiene un alto poder de predicción y es casi diez veces más rápido que las otras técnicas de *gradient boosting* (Singh, 2018). También incluye una variedad de regularización que reduce el sobreajuste y mejora el rendimiento general, de ahí que también se conozca como *técnica de refuerzo regularizado*.

El algoritmo XGBoost fue el resultado de un proyecto de investigación organizado en la Universidad de Washington. Carlos Guestrin y Tianqi Chen presentaron su trabajo en la conferencia SIGKDD del 2016. Desde entonces, XGBoost se ha convertido en un renombrado algoritmo que revolucionó el panorama del aprendizaje de las máquinas. Cabe señalar que permite utilizar una amplia gama de aplicaciones para resolver problemas de predicción, clasificación y regresión definidos por el investigador.

Aparte de los algoritmos descritos, a fin de proveer a la solución que se va a implementar de un enfoque integrador y holístico, se plantea como marco de trabajo el concepto de arquitectura empresarial (AE). Para Goethals et al. (2006), la AE es un acercamiento holístico para el manejo y gestión de una organización, el cual adopta una vista integral que cubre sus procesos de negocio, los sistemas de información, los datos e información y la infraestructura tecnológica. De acuerdo con Lankhorst (2005), la AE es un conjunto coherente de principios, métodos y modelos que se utilizan en el diseño y la realización de la estructura organizacional de una empresa, los procesos de negocio, los sistemas de información y la infraestructura. Finalmente, Arango Serna et al. (2010) explican que el campo de conocimiento de la AE ha evolucionado con el objeto de hacer frente a dos problemas importantes: el primero es la capacidad de gestionar la creciente complejidad de los sistemas de información en las organizaciones, y el segundo se refiere al incremento en la dificultad en la generación de valor real por parte de los sistemas de información para las empresas.

3. METODOLOGÍA

3.1 Descripción del conjunto de datos

- Número de observaciones: 41 188.
- Número de muestras realizadas para el presente estudio:
 - Entrenamiento: 28 831 observaciones, correspondientes al 70 % del conjunto de datos.
 - Prueba: 12 357 observaciones, correspondientes al 30 % del conjunto de datos.
 - Número de variables de entrada: 20.
 - Número de variables de salida: 1.

La Tabla 1 muestra las variables del conjunto de datos.

Tabla 1

Descripción de variables

Tipo de datos	N.º	Variable	Escala	Tipo de variable
Datos del cliente	1	Edad del cliente	Cuantitativa discreta	Entrada
	2	Ocupación	Cualitativa politómica nominal	Entrada
	3	Estado civil	Cualitativa politómica nominal	Entrada
	4	Educación	Cualitativa politómica nominal	Entrada
	5	Mora bancaria	Cualitativa politómica nominal	Entrada
	6	Crédito hipotecario	Cualitativa politómica nominal	Entrada
	7	Crédito personal	Cualitativa politómica nominal	Entrada
Datos últimos de contacto de esta campaña	8	Medio de contacto	Cualitativa dicotómica nominal	Entrada
	9	Mes de último contacto	Cualitativa politómica nominal	Entrada
	10	Día de la semana del último contacto	Cualitativa politómica nominal	Entrada
	11	Duración del último contacto	Cuantitativa discreta	Entrada
Otros datos	12	Contactos en esta campaña	Cuantitativa discreta	Entrada
	13	Número de días desde que fue contactado la última vez	Cuantitativa discreta	Entrada
	14	Número de contactos antes de esta campaña	Cuantitativa discreta	Entrada
	15	Resultado de campañas previas	Cualitativa politómica nominal	Entrada
Datos del contexto social y económico	16	Tasa de variación del empleo	Cuantitativa continua	Entrada
	17	Índice de precios al consumo	Cuantitativa continua	Entrada
	18	Índice de confianza del consumidor	Cuantitativa continua	Entrada
	19	Índice euribor a 3 meses	Cuantitativa continua	Entrada
	20	Número de empleados	Cuantitativa discreta	Entrada
Clase	21	Cliente suscribió depósito en cuenta	Cualitativa dicotómica nominal	Salida

3.2 Preparación de los datos

Teniendo en cuenta que el número de clientes que aceptan campañas de *telemarketing* es mucho menor que aquellos que las rechazan, el *dataset* se encontraba

desbalanceado. Por este motivo fue necesario balancearlo utilizando la técnica estadística de sobremuestreo.

Debido a que varios de los modelos de clasificación utilizados requieren operar con datos numéricos en lugar de cadenas de caracteres, también se convirtieron los atributos de tipo cualitativo dicotómico o politómico a escala numérica. Se filtraron adicionalmente en el conjunto de datos final atributos invariables de contexto social y económico.

4. PRUEBAS

A partir del conjunto de datos definidos, se implementaron los modelos de *decision tree*, KNN, *naive Bayes*, ANN (RPROP) y XGBoost. Estos modelos fueron seleccionados con el objetivo de obtener el modelo predictivo de la mayor exactitud posible.

La Tabla 2 muestra la matriz de confusión de los cinco modelos desarrollados y sus respectivas gráficas en la Figura 2. Por su parte, en la Tabla 3 se aprecian las métricas de desempeño de los algoritmos evaluados. En ambos casos, la información se obtuvo luego de balancear el *dataset* de origen.

Tabla 2

Matriz de confusión de los algoritmos decision tree, KNN, naive Bayes, ANN (RPROP) y XGBoost

	Algoritmo <i>decision tree</i>		Algoritmo KNN		Algoritmo <i>naive Bayes</i>		Algoritmo ANN (RPROP)		Algoritmo XGBoost	
	No	Sí	No	Sí	No	Sí	No	Sí	No	Sí
No	9950	966	9112	1804	9656	1260	8963	1953	10 071	788
Sí	589	10 424	224	10 789	6477	4536	1557	9456	663	10 407

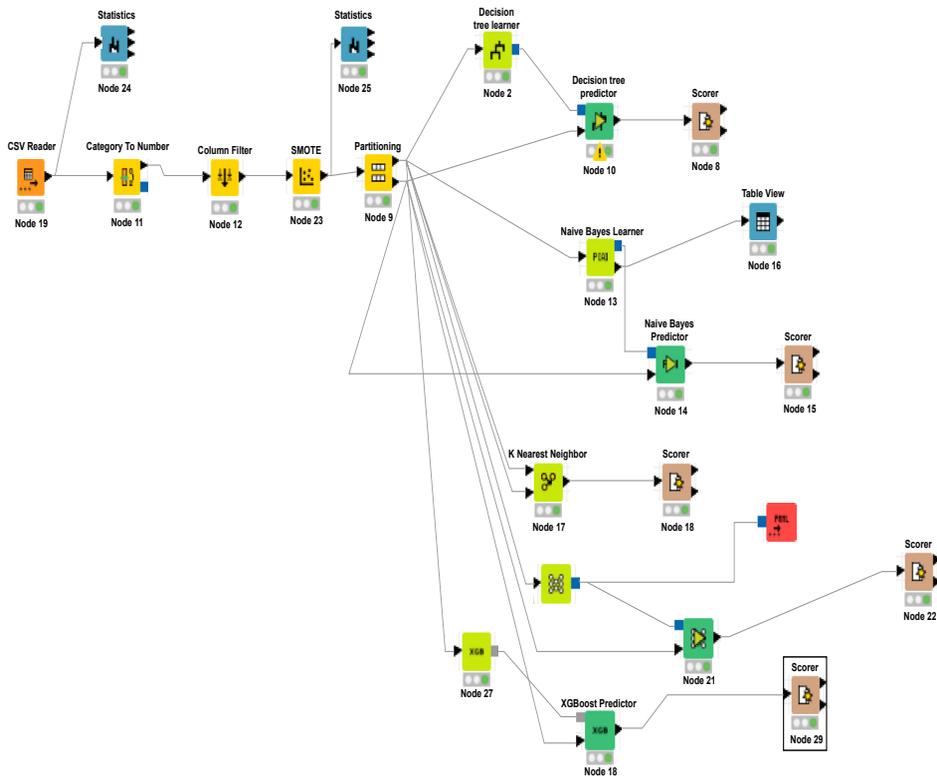
Tabla 3

Resultados de métricas de desempeño de los algoritmos decision tree, KNN, naive Bayes, ANN (RPROP) y XGBoost

	Algoritmo <i>decision tree</i>	Algoritmo KNN	Algoritmo <i>naive Bayes</i>	Algoritmo ANN (RPROP)	Algoritmo XGBoost
Accuracy	0,929	0,908	0,647	0,840	0,934
Recall	0,947	0,980	0,412	0,859	0,940
Precision	0,915	0,857	0,783	0,829	0,930
F-measure	0,931	0,914	0,540	0,843	0,935

Figura 3

Modelo del flujo de información analizada



5. ANÁLISIS DE RESULTADOS

En este artículo se realizó una revisión completa de los algoritmos *decision tree*, KNN, *naive Bayes*, ANN (RPROP) y XGBoost, así como su aplicación para la predicción de decisiones en la priorización de contactos con clientes en campañas comerciales para la venta del producto de depósitos a plazo. Se descubrió que los algoritmos de *machine learning* tienen diferentes aproximaciones teóricas y formas de aplicación. En términos generales, los modelos *decision tree*, KNN, *naive Bayes*, ANN (RPROP) y XGBoost son sensibles a los datos desequilibrados y son adecuados para muestras de gran tamaño.

Para evaluar el desempeño de los cinco algoritmos en análisis, se seleccionaron las métricas de *accuracy*, *recall*, *precision* y la medida F1 para la clase objetivo, es decir, aquellos clientes de interés para el banco por ser potenciales ahorristas y

que estarían dispuestos a aceptar la campaña. Los resultados de esta evaluación se muestran en la Tabla 3.

Para la métrica *accuracy*, el algoritmo XGBoost es el que muestra un mayor valor con 0,934, seguido de *decision tree* con 0,929; luego KNN con 0,908; ANN (RPROP) con 0,840; y, finalmente, *naive Bayes* con 0,647. Para la métrica *recall*, el algoritmo KNN obtiene el mayor valor con 0,980, seguido por *decision tree* con 0,947; luego ANN (RPROP) con 0,859; y, finalmente, *naive Bayes* con 0,412. En caso de la métrica *precision*, el algoritmo XGBoost obtiene 0,930, seguido de *decision tree* con 0,915; luego KNN con 0,857; luego ANN (RPROP) con 0,829; y, finalmente, *naive Bayes* con 0,783. Para la métrica F1, el valor mayor es del algoritmo XGBoost con 0,935; luego *decision tree* con 0,931, seguido de KNN con 0,914; luego ANN (RPROP) con 0,843; y, finalmente, *naive Bayes* con 0,540. Asimismo, para llevar a la práctica el modelo, se implementó una aplicación web en la nube, la cual fue probada como piloto y dio los resultados esperados.

En la siguiente sección, se presentan recomendaciones para extender y masificar este modelo, integrándolo a los procesos y aplicativos de una institución financiera.

6. FUTUROS TRABAJOS

Uno de los objetivos de los sistemas de información es su alineamiento con las estrategias de negocios, y la arquitectura empresarial es la disciplina que permite a la organización que sus recursos de tecnologías de la información respondan adecuadamente a las fuerzas disruptivas de su entorno. La arquitectura empresarial define un marco de trabajo con una organización lógica en capas que describe los diferentes elementos de los procesos de negocio y tecnologías de la información (Saleem & Fakieh, 2020). En este sentido, consideramos que un análisis bajo el enfoque de la arquitectura empresarial es necesario para identificar los siguientes pasos en la evolución de soluciones de campañas de *marketing* en instituciones financieras, para extender el modelo más allá de los modelos algorítmicos de *machine learning*, a fin de integrar este componente con las demás aplicaciones y procesos de la institución que se relacionan con la comercialización de productos bancarios.

En la capa de información y datos, se puede evaluar el incorporar la captura de variables relacionadas con el uso y preferencias de los usuarios actuales que aceptaron la campaña y compraron el producto, de manera que se refinan los modelos para facilitar la venta recurrente a los mismos clientes, con campañas y atributos personalizados.

En la capa de aplicación, se puede analizar la manera de integrar la solución con otros canales de la institución y las aplicaciones que las soportan. De manera que el proceso de comercialización no esté solamente integrado a campañas de *telemarketing* telefónico, sino que, durante el uso de los canales transaccionales, y de manera contextual, los clientes puedan recibir propuestas comerciales que eventualmente sean calificadas y desembolsadas en línea.

Finalmente, en la capa de infraestructura, si bien las instituciones financieras usualmente han favorecido la utilización de recursos *on-premise*, las opciones en nube permiten acceder a recursos para atender integralmente un proceso de campañas. La habilitación de elementos de integración a las aplicaciones *core* financieras permitiría que una capa de campañas que resida en la nube pueda acceder a la información para identificar eventos que gatillen las ofertas a los clientes. La utilización de recursos como *data lakes* de *marketing* permitiría consolidar masivamente la información de manera que pueda ser aprovechada en múltiples usos posteriores, sin necesidad de una transformación previa (Miloslavskaya & Tolstoy, 2016).

7. CONCLUSIONES

La clasificación de datos mediante técnicas de *machine learning* se puede utilizar para mejorar la eficacia en la toma de decisiones de los responsables de las áreas comerciales en las entidades financieras, según las variables seleccionadas y sus ponderaciones.

Dado que el conjunto de datos de entrada es desbalanceado, por el menor ratio de clientes que acepta la campaña comercial respecto a los que la rechazan, fue necesario realizar un balanceo de los datos utilizando la técnica SMOTE (Martínez Heras, 2020).

Sobre la base del desempeño de la métrica *accuracy*, se recomienda el algoritmo XGBoost para las predicciones en este tipo de campañas comerciales.

Teniendo en cuenta que el ratio de éxito actual de las campañas que realiza el banco es del 11 % (4640 clientes de la muestra total), el modelo recomendado incrementa este ratio al 94 % de acuerdo con la métrica *recall*, que mide la capacidad de identificar a aquellos clientes que realmente están interesados en los productos ofertados.

REFERENCIAS

- Arango Serna, M. D., Londoño Salazar, J. E., & Zapata Cortés, J. A. (2010). Arquitectura empresarial: una visión general. *Revista Ingenierías Universidad de Medellín*, 9(16), 101-111.
- Asha, R. B., & Kumar, K. R. (2021). Credit card fraud detection using artificial neural network. *Global Transitions Proceedings*, 2(1), 35-41. <https://doi.org/10.1016/j.gltp.2021.01.006>
- Berrar, D. (2018). Bayes' theorem and naive Bayes classifier. En *Encyclopedia of bioinformatics and computational biology: ABC of bioinformatics* (vol. 1, pp. 403-412). Elsevier. <https://doi.org/10.1016/B978-0-12-809633-8.20473-1>
- Data Science Team. (2019). *XGBoost. ¿Qué es?* <https://Datascience.Eu/Es/Programacion/Xgboost-4/>
- Dutta, S., Bose, P., Goyal, V., & Bandyopadhyay, S. K. (2021). Applying convolutional-GRU for term deposit likelihood prediction. *International Journal of Engineering and Management Research*, 11(3), 265-272. <https://doi.org/10.20944/preprints202007.0101.v1>
- Goethals, F. G., Snoeck, M., Lemahieu, W., & Vandenbulcke, J. (2006). Management and enterprise architecture click: The FAD(E)E framework. *Information Systems Frontiers*, 8(2), 67-79. <https://doi.org/10.1007/s10796-006-7971-1>
- Lankhorst, M. (2005). *Enterprise architecture at work. Modelling, communication and analysis*. Springer.
- Martínez Heras, J. (2020, 9 de octubre). *Precision, recall, F1, accuracy en clasificación*. IArtificial.net. <https://www.iartificial.net/precision-recall-f1-accuracy-en-clasificacion/>
- Miloslavskaya, N., & Tolstoy, A. (2016). Big data, fast data and data lake concepts. *Procedia Computer Science*, 88, 300-305. <https://doi.org/10.1016/j.procs.2016.07.439>
- Moro, S., Laureano, R., & Cortez, P. (2011). Using data mining for bank direct marketing: An application of the CRISP-DM methodology. En P. Novais, J. Machado, C. Analide & A. Abelha (Ed.), *ESM'2011. The 2011 European Simulation and Modelling Conference* (pp. 117-121). EUROISIS.
- Murty, M. N., & Devi, V. S. (2011). *Pattern recognition an algorithmic approach*. Springer.
- Riedmiller, M., & Braun, H. (1993). A direct adaptive method for faster backpropagation learning: The RPROP algorithm. En *IEEE International Conference on Neural Networks* (vol. 1, pp. 586-591). DOI: 10.1109/ICNN.1993.298623

- Rokach, L. (2016). Decision forest: Twenty years of research. *Information Fusion, 27*, 111-125. <https://doi.org/10.1016/j.inffus.2015.06.005>
- Saleem, F., & Fakieh, B. (2020). Enterprise architecture and organizational benefits: A case study. *Sustainability, 12*(19), 8237. <https://doi.org/10.3390/su12198237>
- Singh, A. (2018, 18 de junio). *A comprehensive guide to ensemble learning*. Analytics Vidhya. <https://www.analyticsvidhya.com/blog/2018/06/comprehensive-guide-for-ensemble-models/>.
- Zhang, C., Zhang, H., & Hu, X. (2019). A contrastive study of machine learning on funding evaluation prediction. *IEEE Access, 7*, 106307-106315. <https://doi.org/10.1109/ACCESS.2019.2927517>

ARTÍCULOS DE REVISIÓN

DISEÑO DE INTERFACES DE SISTEMAS INTERACTIVOS UTILIZANDO TÉCNICAS DE *MACHINE LEARNING*: UNA REVISIÓN DEL DISEÑO Y LA USABILIDAD

JULIO VLADIMIR QUISPE SOTA
julio.quispe17@unmsm.edu.pe
<https://orcid.org/0000-0003-1413-5160>
Universidad Nacional Mayor de San Marcos, Perú

RESUMEN. En el presente trabajo se exponen algunos enfoques planteados para el diseño de interfaces de usuario mediante técnicas de *machine learning*. En la primera parte, se hace una revisión de diversos enfoques para el diseño de interfaces, tales como los optimizadores combinatoriales, el uso de *frameworks* y el diseño de interfaces libres de texto. En cuanto al diseño de interfaces con técnicas de *machine learning*, se observa que tienen como base la usabilidad y la experiencia del usuario o UX. Asimismo, el proceso de diseño utiliza interacciones que son almacenadas en sistemas de persistencia o bases de datos, que luego se analizan con técnicas de *machine learning*. Otro enfoque de diseño usa bosquejos y maquetados gráficos y, luego de evaluar su usabilidad, emplea algoritmos de reconocimiento de imágenes para generar las interfaces; estos diseños generalmente son para dispositivos móviles. Por otro lado, existen técnicas que también analizan la usabilidad, pero se enfocan más en las funciones corporales del usuario (movimiento, funciones biológicas como presión sanguínea, latidos del corazón, etcétera); estos datos también pueden analizarse con algoritmos de *machine learning* para generar interfaces de usuario.

PALABRAS CLAVE: *user interface* (UI), experiencia del usuario (UX), *machine learning* (ML), usabilidad

DESIGN OF INTERACTIVE SYSTEM INTERFACES USING MACHINE LEARNING TECHNIQUES: A REVIEW OF DESIGN AND USABILITY

ABSTRACT. This article presents different approaches to user interface design through machine learning techniques. It reviews various approaches to interface design, such as combinational optimizers, frameworks, and free-of-text interface design. Moreover, it shows how interface design with machine learning techniques is based on usability and user experience (UX). Likewise, the design process uses interactions stored in persistence systems or databases, which are then analyzed with machine learning techniques. Another design approach uses sketches and graphic layouts and, after evaluating their usability, uses image recognition algorithms to generate the interfaces; these designs are generally for mobile devices. Some techniques also analyze usability but focus more on the user's bodily functions (movement, biological functions such as blood pressure, heartbeat, etcetera); this data can also be analyzed with machine learning algorithms to generate user interfaces.

KEYWORDS: user interface (UI), user experience (UX), machine learning (ML), usability

1. INTRODUCCIÓN

El diseño de interfaces de usuario para sistemas interactivos es bastante estudiado y cada vez existen mejores técnicas para generar interacciones mucho más intuitivas y cómodas para los usuarios. Pressman (2010) plantea que la flexibilidad de la interfaz es un aspecto muy importante para el diseño de la interfaz de usuario (UI, por sus siglas en inglés), dada la diversidad de usuarios que utilizarán el sistema. Por otro lado, Braun et al. (2021) muestran que las emociones son una gran fuente de entradas y salidas al interactuar con el sistema. También, mediante muestreo y obtención de datos, se puede producir un ciclo de retroalimentación para generar UI más adaptativas al usuario, sin dejar de prestar atención a aspectos tan diversos como los culturales y sociales. Yang (2017), por su parte, señala que aún no existen metodologías estandarizadas basadas en *machine learning* para el diseño de interfaces.

Para Mezhoudi (2013), la experiencia del usuario (UX, por sus siglas en inglés) se puede capturar mediante las interacciones del usuario con el sistema, para luego ser almacenada y crear un conjunto de datos o *dataset*, donde una técnica de *machine learning* podría determinar patrones de comportamiento del usuario y, por consiguiente, se puedan generar interfaces más intuitivas o adaptativas. Es importante mencionar que las necesidades específicas de los usuarios pueden ser reconocidas con la ayuda de los algoritmos de *machine learning*.

Este trabajo está estructurado de la siguiente manera: la sección 2 explica la metodología del estudio de revisión, la sección 3 describe los resultados del estudio y, en la sección 4, se plantean las discusiones. Finalmente, se encuentran las conclusiones y las referencias.

2. METODOLOGÍA

En este estudio, se plantearon las siguientes preguntas, que se orientan al conocimiento del estado del arte de las investigaciones sobre el diseño de interfaces de usuario y el uso de técnicas de *machine learning*:

- P1: ¿qué enfoques existen para el diseño de interfaces de usuario?
- P2: ¿qué es la usabilidad en un sistema interactivo?
- P3: ¿hasta qué punto es aplicable *machine learning* en el diseño de interfaces de usuario?
- P4: ¿es necesaria la experiencia de usuario como base para la aplicación de una técnica de *machine learning*?
- P5: ¿de qué manera se aplica *machine learning* en el diseño de interfaces de usuario?

2.1 Búsqueda de artículos

Para responder estas preguntas, se realizó la búsqueda de artículos en Google Scholar utilizando la siguiente cadena de búsqueda: design usability “user experience” UX “User Interface” UI “machine learning” ML “Artificial Intelligence” software engineering.

El resultado fueron 479 trabajos, de los cuales se seleccionaron solo los que tenían que ver con el tema de estudio, eran de acceso abierto y, además, fueron publicados desde el año 2017. Los artículos de tipo revisión no fueron seleccionados. La Tabla 1 muestra un resumen de la cantidad de artículos obtenidos y su fuente principal.

Tabla 1

Cantidad de artículos seleccionados para el estudio

Fuente	Número de artículos
IEEE Xplore	17
SciELO	1
ResearchGate	3
Otras fuentes	13
Total	34

Se utilizó un texto de consulta sobre ingeniería de *software* para establecer el marco conceptual sobre el cual se define el proceso de diseño de interfaces de usuario. Una de las limitaciones de este estudio fue la falta de acceso a repositorios comerciales como Scopus, Web of Science y otros.

3. RESULTADOS

A partir de la revisión a los trabajos seleccionados en esta investigación, se puede responder a las cinco preguntas planteadas con las siguientes propuestas.

3.1 Los enfoques para el diseño de interfaces de usuario

En la actualidad, los tipos de interfaces de usuario (IU o UI, por sus siglas en inglés) son muy diversos, tales como gráficos interactivos, comandos de selección, interacción 3D, diseño de documentos, etcétera. Para esta diversidad de interfaces, se plantean métodos de diseño, como la optimización combinatoria propuesta por Oulasvirta et al. (2020). El trabajo se enfoca en la selección de las funciones de programa que son presentadas y manipuladas por el usuario, así como de los tipos de componentes y

propiedades que poseen; también, en asociar interacciones con cambios de estado en el programa y en la organización de los componentes en los contenedores para formar una jerarquía.

La utilización de *frameworks* para el diseño de UI es, asimismo, un método de diseño. Por ejemplo, Sanctorum et al. (2019) plantean un *framework* que genera UI adaptativas, distribuidas e híbridas. Aplicando el *framework*, el diseño de la UI se divide en la capa activa, que contiene el código de las acciones que realizarán los componentes; otra capa que contiene los elementos de la UI; y, finalmente, la capa de modelos que vincula la capa activa y la capa de UI.

Las UI cuyo principal medio de interacción es el ingreso de textos o comandos tienen una limitante para los usuarios por el idioma, sobre todo si la UI no utiliza el alfabeto latino o, peor aún, si los usuarios son analfabetos. En ese sentido, Ilyas et al. (2021) definen el concepto de interfaces libres de texto, las cuales no limitan al usuario a las interacciones de solo texto. Así, algunos *frameworks* permiten que el usuario pueda ubicar sectores de interfaz de usuario con texto, imágenes o símbolos.

Por otro lado, Sboui et al. (2017) utilizan la combinación del desarrollo de interfaces de usuarios basado en modelos (MBUID) y la línea de productos de *software* dinámico (DSPL) para plantear el modelo UI-DSPL, que permite generar una UI adaptable a un contexto. Este modelo de diseño plantea dos fases: la primera, dedicada al desarrollo de una UI inicial, mientras que la segunda funciona en tiempo de ejecución; esta permitirá adaptar la UI a los cambios del contexto que provienen por parte del usuario.

3.2 La usabilidad en un sistema interactivo

La usabilidad de la interfaz de usuario es esencial para la interacción de los usuarios con el sistema. Para Ferreira et al. (2020), “la usabilidad es el grado en que un producto o sistema puede ser utilizado por usuarios específicos para lograr objetivos específicos con eficacia, eficiencia y satisfacción en un contexto de uso específico”. Por otro lado, según Sánchez (2015), el grado de usabilidad puede ser medido empíricamente usando las pruebas o experimentos con los usuarios. El grado de usabilidad no define que una interfaz sea buena o mala, sino que satisfaga a un grupo mayoritario de una población objetivo. De acuerdo con Sackl et al. (2020), los usuarios —tanto novatos como experimentados— muestran preferencia por interfaces optimizadas, que utilizan descripción en los componentes.

3.3 La aplicación de *machine learning* en el diseño de interfaz de usuario

Para Yang (2017), mediante *machine learning* (ML), se busca optimizar la experiencia del usuario (UX) en sistemas interactivos, de tal forma que se puedan generar

interfaces más adaptativas que reduzcan el esfuerzo de navegación de los usuarios. Según Yang (2017), los expertos en UX deben construir interfaces de tal manera que se puedan rastrear las interacciones de los usuarios, y estas se puedan usar como fuente de datos para los algoritmos de ML. En ese sentido, ML provee un conjunto de patrones de interfaz de usuario a fin de adaptar o automatizar la navegación en la interfaz y poder cambiar a varios contextos para diferentes usuarios en tiempo de ejecución.

Otra propuesta es la de Akinsola et al. (2021), que consiste en la fusión de una técnica de inteligencia artificial con una interfaz de usuario para generar un híbrido denominado *intelligent user interface* (IUI). Esta nueva interfaz es capaz de minimizar los errores que pueda cometer el usuario al momento de utilizar el sistema, prevenir riesgos de ciberseguridad y funcionar como un asistente para el usuario. Las técnicas para el diseño de la IUI se pueden resumir en tres puntos:

1. Obtención de entradas por parte del usuario de manera inteligente, lo que se puede realizar a través de detección de movimientos, reconocimiento de rostro y expresiones, procesamiento de lenguaje natural, entre otros.
2. Modelado de usuario, que indica todas las técnicas de comunicación que permitan la adaptación de la interacción humano-máquina en diferentes entornos.
3. Generación de explicación, técnica que se enfoca en la capacidad de habilitar un sistema de explicación para los usuarios como agentes de UI, retroalimentación, entre otros.

En la actualidad, la interacción con los sistemas computacionales es muy diversa. Las fuentes de entrada son más naturalistas, tales como el habla, los gestos o la mirada. Además, el estudio de la interacción no debe limitarse a computadoras o dispositivos portables, como celulares, tabletas, etcétera. Braun et al. (2021) plantean el concepto de UI afectiva para la conducción de automóviles. Tomando como base la detección de emociones que provienen del conductor, este modelo considera las preferencias y características permanentes de los usuarios. En ese sentido, los sistemas pueden adaptarse a largo plazo mediante la recopilación de conocimientos y estados de usuario para que puedan reaccionar espontáneamente, a fin de mejorar una situación.

Otra aplicación de ML es anticipar la navegabilidad en las aplicaciones. Joo y Lee (2019) proponen el uso de un *framework* denominado *WebProfiler* para la captura de las interacciones (clics o navegación entre páginas) de los usuarios, de tal modo que se pueda predecir la navegabilidad en la aplicación web. En esa línea, Cruz-Benito et al. (2017) aplican técnicas de ML y el enfoque de la prueba A/B para detectar

las preferencias de usuario, el comportamiento y la versión óptima del formulario para todas las clases de usuarios que utilicen la aplicación. Además, se construyen modelos usando algoritmos de ML, con los cuales se extraen las características más importantes y luego se agrupa a los usuarios según sus similitudes; a partir de ello, se generan las reglas heurísticas entre las diferentes versiones de los formularios web para obtener el más adecuado por cada grupo de usuarios.

3.4 El rol de la experiencia del usuario para la aplicación de una técnica de *machine learning* en el diseño

Yang (2018) plantea las siguientes maneras de implementar el rol de la UX cuando se desea aplicar una técnica de ML para el diseño de UI.

3.4.1 Agregar un machine learning a una experiencia de usuario existente

Según Yang (2018), muchos especialistas diseñadores no consideran la aplicación de ML en sus diseños, además de que esta no es muy clara, a pesar de que es muy utilizada en diversas áreas. Asimismo, la experiencia de usuario (UX) no puede ser una ocurrencia tardía para la aplicación de ML, dado que, para la comunidad de especialistas en diseño y desarrollo, la usabilidad y la UX deben ser consideradas desde etapas tempranas en el proceso. Los diseñadores que utilizan con regularidad una técnica ML, primero, ejecutan la aplicación y a partir de allí se buscan patrones de comportamiento del usuario.

3.4.2 Agregar una experiencia de usuario a un machine learning existente

Para Yang (2018), a través del análisis del contexto tanto cultural como social de los usuarios, se determinará el éxito del uso de ML. En la industria, encontrar el diseño adecuado no será muy simple si se utilizan procedimientos tradicionales; sin embargo, un modelo de ML no garantiza una interfaz óptima, pero sí facilita muchos aspectos de navegabilidad, usabilidad y adaptabilidad. Este planteamiento es reforzado por otros autores que sostienen que la aplicación del modelo ML puede no ser suficiente para que un usuario tenga un enganche con el sistema. Hay casos en los cuales existe una disrupción entre un ML y la UX, por cuestiones sociales, culturales, etcétera.

3.5 Las técnicas de *machine learning* aplicadas en el diseño de interfaz de usuario

Las propuestas revisadas abarcan el diseño de estructuras complejas capaces de interpretar y plantear reglas de comportamiento a partir del almacenamiento de interacciones del usuario con el sistema. Interacciones que utilizan reconocimiento

de movimientos para la navegabilidad en el sistema y algoritmos de reconocimiento de imágenes, tales como las redes neuronales convolucionales, para diseñar la interfaz a partir de una maqueta en papel o en imagen digital.

3.5.1 *Adaptación de la interfaz de usuario basada en la retroalimentación y machine learning*

Mezhoudi (2013) se enfoca en la premisa de asegurar la adaptabilidad del aprendizaje mediante la interacción del usuario. Para llegar a este objetivo, se necesitan estructuras más complejas como un módulo de gestión de adaptación de reglas (ARM, por sus siglas en inglés). Las reglas se almacenan en un sistema de persistencia, del cual se pueden obtener los conjuntos de datos y, mediante el ARM, se inicia el proceso de entrenamiento de acuerdo a la UX del sistema. La retroalimentación hace posible recapitular las experiencias de usuarios registradas, en las que se pueden determinar tanto sus preferencias como sus necesidades. Un concepto que implica adaptabilidad es la interfaz de usuario abstracta, y los modelos ML tienen la potencialidad para generar UI adaptativas. Por ejemplo, los *widgets* seleccionados pueden ser gestionados desde un sistema operativo mediante un algoritmo de ML.

Para finalizar esta sección, cabe mencionar que el aprendizaje adaptativo es propio del usuario, es decir, de su interacción y la retroalimentación, que está disponible en una capa de persistencia en el ARM, sobre la cual se puede realizar un proceso de ML para mejorar aspectos de navegabilidad y usabilidad.

3.5.2 *El prototipado de una interfaz de usuario y el reconocimiento de imágenes*

Buschek et al. (2020) plantean integrar ML en el proceso de diseño de interfaces y utilizan un *framework* denominado *Paper2Wire*, que tiene por objetivo integrar el proceso de maquetado y su conversión a la versión digital del diseño.

Moran et al. (2018) proponen un modelado basado en bosquejos en papel que puede implementar una UI. Los autores plantean usar un modelo de CNN o red neuronal convolucional que utiliza 8878 aplicaciones más populares de Google Play ejecutadas con *GUI-ripping* (generación de ejecuciones automáticas); luego se extraen las capturas de pantalla más populares con las que la CNN fue entrenada. En resumen, es una aplicación de ingeniería reversa. Los aspectos de análisis del correcto diseño de la UI usan experiencias basadas en adaptabilidad y usabilidad de sistemas interactivos. Este prototipado plantea tres fases para el diseño. La primera es la detección de componentes mediante análisis o parseo de diferentes formatos de archivos de imágenes o técnicas de visión computacional para reconocer los componentes GUI que están siendo utilizados en el maquetado de la UI; en esta etapa se establecen las jerarquías de los componentes. La siguiente fase es la clasificación

de componentes GUI detectados del maquetado en elementos específicos de GUI mediante CNN. La última etapa es el ensamblaje en una aplicación; después de determinar la correspondencia de los componentes GUI, estos son llevados a código fuente a manera de nodos de un árbol GUI. Al finalizar todo el proceso, se debe obtener una UI similar al modelado desarrollado en una maqueta.

3.5.3 La usabilidad y los dispositivos de pantallas pequeñas

Shackel y Richardson (1991) definen la usabilidad como la combinación de eficiencia, eficacia y satisfacción. En todo momento, la usabilidad es evaluada por el usuario cada vez que interactúa con el sistema.

La usabilidad entra en juego cuando el espacio para implementar los controles o componentes de la UI es muy limitado. Lee et al. (2018) plantean una alternativa para implementar una UI que cumpla con los requisitos de usabilidad, que consiste en utilizar mecanismos de detección de movimientos implementados en estos dispositivos. Cada movimiento realizado por el usuario es detectado y se le empareja una acción que debe realizar el sistema; es decir, se define navegabilidad y selección. Incluso se hace una detección de errores cuando un movimiento no es considerado como entrada para una acción del sistema. Una UI con este diseño plantea una interacción naturalista con el sistema. Ejemplos de estos dispositivos son los relojes inteligentes, reproductores de música o dispositivos enfocados en la detección de funciones biométricas, como detección de pulso o medición de ritmo cardíaco.

4. DISCUSIÓN

En primera instancia, se revisaron los enfoques actuales para el diseño de interfaces sin la utilización de ML como parte de su proceso; entre estos, se pueden destacar los modelos como la optimización combinatoria o la utilización de *frameworks* para el diseño de la UI. La siguiente propuesta plantea crear interfaces libres de texto, es decir, no utilizar como principal fuente de entradas para realizar la interacción con el sistema textos o cadenas de caracteres, pues el usuario puede tener limitaciones de idioma u otro tipo, tales como lectura o escritura, por lo que conviene enfocarse más en interacciones a través de símbolos e imágenes. El último enfoque revisado muestra el diseño como una línea de producción de *software* dinámico: mediante la combinación de los conceptos de MBUID y DSLP, los diseñadores son capaces de diseñar UI que pueden adaptarse al contexto del usuario, priorizando la reusabilidad y abstracción. Luego se plantean los enfoques en los cuales el ML es parte del proceso de diseño de la UI y la UX, como el registro de la UX para detectar patrones de comportamiento en la utilización del sistema. Este planteamiento podría ser la base para implementar sistemas interactivos en versiones superiores; es decir, cuando

un sistema es implementado por primera vez, es posible registrar o al menos tener una bitácora de las interacciones del usuario con el sistema. Esta acción nos permite determinar patrones y plantear una interfaz más adaptativa para el usuario en una siguiente versión. Ejemplos de esto son las *cookies* de las webs que nos indican preferencias o costumbres de navegación.

El maquetado, como base del diseño de interfaces a partir de repositorios de imágenes con los cuales se entrenarán algoritmos de CNN y sistemas como REDRAW en combinación con el algoritmo KNN, es capaz de crear prototipos de GUI para dispositivos móviles. La aplicación de este enfoque puede tener inconvenientes en el diseño, puesto que es muy dependiente de los especialistas en UX y UI. Un diseñador de muy poca experiencia podría omitir aspectos de usabilidad o funcionalidad, además de que un ML usando visión computacional o técnicas de parseo tiene efectividad limitada en algunos casos.

La navegabilidad y la interacción naturalistas mediante detección de movimiento en dispositivos de pantalla pequeña nos otorgan interacciones más fáciles. Otro aspecto de usabilidad muy importante, la navegabilidad y selección, se atiende con este enfoque; movimientos que realizan nuestras manos o alguna parte de nuestro cuerpo sirven como entradas para la interacción con el sistema.

La disponibilidad de un histórico de interacciones debería permitir la capacidad de implementar una interfaz más adaptativa al usuario, con la posibilidad de determinar posiciones de los controles o qué tipo de estos se pueden usar en una UI para que más se acomoden al usuario. Como menciona Pressman (2010) en el capítulo de interfaces, crear una interfaz que sea capaz de leer la mente del usuario.

5. CONCLUSIONES

Las propuestas que no usan ML en el proceso de diseño de UI plantean emplearlo en trabajos futuros. Las interacciones basadas en texto presentan limitaciones para muchos usuarios por el idioma o incluso el analfabetismo. La tecnología de ML aún no es un estándar para el diseño de interfaces, puesto que los diseñadores de UI y UX no proponen una metodología de aplicación en el diseño de interfaces de usuario. El maquetado debe considerar desde el principio del proceso la UX. Las técnicas de detección de imágenes como base para el diseño de interfaces generan buenas propuestas de GUI, siempre que el diseñador responsable del maquetado considere todas las características de usabilidad. Las interfaces más sofisticadas usan como medios de interacción con el sistema funciones corporales o biológicas, tales como movimiento de brazos, piernas, gestos e incluso emociones. La utilización de *machine learning* para aplicaciones, como distribución de elementos de GUI o selección de

elementos GUI más convenientes, colores o jerarquía de distribución de módulos más complejos, no se ha desarrollado como metodología de diseño.

REFERENCIAS

- Akinsola, J. E. T., Akinseinde, S., Kalesanwo, O., Adeagbo, M., Oladapo, K., Awoseyi, A., & Kasali, F. (2021). Application of artificial intelligence in user interfaces design for cyber security threat modeling. En L. M. Castro, D. Cabrero & R. Heimgärtner (Eds.), *Software usability*. IntechOpen. <https://www.intechopen.com/chapters/76094>
- Braun, M., Weber, F., & Alt, F. (2021). Affective automotive user interfaces. Reviewing the state of driver affect research and emotion regulation in the car. *ACM Computing Surveys (CSUR)*, 54(7), 1-26. <https://doi.org/10.1145/3460938>
- Buschek, D., Anlauff, C., & Lachner, F. (2020, September). Paper2Wire: A case study of user-centred development of machine learning tools for UX designers. En *Proceedings of the Conference on Mensch und Computer* (pp. 33-41). <https://dl.acm.org/doi/abs/10.1145/3404983.3405506>
- Cruz-Benito, J., Vázquez-Ingelmo, A., Sánchez-Prieto, J. C., Therón, R., García-Peñalvo, F. J., & Martín-González, M. (2017). Enabling adaptability in web forms based on user characteristics detection through A/B testing and machine learning. *IEEE Access*, 6, 2251-2265. <https://ieeexplore.ieee.org/abstract/document/8240912>
- Ferreira, J.M., Acuña, S.T., Dieste, O., Vegas, S., Santos, A., Rodríguez, F., & Juristo, N. (2020). Impact of usability mechanisms: An experiment on efficiency, effectiveness and user satisfaction. *Information and Software Technology*, 117, 106195. <http://jultika.oulu.fi/files/nbnfi-fe202002054561.pdf>
- Ilyas, Q. M., Ahmad, M., Zaman, N., Alshamari, M. A., & Ahmed, I. (2021). Localized text-free user interfaces. *IEEE Access*, 10, 2357-2371. <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9665747>
- Joo, M., & Lee, W. (2019). WebProfiler: User interaction prediction framework for web applications. *IEEE Access*, 7, 154946-154958. <https://ieeexplore.ieee.org/abstract/document/8880603>
- Lee, K. T., Yoon, H., & Lee, Y. S. (2018). Implementation of smartwatch user interface using machine learning based motion recognition. En *2018 International Conference on Information Networking (ICOIN)* (pp. 807-809). <http://csc.villanova.edu/~beck/csc8570/papers/lee.pdf>

- Mezhoudi, N. (2013). User interface adaptation based on user feedback and machine learning. En *Proceedings of the companion publication of the 2013 International Conference on Intelligent User Interfaces Companion* (pp. 25-28). https://dial.uclouvain.be/pr/boreal/object/boreal%3A153816/datastream/PDF_01/view
- Moran, K., Bernal-Cárdenas, C., Curcio, M., Bonett, R., & Poshyvanyk, D. (2018). Machine learning-based prototyping of graphical user interfaces for mobile apps. *IEEE Transactions on Software Engineering*, 46(2), 196-221. <https://ieeexplore.ieee.org/abstract/document/8374985>
- Oulasvirta, A., Dayama, N. R., Shiripour, M., John, M., & Karrenbauer, A. (2020). Combinatorial optimization of graphical user interface designs. *Proceedings of the IEEE*, 108(3), 434-464. <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9000519>
- Pressman, R. S. (2010). *Ingeniería del software. Un enfoque práctico*. McGraw-Hill.
- Sackl, M., Steinmaurer, A., Cheong, C., Cheong, F., Filippou, J., & Gütl, C. (2020). sCool: Impact on human-computer interface improvements on learner experience in a game-based learning platform. En *International Conference on Interactive Collaborative Learning* (pp. 439-451). Springer. https://www.researchgate.net/profile/Alexander-Steinmaurer-3/publication/350001425_sCool_Impact_on_Human-Computer_Interface_Improvements_on_Learner_Experience_in_a_Game-Based_Learning_Platform/links/606cb77092851c4f268661c4/sCool-Impact-on-Human-Computer-Interface-Improvements-on-Learner-Experience-in-a-Game-Based-Learning-Platform.pdf
- Sánchez, W. O. (2015). La usabilidad en ingeniería de *software*: definición y características. *Ing-novación*, 2, 7-21. <http://www.redicces.org.sv/jspui/bitstream/10972/1937/1/2.%20La%20usabilidad%20en%20Ingenieria%20de%20Software-%20definicion%20y%20caracteristicas.pdf>
- Sanctorum, A., & Signer, B. (2019). A unifying reference framework and model for adaptive distributed hybrid user interfaces. En *2019 13th International Conference on Research Challenges in Information Science (RCIS)* (pp. 1-6). <https://ieeexplore.ieee.org/abstract/document/8877048>
- Sboui, T., Ayed, M. B., & Alimi, A. M. (2017). A UI-DSPL approach for the development of context-adaptable user interfaces. *IEEE Access*, 6, 7066-7081. <https://ieeexplore.ieee.org/abstract/document/8194842>
- Shackel, B., & Richardson, S. J. (Eds.). (1991). *Human factors for informatics usability*. Cambridge University Press.

- Yang, B., Wei, L., & Pu, Z. (2020). Measuring and improving user experience through artificial intelligence-aided design. *Frontiers in Psychology, 11*, 595374. <https://www.frontiersin.org/articles/10.3389/fpsyg.2020.595374/full>
- Yang, Q. (2017). The role of design in creating machine-learning-enhanced user experience. En *The AAAI 2017 Spring Symposium on Designing the User Experience of Machine Learning Systems* (pp. 406-411). <https://www.aaai.org/ocs/index.php/SSS/SSS17/paper/viewFile/15363/14575>
- Yang, Q. (2018). *Machine learning as a UX design material: How can we imagine beyond automation, recommenders, and reminders?* AAAI Spring Symposia. https://www.researchgate.net/profile/Qian-Yang-19/publication/324077664_Machine_Learning_as_a_UX_Design_Material_How_Can_We_Imagine_Beyond_Automation_Recommenders_and_Reminders/links/5b3cfe2ca6fdcc8506f560e1/Machine-Learning-as-a-UX-Design-Material-How-Can-We-Imagine-Beyond-Automation-Recommenders-and-Reminders.pdf

DISPOSITIVOS *WEARABLES* Y LOS RIESGOS A LA PRIVACIDAD: UNA REVISIÓN DE LA LITERATURA

ANGELO RODRIGO TACO JIMENEZ
angelotacoj@gmail.com
<https://orcid.org/0000-0001-9806-5379>
Universidad de Lima, Perú

RESUMEN. En la actualidad, está en auge del uso de dispositivos que se pueden llevar puestos (*wearable devices*, en inglés): tanto niños como adultos han incorporado en su día a día este tipo de dispositivos adicionales al celular, que hoy son una extensión más del ser humano. Esta investigación hace una revisión de qué son los *wearables*, su uso y los riesgos que implican para el usuario en cuanto a su privacidad y su seguridad. Se eligió la metodología de identificación y control de riesgos entre los resultados de una revisión de alcance en la literatura pertinente.

PALABRAS CLAVE: *wearables*, riesgos, privacidad, seguridad, controles

WEARABLE DEVICES AND PRIVACY RISKS: A LITERATURE REVIEW

ABSTRACT. Currently, the use of wearable devices is booming: both children and adults have incorporated this type of device in addition to cell phones into their day-to-day lives, which today are another extension of the human being. This research reviews what wearable devices are, their use, and the risks they imply for the user in terms of privacy and security. The risk identification and control methodology was chosen from the results of a scoping review in the relevant literature.

KEYWORDS: wearable devices, risks, privacy, security, controls

1. INTRODUCCIÓN

Según Wells (2019), en el año 2018, cerca de 3,7 billones de dispositivos con tecnología *bluetooth* han sido distribuidos alrededor del mundo. Además, en años recientes, la producción de estos dispositivos ha aumentado notablemente y han crecido sus ventas desde el año 2015 hasta el año 2020 (Laricchia, 2022).

Los dispositivos *wearables* pueden favorecer e impulsar el crecimiento tecnológico; sin embargo, al no controlar este crecimiento exponencial, pueden surgir problemas en la garantía de estándares de calidad, seguridad y privacidad. A diferencia de otros tipos de dispositivos *wearables*, los actuales tienen la capacidad de recopilar información del usuario y, además, funcionan como portadores de internet de las cosas (IoT, por sus siglas en inglés), a través de la cual se interconectan los objetos físicos entre sí (Gubbi et al., 2013).

La transformación digital, en el año 2020, provocó que se generara poco más de un billón de megabytes de datos, y esta cantidad se duplica cada cierto tiempo (González, 2020). A ello se suma el negocio de la compra y venta de los datos personales de los usuarios, por lo que estos aparecen en numerosas bases de datos. Como consecuencia, existe un riesgo notable para la privacidad de los usuarios, debido a que en un momento dado no saben cómo su información llega a ser conocida por terceros.

Por lo anteriormente mencionado, el propósito del siguiente trabajo de investigación es aprender los conceptos relacionados con los dispositivos *wearables* y analizar los riesgos que implica el uso de este tipo de dispositivos, a fin de identificar medidas para controlarlos.

2. MARCO TEÓRICO

2.1 Dispositivos *wearables*

Estos dispositivos son utilizados por usuarios de manera externa, de ahí el término *wearable*, que en español significa "llevable". Permiten que el usuario pueda interactuar con sus registros personales y, a su vez, intercambiar o compartir datos como estos:

- Frecuencia cardíaca
- Hábitos de sueño
- Geolocalización
- Parámetros fisiológicos

- Parámetros clínicos
- Hábitos alimenticios
- Calendario

Este tipo de dispositivos tiene mayor exposición en el deporte o en el entretenimiento, y por ese motivo las empresas apuestan por su constante desarrollo y crecimiento, y buscan que usarlos sea lo más sencillo posible para el usuario. Además, las empresas desarrolladoras de aplicaciones también son un pilar importante, ya que al crear aplicaciones más intuitivas, accesibles en precio y óptimas, los usuarios hallan una motivación para llevar a cabo actividades físicas, pues pueden tener un control de su progreso.

2.2 Riesgo

La palabra *riesgo* tiene distintas definiciones que se adaptan a cada situación. La Organización Internacional de Normalización (ISO, 2015) la define como un efecto de la incertidumbre, mientras que para la ISO (2018) es el efecto de la incertidumbre sobre el logro de los objetivos. También es definida como la combinación de que se lleve a cabo un evento y sus posibles consecuencias positivas y negativas. Dentro del concepto de riesgo también encontramos dos factores importantes:

2.2.1 Amenaza

Es un fenómeno que potencialmente puede ocasionar daños y pérdidas. Se determina en función de su frecuencia e intensidad. Son ejemplos de amenazas:

- *Malware, ransomware, virus, etcétera*
- Fallas técnicas u operativas

2.2.2 Vulnerabilidad

Se refiere a las características y circunstancias de un sistema que pueden hacerlo susceptible a los efectos provocados por una amenaza latente.

Con los dos factores mencionados, amenaza y vulnerabilidad, se formula la siguiente ecuación:

$$\text{Riesgo} = \text{vulnerabilidad} \times \text{amenaza} \quad (1)$$

2.3 Privacidad

El derecho a la privacidad se ha visto afectado, como explica Laricchia (2022), debido a la creación de nuevas tecnologías y el constante desarrollo de internet, que

genera nuevos retos a las empresas en el ámbito de protección de datos personales de los usuarios. Según información recopilada por Argentina Cibersegura (2020), la privacidad se entiende como el control que ejerce un usuario respecto a su información y datos, para poder limitar el acceso a terceros o instituciones.

2.4 Gestión del riesgo

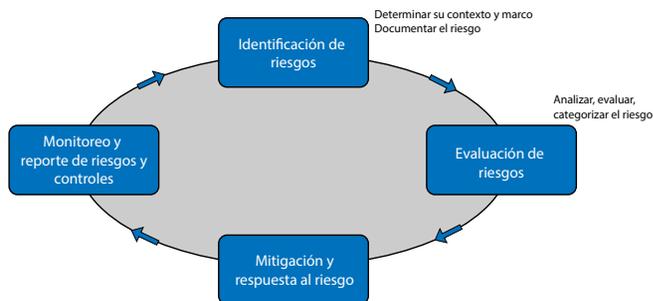
Es un conjunto de técnicas y herramientas que ayudan en la toma de las decisiones más apropiadas para las empresas o personas, según criterios que tengan en cuenta la incertidumbre, la posibilidad del futuro y los efectos de los objetivos. Su fin es la prevención de los riesgos, en lugar de mitigarlos o corregirlos, lo que es beneficioso para las empresas u organizaciones que pongan en uso la ISO 9000 (2015).

2.5 Ciclo de vida de la gestión del riesgo

Para definir el ciclo de vida de la gestión del riesgo, se debe tener en cuenta cuán importante es la toma de decisiones estratégicas, pues estas son, en muchas ocasiones, la causa principal de los desaciertos en escenarios de las empresas u organizaciones. En la Figura 1 se ilustra los cuatro elementos del ciclo de vida de la gestión del riesgo.

Figura 1

Ciclo de vida de la gestión del riesgo



Nota. Reproducido de *Certified in Risk and Information Systems Control (CRISC)*, por ISACA Madrid Chapter, 2010.

3. ANTECEDENTES

La presente investigación se centrará en el uso de una metodología para el descubrimiento y tratamiento de los riesgos para la privacidad y seguridad de los usuarios que acarrean los dispositivos *wearables*. Este tipo de dispositivos son

capaces de proporcionar servicios inteligentes, como la recopilación de información, procesamiento de datos y salida de información (Fernández-Caramés & Fraga-Lamas, 2018); pero el problema radica en la gran cantidad de datos de cada usuario en todo el mundo que puede estar disponible para terceros. Esto plantea una preocupación en la comunidad de investigadores de seguridad y privacidad.

Por tanto, se ha llevado a cabo una búsqueda y análisis de la literatura en bases de datos existentes que ahondan en las vulnerabilidades de la seguridad y de la privacidad que se encuentran en los dispositivos *wearables* y en sus aplicaciones.

En primer lugar, Cyr et al. (2014) hicieron un análisis de las características de la seguridad y de la privacidad de los datos de los usuarios en dispositivos de Fitbit, ahondando específicamente en las debilidades de la seguridad en estos, que usaban *bluetooth* y una aplicación de *smartphone* durante la sincronización del tráfico. El resultado arrojó que Fitbit recopilaba datos del usuario sin su consentimiento y, además, que la dirección MAC de estos dispositivos no cambiaba jamás, lo que facilitaba cualquier ataque de un tercero.

Luego, Seneviratne et al. (2017) realizaron un estudio y una clasificación de dispositivos *wearables* que están disponibles en el mercado; asimismo, analizaron las amenazas en términos de confidencialidad y la disponibilidad de la información que hay en estos dispositivos. Usando el BLE (*Bluetooth Low Energy*) como medio de comunicación, se concluye que existen hasta tres tipos de ataques a los que son vulnerables los dispositivos *wearables*:

- Espionaje
- Análisis de tráfico
- Compilación de información

También Hale et al. (2019) llevaron a cabo una investigación en la que analizaron tres dispositivos: Pebble, Fitbit y Jawbone. Detectaron que los tres expusieron su conexión, por lo que volvieron vulnerable al servidor y abrían la posibilidad de que un atacante pueda seguir la conexión luego de iniciar.

Finalmente, se revisó un informe de Wu y Luo (2019) acerca de la confidencialidad y la seguridad de los datos de los usuarios, el cual se realizó para garantizar el cumplimiento de la regulación de datos, por la vulnerabilidad sensible de dispositivos *wearables*. Los autores identificaron que terceros podrían tener acceso privilegiado a los dispositivos por descifrar la clave de acceso, debido a un débil sistema de seguridad. Por tanto, sugirieron que, siguiendo las pautas de la HIPAA (Health Insurance Portability and Accountbaility Act), la distribución de la clave de seguridad de los usuarios sea parte de un proceso adicional de autenticación más sólido, con

el fin de tener un estándar de mayor de calidad en privacidad y seguridad para los usuarios.

4. METODOLOGÍA

Se hizo una revisión de alcance para introducir la definición de riesgos y conceptos dependientes a fin de abordar el tema de la privacidad en dispositivos *wearables*. La revisión de alcance usualmente tiene el objetivo de examinar la literatura que esté disponible sin evaluaciones de calidad formales y, a su vez, ayuda a decidir si es necesario llevar a cabo una revisión sistemática (Grant & Booth, 2009; Peters et al., 2015). Según la propuesta de Arksey y O'Malley (2005) de pasos de revisión para la definición de alcance, el presente trabajo realizó búsquedas de literatura existente en tres bases de datos especializadas: IEEE, Web of Science y Scopus.

4.1 Planificación del estudio

En esta subsección se presentan las preguntas de relevancia para la revisión:

- ¿Cuáles son las definiciones de los dispositivos *wearables*?
- ¿Cuál es el riesgo de los dispositivos *wearables* para la privacidad de los usuarios?
- ¿Cuál es el riesgo de los dispositivos *wearables* para la seguridad de los usuarios?
- ¿Cómo identificar los riesgos de dispositivos *wearables*?

Tabla 1

Keywords y query de búsqueda

Bases de datos	Palabras clave	Cadena de búsqueda	Cantidad de artículos
IEEE	"risk", "wearables", "privacy"	TITLE-ABS-KEY (risk AND wearables AND privacy) AND (LIMIT-TO (PUBYEAR , 2022) OR LIMIT-TO (PUBYEAR , 2021) OR LIMIT-TO (PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR , 2018)) AND (LIMIT-TO (EXACTKEYWORD , "Data privacy") OR LIMIT-TO (EXACTKEYWORD , "Data Privacy") OR LIMIT-TO (EXACTKEYWORD , "Wearables computers") OR LIMIT-TO (EXACTKEYWORD , "Internet Of Things"))	8

(continúa)

(continuación)

Web of Science	"risk", "wearables", "privacy"	TITLE-ABS-KEY (risk AND wearables AND privacy) AND (LIMIT-TO (PUBYEAR , 2022) OR LIMIT-TO (PUBYEAR , 2021) OR LIMIT-TO (PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR , 2018)) AND (LIMIT-TO (DOCTYPE , "ar"))	6
Scopus	"user", "risk", "wearables", "privacy"	TITLE-ABS-KEY (risk AND wearables AND privacy AND user) AND (LIMIT-TO (PUBYEAR , 2022) OR LIMIT-TO (PUBYEAR , 2021) OR LIMIT-TO (PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR , 2018)) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (SUBJAREA , "COMP")) AND (LIMIT-TO (EXACTKEYWORD , "Wearable Technology") OR LIMIT-TO (EXACTKEYWORD , "Data Privacy") OR LIMIT-TO (EXACTKEYWORD , "Wearables") OR LIMIT-TO (EXACTKEYWORD , "Internet Of Things"))	10

Para la búsqueda de la literatura en las bases de datos, se usaron palabras clave con relación a dispositivos *wearables* y a riesgos, además de priorizar los artículos más recientes.

Se utilizaron distintos criterios de exclusión y de inclusión para seleccionar los artículos. El primer criterio de exclusión se refiere a los artículos relacionados con el ámbito de la salud y que estudian cómo los dispositivos *wearables* apoyan en el tratamiento o prevención de enfermedades, pero sí se incluye a aquellos que hacen una revisión de conceptos importantes para el desarrollo de la investigación. El segundo criterio comprende a los artículos que describen vulnerabilidades y amenazas del uso de dispositivos *wearables*, pero se excluye a aquellos que tenían información muy similar y que no aportaban lo suficiente. Finalmente, luego del proceso de exclusión e inclusión, fueron seleccionados diez artículos; cabe destacar que la revisión se limitó a publicaciones de artículos en español y en inglés.

Tabla 2*Literatura revisada para dar respuesta a las preguntas*

ID	Año	Nombre	Autores
01	2019	Threat Analysis for Wearable Health Devices and Environment Monitoring Internet of Things Integration System	Tzu Wei Tseng, Chia Tung Wu, Feipei Lai
02	2018	Data Transfusion: Pairing Wearable Devices and Its Implication on Security for Internet of Things	Youngjoo Lee, Wonseok Yang, Taekyoung Kwon

(continúa)

(continuación)

03	2019	"Worth One Minute": An Anonymous Rewarding Platform for Crowd-Sensing Systems	Lorenz Cuno Klopfenstein, Saverio Delpriori, Alessandro Aldini, Alessandro Bogliolo
04	2020	Preserving Privacy in Mobile Health Systems Using Non-Interactive Zero-Knowledge Proof and Blockchain	Antonio Emerson Barros Tomaz, José Cláudio do Nascimento, Abdelhakim Senhaji Hafid, José Neuman de Souza
05	2020	Security Analysis of IoT Devices by Using Mobile Computing: A Systematic Literature Review	Bin Liao, Yasir Ali, Shah Nazir, Long He, Habib Ullah Khan
06	2019	A Review on Intelligent Wearables: Uses and Risks	Yukang Xue
07	2021	Are Serious Games Too Serious? Diffusion of Wearable Technologies and the Creation of a Diffusion of Serious Games Model	Ton A. M. Spil, Vincent Romijnders, David Sundaram, Nilmini Wickramasinghe, Björn Kijl
08	2022	Análisis de problemas de seguridad y privacidad en <i>wearables</i> usados por menores	Sonia Solera-Cotanilla, Jaime Fúster, Jaime Pérez, Rafael Palacios, Mario Vega-Barbas, Manuel Álvarez-Campana, Gregorio López
09	2018	A Review on Smart Home Present State and Challenges: Linked to Context-Awareness Internet of Things (IoT)	Zahrah A. Almusaylim, Noor Zaman
10	2021	Systematically Quantifying IoT Privacy Leakage in Mobile Networks	Shuodi Hui, Zhenhua Wang, Xueshi Hou, Xiao Wang, Huandong Wang, Yong Li, Depeng Jin

5. RESULTADOS

5.1 ¿Cuáles son las definiciones de los dispositivos *wearables*?

Lee et al. (2018) definen los dispositivos *wearables* como dispositivos informáticos portátiles; los describen como productos electrónicos en miniatura que son usados por los usuarios y, según los autores, estos representan la siguiente era de la internet de las cosas (IoT, por sus siglas en inglés). Asimismo, aclaran que estos dispositivos poseen una interfaz de usuario particular por sus condiciones y que su tipo de conectividad es ágil.

Según Klopfenstein et al. (2019), estos dispositivos son un poderoso instrumento con capacidad de detección, poder de cómputo y medio de comunicación, que además ayuda en la salud y el estado físico; igualmente, cada uno de estos dispositivos tiene la capacidad de recopilar, procesar y transmitir datos. Tomaz et al. (2020) los definen como accesorios tecnológicos vestibles o implantables usados para el monitoreo de salud de un paciente y que permiten ofrecer servicio médico a distancia. Para Xue (2019), los dispositivos *wearables* se pueden usar en prendas o como accesorio

personal adherido al cuerpo y son capaces de recopilar información, realizar procesamiento de datos y salida de información procesada.

5.2 ¿Cuál es el riesgo de los dispositivos *wearables* para la privacidad de los usuarios?

El artículo publicado por Tomaz et al. (2020) corresponde a una investigación en el ámbito de la salud, y hace una crítica a las empresas creadoras y distribuidoras de dispositivos *wearables* que no se interesan por la privacidad de sus usuarios; se asume por las empresas que los datos recibidos por los teléfonos inteligentes no son manipulados y que su origen no representa un riesgo. La propuesta de Tomaz et al. (2020) es la creación de una asociación exclusiva entre un dispositivo *wearable* y una aplicación de salud; el factor diferenciador se basa en la no dependencia de terceras empresas que se encarguen del almacenamiento de datos de pacientes y que se integre una arquitectura con *blockchain*/IPFS, donde los datos privados de pacientes solo puedan ser administrados por ellos mismos.

Un estudio realizado por Spil et al. (2020) realizó una encuesta a 97 usuarios respecto a la difusión de los juegos serios y cómo impacta en ellos. De esta encuesta, se rescata que solo el 40 % de la población entrevistada mostró interés y preocupación por su privacidad frente al manejo de información de dispositivos *wearables* compatibles con los juegos serios.

Por su parte, Lee et al. (2018) acuñan el término *data transfusion* para referirse a la información que se comparte de un dispositivo móvil a un reloj inteligente cuando se emparejan. Comentan que es inevitable la transferencia de datos cuando este hecho ocurre, ya que es necesario para la inicialización del dispositivo. Lo preocupante es que los usuarios no tienen conocimiento de qué tipo de datos son los que se comparten. Concluyen que, en comparación con los dispositivos móviles, los usuarios no toman conciencia del riesgo que conlleva el desconocimiento de datos compartidos y de que esto es un problema para su privacidad.

5.3 ¿Cuál es el riesgo de los dispositivos *wearables* para la seguridad de los usuarios?

Los dispositivos *wearables* están expuestos, sin excepción, a distintos tipos de problemas de seguridad, los cuales se componen de amenazas, vulnerabilidades y ataques. De acuerdo con Girma (2018), la relación entre la comunicación y la conectividad es la amenaza principal para la seguridad de los usuarios y, en consecuencia, provoca una preocupación en ellos. Por su parte, Liao et al. (2020) señalan que los dispositivos *wearables* son muy susceptibles a sufrir amenazas en su seguridad debido a su bajo costo de producción, por lo que tienen bajo rendimiento

y potencia en comparación con teléfonos móviles y computadoras de escritorio. En el ámbito de IoT, donde trabajan los dispositivos *wearables*, los datos generados por el uso del usuario se recopilan y procesan la mayoría de veces, a excepción de algunas compañías, por terceras empresas, sin que el usuario tenga conocimiento de ello (Almusaylim & Zaman, 2018); así, un atacante puede recibir la comunicación y extraer contenido sin cifrar que se encuentre disponible en el tráfico (Dziubinski & Bandai, 2020).

En la literatura revisada, Hui et al. (2021) reflexionan acerca de la seguridad de los datos de los usuarios y cómo estos son manipulados y compartidos entre proveedores de servicios y terceros maliciosos, lo que da como resultado un monitoreo ilegal, riesgos financieros o amenazas personales. Por ello, proponen un método para cuantificar sistemáticamente la fuga de privacidad de IoT en el tráfico de red móvil; esto implica que los datos de los usuarios están expuestos y son de fácil acceso para cualquier tercero malicioso en el tráfico de red. Con base en su marco de trabajo propuesto, analizan estos datos y presentan estudios de caso que concluyen que los dispositivos *wearables* tienen mayor escala de fuga que cualquier otra entidad de red. Finalmente, se recomienda a usuarios y empresas desarrolladoras que no dejen pasar por alto la importancia de su seguridad en el mundo real, pues cualquier tipo de filtración a pequeña escala puede desencadenar riesgos de suma importancia para los usuarios y sus intereses.

5.4 ¿Cómo identificar riesgos de dispositivos *wearables*?

Tseng et al. (2019) mencionan en su investigación distintos métodos de prueba, como el escaneo de puertos y el escaneo de vulnerabilidades de seguridad y privacidad con herramientas especializadas, por ejemplo, Wireshark (2022), Nmap (2022) y OpenVAS (2022). Adicionalmente, especifican que los dispositivos IoT se clasifican de dos modos: *bluetooth* y *wifi*; y que en estos se implementan diferentes protocolos. La literatura de riesgos revisada por Tseng et al. (2019) resalta que la evaluación de riesgos se clasifica en dos tipos: cuantitativa y cualitativa. Ahonda en los métodos de evaluación para identificar la probabilidad e impacto de un riesgo, así como en estudios de evaluación de riesgos que se basan en el modelo STRIDE, que identifica seis amenazas: (i) *spoofing*, (ii) *tampering*, (iii) *repudiation*, (iv) *information disclosure*, (v) *denial of service* y (vi) *elevation of privilege*.

Este método se relaciona con el método DREAD (Altawy & Youssef, 2016), que se usa para calcular la relación de riesgos usando preguntas base para valorar las amenazas como bajas, medias, altas y críticas. Se presenta, a continuación, el método DREAD:

- *Damage*: ¿cuánto daño puede hacer?

- *Reproducibility*: ¿qué tan difícil es realizar el ataque?
- *Exploitability*: ¿qué es necesario para ejecutar un ataque?
- *Affected users*: ¿cuántas personas pueden verse afectadas?
- *Discoverability*: ¿qué tan difícil es descubrir vulnerabilidades?

La literatura presenta diversos casos donde existen potenciales riesgos para las personas. Por ejemplo, para los niños que necesitan supervisión de sus padres o apoderados, hay aplicaciones para supervisar su navegación y limitar el uso de distintas funcionalidades incorporadas en sus dispositivos. Cabe destacar que estas aplicaciones requieren permisos especiales para su correcto funcionamiento, lo que da pie a una nueva preocupación acerca de la privacidad de los niños por los datos expuestos en estas aplicaciones. El estudio de Feal et al. (2020) comparó 46 aplicaciones de este rubro en Google Play Store y demostró que el 34 % de ellas recopilan y envían información personal sin consentimiento del usuario, mientras que el 72 % envía los datos con terceros, todo ello sin informar al usuario en sus términos y condiciones. Por eso, luego de hacer la experimentación, se concluye que no son recomendables por su falta de claridad con los usuarios.

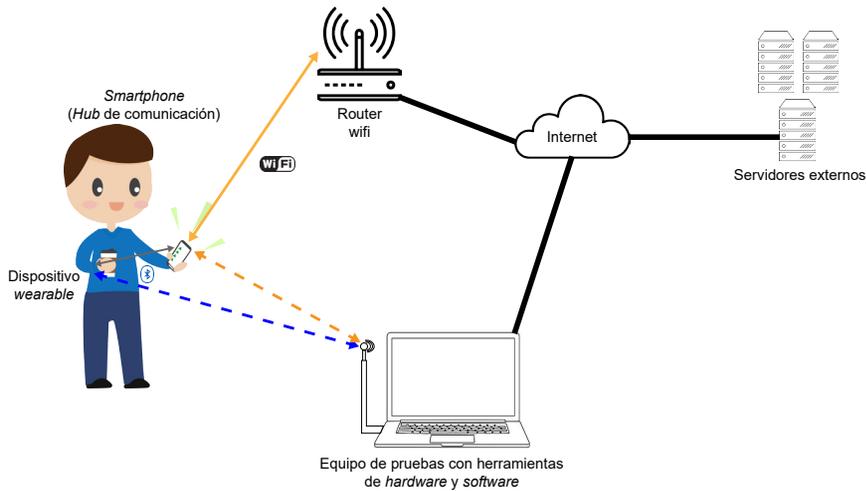
Adicionalmente, con respecto al entorno de Android, Reardon et al. (2019) demostraron que existen aplicaciones de Google Play Store que, con permisos específicos y con la suposición de que el usuario utilice una tarjeta SD (*Secure Digital*), podrían ser utilizadas por un tercero y funcionar como un canal encubierto para compartir el código IMEI del dispositivo móvil.

Aunque los problemas de privacidad parecen ser motivo de gran preocupación, muchas personas no toman medidas o precauciones específicas para mejorar la privacidad (Udoh & Alkharashi, 2016).

Existen diversos aportes metodológicos para la identificación de vulnerabilidades que, en consecuencia, determinan los potenciales riesgos. Por ejemplo, Solera-Cotanilla et al. (2022) proponen una metodología para analizar la seguridad y la privacidad de dispositivos *wearables*; su investigación pone a prueba dispositivos de distintas gamas y características técnicas con el fin de crear conciencia entre fabricantes y usuarios acerca del estado de seguridad y privacidad de estos dispositivos. El esquema de comunicación comúnmente utilizado por los dispositivos *wearables* actuales se muestra en la Figura 2 (Solera-Cotanilla et al., 2022). Un elemento con mayor capacidad de computación (por ejemplo, un *smartphone*) hace de intermediario (*hub*, configurador, etcétera) entre el *wearable* y los servidores externos. La tecnología de comunicación más común entre estos dispositivos es BLE.

Figura 2

Escenario de pruebas



Nota. Reproducido de "Análisis de problemas de seguridad y privacidad en wearables usados por menores" (p. 211), por S. Solera-Cotanilla et al., 2022, en *Actas de las VII Jornadas Nacionales de Investigación en Ciberseguridad (JNIC 2022)*.

En ese sentido, en el escenario se observan tres áreas de comunicación potenciales de análisis: (i) la primera, centrada en la interacción usuario-dispositivo conectado; (ii) la segunda, entre el *wearable* y el *hub* de comunicación; (iii) y la tercera, entre el *hub* de comunicación y los servidores externos o aplicaciones de terceros. En este trabajo, no analizaremos el caso de las conexiones móviles *Long-Term Evolution* (LTE) entre el *hub* de comunicación (por ejemplo, un *smartphone*) y los servidores externos.

6. CONCLUSIONES

Se ha hecho una revisión de los artículos más relevantes para el fin de esta investigación, el cual es conocer qué y cómo se usan los dispositivos *wearables*, así como los riesgos a los que están expuestos los usuarios, tanto niños como adultos, y cómo pueden verse afectados. Al conocer las vulnerabilidades de los principales dispositivos *wearables* mediante metodología de riesgos, se pueden reconocer los riesgos y demás. En futuros trabajos se podrían proponer nuevas metodologías y alternativas a la gestión de riesgos.

REFERENCIAS

- Almusaylim, Z. A., & Zaman, N. (2018). A review on smart home present state and challenges: Linked to context-awareness Internet of Things (IoT). *Wireless Network*, 25, 3193-3204. <https://doi.org/10.1007/s11276-018-1712-5>
- Altawy, R., & Youssef, A. M. (2016). Security tradeoffs in cyber physical systems: A case study survey on implantable medical devices. *IEEE Access*, 4, 959-979. <https://doi.org/10.1109/access.2016.2521727>
- Argentina Cibersegura. (2020). *¿Qué es la privacidad en internet?* https://www.argentinacibersegura.org/admin/resources/files/consejos/33/Gu%C3%ADa_sobre_Privacidad.pdf
- Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, 8(1), 19-32. <https://doi.org/10.1080/1364557032000119616>
- Cyr, B., Horn, W., Miao, D., & Specter, M. A. (2014). Security analysis of wearable fitness devices (Fitbit). *Massachusetts Institute of Technology*, 1. <https://media.kasperskycontenthub.com/wp-content/uploads/sites/43/2015/03/20082016/17-cyrbritt-webbhorn-specter-dmiao-hacking-fitbit.pdf>
- Dziubinski, K., & Bandai, M. (2020). Your neighbor knows what you're doing: Defending smart home IoT device traffic from privacy LAN attacks. En L. Barolli, F. Amato, F. Moscato, T. Enokido & M. Takizawa (Eds.), *Web, Artificial Intelligence and Network Applications. WAINA 2020* (pp. 526-534). Springer. https://doi.org/10.1007/978-3-030-44038-1_48
- Feal, Á., Calciati, P., Vallina-Rodriguez, N., Troncoso, C., & Gorla, A. (2020). Angel or devil? A privacy study of mobile parental control apps. *Proceedings on Privacy Enhancing Technologies*, 2, 314-335. <https://doi.org/10.2478/popets-2020-0029>
- Fernández-Caramés, T., & Fraga-Lamas, P. (2018). Towards the internet of smart clothing: A review on IoT wearables and garments for creating intelligent connected e-textiles. *Electronics*, 7(12), 405. <https://doi.org/10.3390/electronics7120405>
- Girma, A. (2018). Analysis of security vulnerability and analytics of Internet of Things (IOT) platform. En S. Latifi (Ed.), *Information Technology - New Generations. 15th International Conference on Information Technology* (pp. 101-104). Springer. https://doi.org/10.1007/978-3-319-77028-4_16
- González, C. A. (2020). *Introducción a la minería de datos*. Universidad de Valladolid.

- Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, 26(2), 91-108. <https://doi.org/10.1111/j.1471-1842.2009.00848.x>
- Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems*, 29(7), 1645-1660. <https://doi.org/10.1016/j.future.2013.01.010>
- Hale, M. L., Lotfy, K., Gamble, R. F., Walter, C., & Lin, J. (2019). Developing a platform to evaluate and assess the security of wearable devices. *Digital Communications and Networks*, 5(3), 147-159. <https://doi.org/10.1016/j.dcan.2018.10.009>
- Hui, S., Wang, Z., Hou, X., Wang, X., Wang, H., Li, Y., & Jin, D. (2021). Systematically quantifying IoT privacy leakage in mobile networks. *IEEE Internet of Things Journal*, 8(9), 7115-7125. DOI: 10.1109/JIOT.2020.3038639
- ISACA Madrid Chapter. (2010). *Certified in Risk and Information Systems Control (CRISC)*.
- Klopfenstein, L. C., Delpriori, S., Aldini, A., & Bogliolo, A. (2019). "Worth one minute": An anonymous rewarding platform for crowd-sensing systems. *Journal of Communications and Networks*, 21(5), 509-520. <https://doi.org/10.1109/jcn.2019.000051>
- Laricchia, F. (2022). *Wearables unit shipments worldwide by vendor from 1st quarter 2014 to 4th quarter 2021*. Statista.
- Lee, Y., Yang, W., & Kwon, T. (2018). Data transfusion: Pairing wearable devices and its implication on security for Internet of Things. *IEEE Access*, 6, 48994-49006. <https://doi.org/10.1109/access.2018.2859046>
- Liao, B., Ali, Y., Nazir, S., He, L., & Khan, H. U. (2020). Security analysis of IoT devices by using mobile computing: A systematic literature review. *IEEE Access*, 8, 120331-120350. <https://doi.org/10.1109/access.2020.3006358>
- Nmap. (2022). *Free security scanner for network exploration and security audits*. <https://nmap.org/>
- OpenVAS. (2022). *A framework for vulnerability scanning and vulnerability management*. <http://openvas.org/>
- Organización Internacional de Normalización. (2015). *Sistemas de gestión de calidad (ISO 9000:2015)*. <https://www.iso.org/obp/ui/es/#iso:std:iso:9000:ed-4:v1:es>
- Organización Internacional de Normalización. (2018). *Gestión del riesgo (ISO 31000:2018)*. <https://www.iso.org/obp/ui#iso:std:iso:31000:ed-2:v1:es>

- Peters, M. D., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., & Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. *International Journal of Evidence-Based Healthcare*, 13(3), 141-146. <https://doi.org/10.1097/XEB.0000000000000050>
- Reardon, J., Feal, Á., Wijesekera, P., Elazari, A., Vallina-Rodriguez, N., & Egelman, S. (2019). *50 ways to leak your data: An exploration of apps' circumvention of the android permissions system*. USENIX. <https://www.usenix.org/conference/usenixsecurity19/presentation/reardon>
- Seneviratne, S., Hu, Y., Nguyen, T., Lan, G., Khalifa, S., Thilakarathna, K., ... Seneviratne, A. (2017). A survey of wearable devices and challenges. *IEEE Communications Surveys & Tutorials*, 19(4), 2573-2620. <https://doi.org/10.1109/comst.2017.2731979>
- Solera-Cotanilla, S., Fúster, J., Pérez, J., Palacios, R., Vega-Barbas, M., Álvarez-Campana, M., & López, G. (2022). Análisis de problemas de seguridad y privacidad en *wearables* usados por menores. En *Actas de las VII Jornadas Nacionales de Investigación en Ciberseguridad (JNIC 2022)* (pp. 209-215). https://www.researchgate.net/publication/361972187_Analisis_de_Problemas_de_Seguridad_y_Privacidad_en_Wearables_Usados_por_Menores
- Spil, T. A. M., Romijnders, V., Sundaram, D., Wickramasinghe, N., & Kijl, B. (2021). Are serious games too serious? Diffusion of wearable technologies and the creation of a diffusion of serious games model. *International Journal of Information Management*, 58. <https://doi.org/10.1016/j.ijinfomgt.2020.102202>
- Tomaz, A. E. B., Nascimento, J. C. D., Hafid, A. S., & De Souza, J. N. (2020). Preserving privacy in mobile health systems using non-interactive zero-knowledge proof and blockchain. *IEEE Access*, 8, 204441-204458. <https://doi.org/10.1109/access.2020.3036811>
- Torre, I., Sanchez, O. R., Koceva, F., & Adorni, G. (2018). Supporting users to take informed decisions on privacy settings of personal devices. *Personal and Ubiquitous Computing*, 22, 345-364. <https://doi.org/10.1007/s00779-017-1068-3>
- Tseng, T. W., Wu, C. T., & Lai, F. (2019). Threat analysis for wearable health devices and environment monitoring Internet of Things integration system. *IEEE Access*, 7, 144983-144994. <https://doi.org/10.1109/access.2019.2946081>
- Udoh, E. S., & Alkharashi, A. (2016). Privacy risk awareness and the behavior of smartwatch users: A case study of Indiana University students. En *2016 Future Technologies Conference (FTC)* (pp. 926-931). <https://doi.org/10.1109/ftc.2016.7821714>

- Wireshark. (2022). *A network protocol analyzer*. <https://www.wireshark.org/>
- Wells, S. (2019, 17 de julio). *How Fitbits, other bluetooth devices make us vulnerable to tracking*. The Brink. <https://www.bu.edu/articles/2019/fitbit-bluetooth-vulnerability/>
- Wu, M., & Luo, J. (2019). *Wearable technology applications in healthcare: A literature review*. HIMSS. <https://www.himss.org/resources/wearable-technology-applications-healthcare-literature-review>
- Xue, Y. (2019). A review on intelligent wearables: Uses and risks. *Human Behavior and Emerging Technologies*, 1(4), 287-294. <https://doi.org/10.1002/hbe2.173>

Recibido: 5 de julio del 2022 / Aceptado: 8 de octubre del 2022
doi: <https://doi.org/10.26439/interfases2022.n016.6121>

UNA REVISIÓN DE LAS IMPLEMENTACIONES DE SISTEMAS PARA LA IDENTIFICACIÓN DE TENDENCIAS DE LA DIABETES

ROCIO ISABEL BENITES LOJA
2018100484@ucss.edu.pe
<https://orcid.org/0000-0002-9671-554X>
Universidad Católica Sedes Sapientiae, Perú

MARCO ANTONIO CORAL YGNACIO
mcoral@ucss.edu.pe
<https://orcid.org/0000-0001-6628-1528>
Universidad Católica Sedes Sapientiae, Perú

RESUMEN. La diabetes *mellitus* es una enfermedad crónica que aparece cuando el páncreas no secreta suficiente insulina o cuando el organismo no utiliza apropiadamente la insulina que produce. Dado que la insulina es una hormona que regula la concentración de glucosa en la sangre, uno de los efectos más comunes de la diabetes no controlada es la hiperglucemia, que con el tiempo daña gravemente muchos órganos y sistemas del cuerpo. Por ello, es importante el desarrollo de *software* predictivo para el diagnóstico y posterior tratamiento de esta enfermedad, en particular para la diabetes tipo 1 y 2, que concentran la mayoría de los casos. El presente trabajo realiza una revisión sistemática de literatura a fin de determinar los métodos y la problemática en la construcción de sistemas de identificación de tendencias orientados a la diabetes. Los resultados muestran 16 métodos diferentes de construcción utilizados en estos sistemas, de los cuales los más eficientes son la regresión logística y las redes neuronales artificiales.

PALABRAS CLAVE: diabetes *mellitus*, identificación de tendencias, *software* preventivo, métodos de construcción, regresión logística, redes neuronales artificiales

A REVIEW OF SYSTEM IMPLEMENTATIONS FOR DIABETES TREND IDENTIFICATION

ABSTRACT. Diabetes mellitus is a chronic disease that appears when the pancreas does not secrete enough insulin or the body does not properly use the insulin it produces. Insulin is a hormone that regulates glucose concentration in the blood: one of the most common effects of uncontrolled diabetes is hyperglycemia, which seriously damages many organs and body systems over time. In this sense, the development of predictive software, the diagnosis, and subsequent treatment of diabetes, especially of type 1 and 2, which are the most frequent, deserve attention. This paper presents a systematic review of the literature to determine the methods and problems in constructing diabetes-oriented trend identification systems. The results show 16 construction methods used in these systems, the most efficient being logistic regression and artificial neural networks.

KEYWORDS: diabetes mellitus, trend identification, preventive software, construction methods, logistic regression, artificial neural networks

1. INTRODUCCIÓN

La diabetes es un trastorno metabólico considerado una enfermedad crónica o de larga duración que se caracteriza por niveles elevados de glucosa en la sangre durante un extenso periodo. Las complicaciones graves asociadas con la diabetes son, a corto plazo, la cetoacidosis diabética y, a largo plazo, complicaciones como enfermedades cardiovasculares, accidentes cerebrovasculares, insuficiencia renal crónica y otras (Jaiswal et al., 2021). Según estadísticas mundiales brindadas por la Organización Mundial de la Salud, se prevé que para el 2045 el número de pacientes con diabetes *mellitus* alcanzará los 629 millones (Felizardo et al., 2021), lo cual la convierte en un problema de salud mundial.

En este contexto, se hace necesario predecir quiénes son pacientes potenciales que podrían estar en riesgo de adquirir diabetes. Con esto se busca evitar el desarrollo de dicha enfermedad y, en caso de darse, poder aplicar el tratamiento adecuado para prevenir su progresión a gran escala (Jaiswal et al., 2021). Por ello, el sector salud promueve la implementación de sistemas de apoyo al diagnóstico y prevención de la diabetes con el objetivo de reducir sus efectos, utilizando dispositivos inteligentes y sensores (Ray & Chaudhuri, 2021) e identificación de modelos predictivos mediante métodos computacionales asociados a la inteligencia artificial (Castelyn et al., 2021b), la cual tiene el potencial de revolucionar el cuidado de la salud debido a la disponibilidad de métodos y plataformas computacionales que son cada vez más poderosas, junto con las diversas fuentes informativas de datos que se tiene de cada paciente, tanto dentro como fuera de los entornos clínicos (Gautier et al., 2021).

En el campo de la inteligencia artificial existen algoritmos de aprendizaje automático, como las redes neuronales artificiales (ANN), las máquinas de soporte vectorial (SVM), *naive Bayes* (NBC), aprendizaje profundo (DL) y técnicas de minería de datos, que se utilizan para detectar patrones y determinar el diagnóstico y tratamiento de enfermedades (Jaiswal et al., 2021). La literatura especializada muestra pocas evidencias de implementaciones de sistemas de identificación de tendencias para la diabetes; por ello, es importante investigar las diferentes formas de implementación de un sistema de detección de pacientes potenciales y predicción de glucosa en sangre para pacientes con diabetes *mellitus* tipo 1 y 2. Se propone realizar, entonces, una revisión sistemática de literatura con el fin de identificar los métodos de construcción utilizados en las implementaciones de sistemas de predicción de la diabetes a fin de dar respuestas a los objetivos planteados en esta investigación. La revisión se basa en el método de Hanna Snyder, que considera las siguientes fases: diseño de la revisión, realización de la revisión, análisis y redacción de la reseña (Snyder, 2019). Se desarrolla una estrategia de búsqueda para realizar el análisis de artículos científicos en las siguientes bases de datos, por ser de acceso

libre, por la facilidad de búsqueda y la descarga de documentos: ScienceDirect, PubMed e IEEE Xplore. Los resultados de la revisión identificaron 16 métodos diferentes de construcción, 2 métodos considerados muy eficientes y 7 problemas en la implementación de estos sistemas.

El trabajo está organizado de la siguiente manera: después de la introducción en el punto 1, el punto 2 muestra la revisión de literatura; el punto 3 describe la metodología de la revisión, seguida por la presentación de los resultados en el punto 4; el punto 5 detalla los hallazgos, y las conclusiones se encuentran en el punto 6.

2. REVISIÓN DE LITERATURA

2.1 Diabetes *mellitus*

La diabetes *mellitus* (DM) es un trastorno metabólico que resulta de un defecto en la secreción de insulina y provoca el aumento del nivel de glucosa en la sangre, lo que afecta la forma en que el cuerpo convierte los alimentos en energía. Es una enfermedad que se caracteriza por hiperglucemia crónica y se asocia con una fuerte carga en la salud microvascular (retinopatía, nefropatía y neuropatía), así como con complicaciones macrovasculares (cardiovasculares, enfermedad vascular periférica y cerebrovascular) (Peer et al., 2020).

La DM tipo 1 afecta principalmente a las células beta del páncreas, lo que lleva a una deficiencia de insulina e hiperglucemia; por ello, las personas con este tipo de diabetes dependen de la insulina, la cual debe ajustarse según los alimentos que consumen, la actividad física, el estrés y la enfermedad de por sí (Pease et al., 2019). La DM tipo 2 hace que las células del cuerpo no interactúen con la insulina o que las células pancreáticas no produzcan suficiente insulina para regular los niveles de glucosa en la sangre, lo que debilita el metabolismo de los carbohidratos, las grasas y las proteínas, dando como resultado la DM tipo 1 en caso de no haber tenido un buen control de la DM tipo 2 (Ahmed et al., 2022).

2.2 Aplicación de la inteligencia artificial en la diabetes

En la medicina, la inteligencia artificial (IA) ha hecho una gran contribución en la vida humana, gracias a la disponibilidad de diversas plataformas médicas y métodos computacionales (Gautier et al., 2021) basados en aprendizaje automático (ML) y aprendizaje profundo. La diabetes se ha convertido en una de las principales enfermedades que han sido objeto de investigaciones médicas (Kodama et al., 2022) cuya atención está puesta en el aprendizaje automático, que permite desarrollar modelos predictivos (Tuppada & Patil, 2022) y herramientas impulsadas por IA para la

detección de enfermedades (Ahmed et al., 2022) y el control de las complicaciones (De Silva, Lee et al., 2020). Una de las herramientas más utilizadas es la técnica de ML, la cual permite que los sistemas inteligentes construyan modelos apropiados mediante el aprendizaje y la extracción de patrones en los datos (Li et al., 2020), pues es necesario que un conjunto de datos apropiados tenga las características necesarias para el respectivo entrenamiento y validación (Ahmed et al., 2022). Entre las aplicaciones se encuentran sistemas para la predicción de la diabetes tipo 2, que utilizan el promedio ponderado objetivo basado en la distancia, cuyos resultados son prometedores debido a su método de predicción novedoso (Nuankaew et al., 2021). Otras aplicaciones con mejores resultados emplean más de una técnica de aprendizaje automático, integrando la lógica difusa como método de construcción (Ahmed et al., 2022).

3. METODOLOGÍA

La revisión de literatura se lleva a cabo con el propósito de investigar y sintetizar evidencia de los principales métodos que se utilizan en la construcción de sistemas de predicción para la diabetes. En el desarrollo de la investigación, se consideran las siguientes fases (Snyder, 2019): diseño de la revisión, realización de la revisión y análisis.

3.1 Diseño de la revisión

Para realizar la búsqueda de los artículos, se define el tema con la finalidad de delimitar el alcance de la investigación: sistemas de identificación de tendencias para la diabetes. Posteriormente, se determinan los siguientes objetivos en relación con las preguntas de investigación:

- P1: ¿qué métodos de construcción son utilizados para la construcción de sistemas de identificación de tendencias relacionadas con la diabetes?
O1: identificar métodos de construcción que se emplean para la construcción de sistemas de identificación de tendencias relacionadas con la diabetes.
- P2: ¿qué métodos de construcción son los más eficientes para la construcción de un sistema de identificación de tendencias para la diabetes?
O2: identificar los métodos de construcción más eficientes para la construcción de un sistema de identificación de tendencias para la diabetes.
- P3: ¿qué modelos de diagnóstico son empleados para la detección de la diabetes?
O3: identificar los modelos de diagnóstico que son empleados para la detección de la diabetes.

- P4: ¿qué problemas se han encontrado para construir un sistema de identificación de tendencias relacionadas con la diabetes?

O4: determinar cuáles han sido los problemas existentes en la construcción de sistemas de identificación de tendencias relacionadas con la diabetes.

3.2 Estrategia de búsqueda

Los criterios establecidos para la búsqueda y selección de los artículos que se tuvieron en cuenta son los que aparecen en la Tabla 1.

Tabla 1

Criterios para la búsqueda sistemática

Criterios de inclusión	Criterios de exclusión
Artículos cuyos temas de investigación están orientados a la detección de la diabetes.	Artículos que mencionen síntomas no relacionados con la diabetes.
Artículos que mencionen métodos, técnicas y/o algoritmos para la construcción de un sistema predictivo para la diabetes.	Artículos no indexados.
Artículos que guarden relación con la enfermedad de la diabetes.	Informes de casos, estudios de casos, boletines, comentarios.
Artículos publicados del 2019 hasta el 2022.	
Artículos indexados y que tengan DOI.	
Artículos escritos en inglés.	

Posteriormente se identificaron las fuentes o bases de datos bibliográficas, tales como ScienceDirect, PubMed e IEEE Xplore, para ubicar los artículos relevantes a las preguntas de investigación formuladas. Además, se determinó una selección de palabras clave, que fueron las siguientes: *models, methods, algorithms, predictions, trends, diabetes, diagnosis*, la cual sirvió para establecer la siguiente cadena de búsqueda: “diabetes” and (predictions or trends) and (Models or Methods or Algorithms) and diagnosis.

3.3 Selección de estudios

De acuerdo con los criterios definidos para la búsqueda y selección de artículos, en la primera etapa de identificación se encontraron 7329 artículos en relación con la cadena de búsqueda de las diferentes fuentes bibliográficas. En la segunda etapa

de cribado, se revisaron los artículos según los criterios de inclusión y exclusión ya señalados, y se obtuvieron 446 estudios. Finalmente, en la tercera etapa, los artículos se evaluaron y se volvieron a seleccionar de acuerdo con los criterios de elegibilidad a partir del título, las palabras clave y el resumen, considerando solo los que tengan que ver con el tema de estudio. En primera instancia, se seleccionaron 14 artículos de ScienceDirect; de PubMed se obtuvieron 150, pero se eliminaron 123 duplicados y se tienen 27 estudios; de IEEE Xplore, se obtuvieron 80 estudios, de los cuales se eliminaron 70 duplicados, por lo que quedó un total de 10 trabajos. Se concluye la selección de 50 artículos que tratan sobre la construcción de sistemas de identificación de tendencias relacionados con la diabetes.

Tabla 2

Sintaxis de selección de estudios

Base de datos bibliográfica	Etapa 1: identificación	Etapa 2: cribado	Etapa 3: elegibilidad	Artículos eliminados	Estudios seleccionados
ScienceDirect	5487	80	14	0	14
PubMed	531	249	150	123	27
IEEE Xplore	1311	117	80	70	10
Total	7329	446	244	193	50

3.4 Extracción y análisis de datos

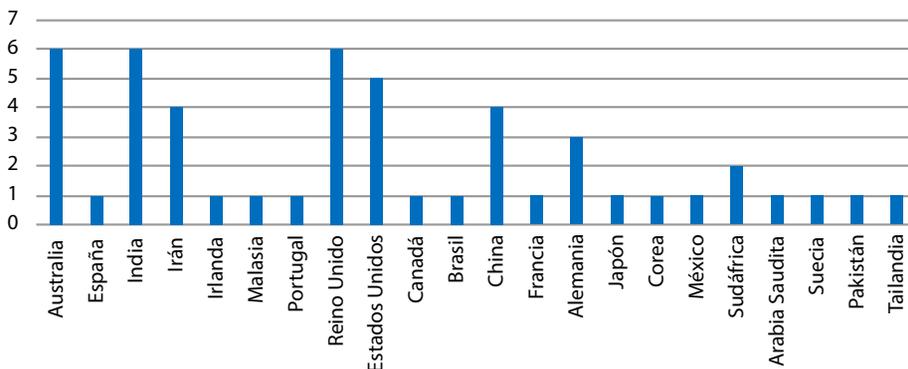
Después de haber realizado la revisión de los 50 artículos obtenidos, se procede a realizar el análisis cuantitativo, que se presentará en diferentes gráficos estadísticos de la siguiente manera:

- Cantidad de publicaciones científicas por país
- Cantidad de artículos más citados
- Cantidad de publicaciones científicas por año de acuerdo con las bases de datos utilizadas
- Cantidad de citas por publicaciones científicas
- Cantidad de repeticiones de palabras clave

El detalle del análisis puede verse desde el siguiente enlace: https://docs.google.com/spreadsheets/d/11W78U_L0xOpYw15FeKQ3HZB5Bnl_ly2s/edit?usp=sharing&uid=114897038577244842540&rtpof=true&sd=true

Figura 1

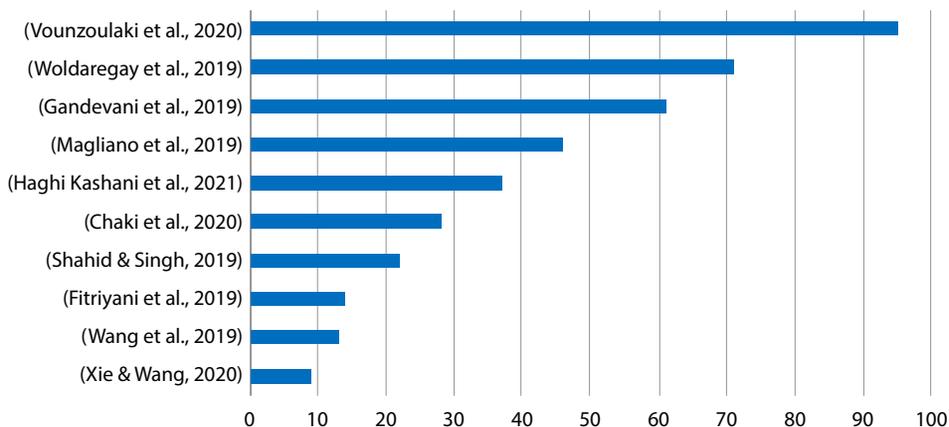
Artículos publicados por país



En la Figura 1, se observa que la mayor parte de los estudios obtenidos provienen de tres países: Australia, India y Reino Unido, los cuales muestran un mayor interés en el tema de investigación. Igualmente, los países con un rango medio de publicaciones y también con cierto interés en la investigación son Irán, Estados Unidos, China y Alemania.

Figura 2

Los diez artículos más citados

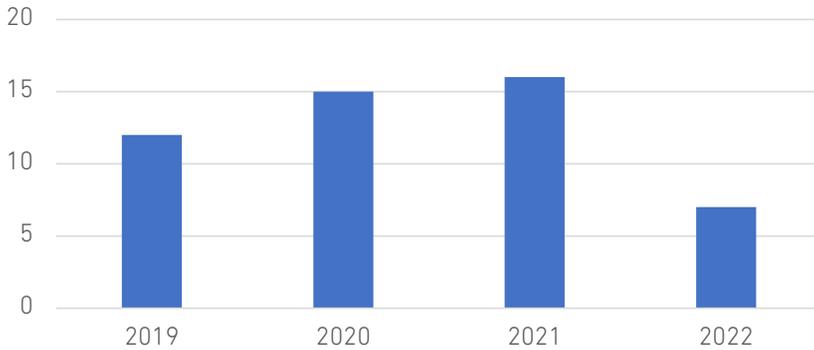


La Figura 2 muestra que la cantidad de trabajos más referenciados y tomados en cuenta por otros autores al momento de realizar un estudio sobre sistemas de predicción de la diabetes son Vounzoulaki et al. (2020) con 95 citaciones, Woldaregay

et al. (2019) con 71 y Gandevari et al. (2019) con 61, mientras que el décimo más citado tiene 9 citaciones.

Figura 3

Artículos publicados por año



La Figura 3 muestra que, de los 50 estudios recabados, la mayor concentración de trabajos realizados fue durante el año 2021. Asimismo, se puede observar que en el año 2020 también hay predominancia de estudios publicados, mientras que en los años 2019 y 2022 las publicaciones relacionadas con el tema de investigación son menores.

Tabla 3

Palabras clave más utilizadas

Palabras clave	Frecuencia
Machine learning	19
Diabetes	9
Diabetes mellitus, systematic review	6
Meta-analysis	5
Artificial intelligence, type 1 diabetes, deep learning, prediction	4
Type 2 diabetes, diagnosis, prognosis, prediabetes	3
Disease prediction, deep neural network, data mining, hypertension, telehealth, eHealth, detection, type 2 diabetes mellitus, prediction model	2

(continúa)

(continuación)

Palabras clave	Frecuencia
Uncertainty, medical data, computational intelligence, disease diagnosis, prognostic model, digital health, gestational diabetes mellitus, telemonitoring, chronic disease, cardiovascular disease, respiratory disease, metabolic syndrome, risk, eHealth, pregnancy, digital health, diabetes technologies, diabetes management, neural networks, validation, cardiovascular risk factors, automation, algorithms, devices, hypoglycaemia, sensors, Middle-East, antidepressants use, new onset, diabetes prediction, review, electronic health records, incidences, population-based study, protocol, stroke, NHANES, feature selection, predictors, Endocrinology, metabolism, osteoporosis, pituitary, thyroid, adrenal, clinical decision support, risk assessment, chronic disease management, natural language processing, physician-patient communication, prediabetes discussions, prediabetes management, type 2, continuous glucose monitor (CGM), glucose prediction, long short term memory (LSTM), diabetic prediction, fuzzy system, fused machine learning model, diabetic symptoms, objective distance, weighting factors, information gain, diabetes mellitus prediction, adaptive synthetic sampling, benchmark, deep neural network, regression, convolutional neural networks, regression analysis, feature extraction, ensemble learning, imbalanced data, glucose management, diabetic complications, thermography, Computer Aided Design (CAD), foot ulcers, image processing, m-Health, Bolus advisors, continuous glucose monitoring, flash glucose monitoring, insulin pumps, network meta-analysis, machine learning configurations, natural language processing, systematic review screening, transfer learning, applied case study, diagnostic criteria, prevalence, TRIPOD, PROBAST, prediction models, gestational diabetes, diabetes prediction, big data, blood glucose dynamics, blood glucose level prediction, SVM, Bayesian network, apriori algorithm, back propagation algorithm, artificial neural network, medical devices, insulin, glucagon, closed-loop, hybrid, algorithm, glucose responsive, islet encapsulation, Internet of Things (IoT), healthcare-health, risk factors, diseases, obesity, overweight, Chronic Kidney Disease (CKD), pancreatic carcinoma, early diagnosis, consensus, hyperglycaemia, multimorbidity, clustering, comorbidities, complications, patterns, population health, risk stratification, data-based algorithms or models, diabetics real data, type 2 diabetes mellitus, hypoglycaemia or hypoglycemia, blood glucose level, performance measure, Scopus database, PubMed database, metabolic syndrome, prediction model	1

La Tabla 3 detalla las palabras clave más utilizadas en los artículos de investigación. Las palabras *machine learning* y *diabetes* han sido las más consideradas en los artículos con 19 y 10 repeticiones, respectivamente. Asimismo, palabras como *diabetes mellitus*, *systematic review*, *meta-analysis*, *artificial intelligence* o *prediction* fueron usadas entre 6 y 4 veces como palabras clave. El resto de palabras fueron empleadas entre 1 y 3 veces. Se observa la cantidad de temas relacionados con el tema de estudio y la conexión del *machine learning* con el tema.

4. RESULTADOS

En esta sección se da respuesta a las preguntas de investigación planteadas en el trabajo. Para ello, es necesario indicar que los métodos de construcción se refieren a los diversos métodos, técnicas o algoritmos utilizados por los diferentes autores para construir la lógica que representa un modelo computacional para predecir una enfermedad. En este sentido, algunos autores usan algoritmos haciendo referencia a la estructura procedimental que genera una lógica computacional, mientras que otros emplean técnicas más complejas o métodos estructurados en pasos o etapas.

4.1 P1: ¿qué métodos de construcción son utilizados para la construcción de sistemas de identificación de tendencias relacionadas con la diabetes?

La Tabla 5 muestra los métodos de construcción más utilizados en las implementaciones de sistemas de predicción, recabados de los artículos de investigación seleccionados.

Tabla 5

Métodos de construcción utilizados en las implementaciones de sistemas de predicción

Referencia	Método de construcción	Descripción
(Schwartz et al., 2022), (De Silva, Jönsson & Demmer, 2020), (De Silva et al., 2021), (Zhu et al., 2021), (Zaitcev et al., 2020), (Tuppada & Patil, 2022), (Fregoso-Aparicio et al., 2021), (Kodama et al., 2022), (Zhang et al., 2022), (Fitriyani et al., 2019), (Hong et al., 2020)	Regresión logística	Es un modelo estadístico que se utiliza para determinar la probabilidad de que ocurra un evento. Además, es considerado uno de los algoritmos de <i>machine learning</i> más simples y más utilizados para tareas de clasificación.
(Ahmed et al., 2022), (Zaitcev et al., 2020), (Khan et al., 2021), (De Silva, Lee et al., 2020), (Jaiswal et al., 2021), (Woldaregay et al., 2019), (Chaki et al., 2020), (Fitriyani et al., 2019), (Shahid & Singh, 2019), (Safaei et al., 2021)	Redes neuronales artificiales (ANN) o perceptrón multicapa	Una red neuronal es un método de la inteligencia artificial que enseña a las computadoras a procesar datos en la forma en que lo hace el cerebro humano; para este caso, está expresado en el conocimiento de un experto.
(Schwartz et al., 2022), (Zhu et al., 2021), (Zaitcev et al., 2020), (Li et al., 2020), (Felizardo et al., 2021)	Redes neuronales convolucionales (CNN)	Son un tipo de redes neuronales artificiales muy efectivas para tareas de visión artificial, como la clasificación y segmentación de imágenes, entre otras aplicaciones.
(Nuankaew et al., 2021)	Distancia objetiva ponderada promedio (AWOD)	Es un método de aplicación a los datos mediante un promedio ponderado.

(continúa)

(continuación)

(Q. Wang et al., 2019)	Algoritmo CART, algoritmo DMP_MI	CART es un algoritmo de clasificación y árboles de regresión ampliamente utilizado en análisis predictivo.
(Jaiswal et al., 2021), (Woldaregay et al., 2019), (Nuankaew et al., 2021), (Zimmerman et al., 2021), (Hong et al., 2020)	<i>Support vector machine</i> (SVM)	Es un modelo de aprendizaje automático supervisado que utiliza algoritmos de clasificación para problemas de clasificación.
(Q. Wang et al., 2019), (Tuppap & Patil, 2022)	<i>Naive Bayes</i> (NB)	Es un modelo de predicción basado en la probabilidad bayesiana, considerada también como un algoritmo clasificador.
(Fitriyani et al., 2019), (Tuppap & Patil, 2022)	Árboles de decisión	Es un modelo de predicción cuyos posibles resultados se obtienen de una serie de decisiones relacionadas.
(Nuankaew et al., 2021)	Vecinos más cercanos (KNN)	Es un método de clasificación que, a pesar de su simplicidad, puede superar a los clasificadores más potentes.
(Q. Wang et al., 2019)	Muestreo sintético adaptativo (ADASYN)	Es un método de sobremuestreo que realiza una especie de balanceo en datos.
(Castelyn et al., 2021a)	Algoritmo heurístico basado en umbrales	Es un algoritmo basado reglas que, a partir de un conocimiento previo, indica qué acción tomar.
(Xie & Wang, 2020)	Método recursivo o directo	Es un método empleado cuando se sabe cómo resolver de manera directa un problema para un cierto conjunto de datos.
(Schwartz et al., 2022)	Sistema basado en reglas	Es un conjunto de reglas que representa de forma natural el conocimiento explícito de los expertos.
(Shahid & Singh, 2019)	AB (AdaBoost), algoritmo evolutivo	AdaBoost es un método basado en reglas, mientras que el algoritmo evolutivo es un método de optimización y búsqueda de soluciones basado en la evolución biológica.
(Vounzoulaki et al., 2020)	Estadísticos: Begg y Egger	Son pruebas estadísticas de estimaciones.
(Schwartz et al., 2022)	Programación neurolingüística (PNL)	Es un campo de la inteligencia artificial que investiga la manera de comunicar las máquinas con las personas.

4.2 P2: ¿qué métodos de construcción son los más eficientes para la construcción de un sistema de identificación de tendencias para la diabetes?

La Tabla 6 muestra el análisis realizado a partir de los resultados y conclusiones de cada trabajo. Se omiten aquellos donde los autores no indican de forma expresa los resultados o los niveles de eficiencia obtenidos con el método de construcción utilizado.

Tabla 6

Métodos más eficientes en la construcción de sistemas de predicción

Método de construcción	Resultados obtenidos de los métodos	Eficiencia
Regresión logística	Se obtuvo un alto rendimiento para predicciones (De Silva, Jönsson & Demmer, 2020); es un método que puede aplicarse dependiendo de las necesidades del sistema por implementar (Schwartz et al., 2022).	Muy eficiente
Redes neuronales artificiales	Se obtuvo un mayor rendimiento en comparación con otros métodos de pruebas, lo que ha generado grandes posibilidades en el reconocimiento de patrones de enfermedades (Safaei et al., 2021).	Muy eficiente
Distancia objetiva ponderada promedio (AWOD)	El método basado en AWOD proporcionó una precisión del 93,22 % y del 98,95 % para el conjunto de datos 1 y el conjunto de datos 2, respectivamente (Nuankaew et al., 2021).	Eficiente
DMP_MI	El algoritmo DMP_MI propuesto ha superado a otros algoritmos en precisión con un 87,10 % y otros indicadores de rendimiento del clasificador, con lo que ha demostrado un gran potencial para la predicción de la diabetes (Q. Wang et al., 2019).	Eficiente
Begg y Egger	Los resultados obtenidos indican que las pruebas de Begg y Egger no fueron estadísticamente significativas (Vounzoulaki et al., 2020).	Poco eficiente

4.3 P3: ¿qué modelos de diagnóstico son empleados para la detección de la diabetes?

La Tabla 7 muestra los modelos de diagnóstico computacionales y no computacionales empleados para la detección de enfermedades como la diabetes. Entre los computacionales están los modelos predictivos construidos mediante las técnicas de inteligencia artificial. Sin embargo, se encontraron también encuestas y criterios

de diagnóstico considerados modelos no computacionales para su aplicación en sistemas de predicción.

Tabla 7

Modelos de diagnóstico para la detección de la diabetes

Referencia	Modelos de diagnóstico
(Gautier et al., 2021)	Encuestas autoinformadas.
(Gautier et al., 2021), (Jaiswal et al., 2021), (Shahid & Singh, 2019), (Felizardo et al., 2021), (Hong et al., 2020), (Schwartz et al., 2022), (Zhang et al., n.d.2022), (Kodama et al., 2022), (Fregoso-Aparicio et al., 2021), (Tuppad & Patil, 2022), (Nuankaew et al., 2021), (Xie & Wang, 2020)	Modelos predictivos construidos mediante técnicas de inteligencia artificial como ANN, RL, <i>naive Bayes</i> , árboles de decisión, SVM, entre otros.
(Y. Wang et al., 2021), (Safaei et al., 2021), (Galbete et al., 2022)	Uso de criterios de diagnóstico de diabetes, como sexo, edad, consumo de alcohol, depresión, obesidad, IMC, hipertensión, etcétera.
(Madhava & Verma, 2019)	Termografía mediante el sistema detección asistido por ordenador (CAD).

4.4 P4: ¿qué problemas se han encontrado para construir un sistema de identificación de tendencias relacionadas con la diabetes?

La Tabla 8 muestra el análisis realizado a los resultados y conclusiones de cada trabajo estudiado; se rescatan solo aquellos donde se expresan de forma puntual problemas relacionados con la implementación y construcción de sistemas de predicción sobre la diabetes.

Tabla 8

Problemas en la construcción de sistemas de predicción relacionados con la diabetes

Referencia	Problemática
(Safaei et al., 2021)	Se menciona como problema que la implementación de un sistema de predicción basado en redes neuronales artificiales solo se puede aplicar para un conjunto de parámetros conocidos, es decir, las variables de entrada.
(De Silva et al., 2021)	Se evidencia cierto grado de comparabilidad en los resultados al emplear solo tres algoritmos de desempeño, pero existen ciertas diferencias identificadas que son notables. Debido a esto, se recalca la importancia de aplicar una variedad de algoritmos para obtener mejores resultados.
(Jaiswal et al., 2021)	Se menciona como problema que, si no se fusiona un conjunto de datos para lograr una mejor precisión, no se tendrá un modelo confiable.

(continúa)

(continuación)

(Schwartz et al., 2022)	Se menciona como problema que un sistema simple basado en reglas es inadecuado para identificar discusiones de prediabetes debido a la poca especificidad.
(Q. Wang et al., 2019)	Se menciona como problema que, debido a la falta de sensibilidad de la precisión para los datos desequilibrados, se consideran indicadores integrales para la evaluación del desempeño, como la precisión, la recuperación, la puntuación F1 y el área bajo la curva (AUC).
(Gautier et al., 2021)	Pocos estudios investigan problemas de una población en particular; los estudios se reducen a hospitales o clínicas, por lo que la calidad y la suficiencia de los datos es un obstáculo para el modelo que se va a utilizar.
(Asgari et al., 2021)	Se evidenció que el modelo desarrollado no aumentó la capacidad en el tamizaje/predicción de la DM tipo 2, principalmente en la parte de análisis, debido a la falta de validación externa de los modelos de predicción.

5. HALLAZGOS

Los resultados obtenidos a partir de los estudios seleccionados mencionan que los métodos de construcción más eficientes son la regresión logística, considerada un método sólido como otros algoritmos de clasificación, y las redes neuronales, que también se ajustan a sistemas de predicción, pero es necesario iterar varias veces en el entrenamiento para que el modelo o la arquitectura de red propuesta busque minimizar los errores y los resultados sean muy próximos a los del experto humano. Además, se halla que los modelos de diagnóstico más utilizados por la medicina son los computacionales, mediante los diversos modelos predictivos que provee la inteligencia artificial.

Los resultados obtenidos confirman que la inteligencia artificial es la base de nuevas herramientas que ayudan a diagnosticar y prevenir diversas enfermedades.

6. CONCLUSIONES

Los métodos de construcción utilizados en implementaciones de sistemas de predicción relacionados con la diabetes se basan en técnicas y algoritmos de inteligencia artificial en su mayoría, aunque también se utilizan las pruebas estadísticas para este fin.

Se concluye que la complejidad de cada método depende de los datos que sean manejados para inferir la solución. Generalmente, los modelos de aprendizaje automático dependen de las características del conjunto de datos que se tiene como entrada.

De acuerdo con el análisis realizado, se puede determinar que la regresión logística y las redes neuronales artificiales son los métodos de construcción que mejores resultados han mostrado en las predicciones realizadas.

También es posible afirmar que existen modelos no computacionales que pueden ser utilizados en la detección y diagnóstico de la diabetes, los cuales han presentado buenos resultados. Respecto a los problemas, se concluye que la mayoría se centra en los datos que usarán los modelos de predicción, siendo la calidad y la confiabilidad los factores más importantes en este aspecto.

REFERENCIAS

- Ahmed, U., Issa, G. F., Khan, M. A., Aftab, S., Khan, M. F., & Said, R. A. T. (2022). Prediction of diabetes empowered with fused machine learning. *IEEE Access*, *10*, 8529-8538. <https://doi.org/10.1109/ACCESS.2022.3142097>
- Asgari, S., Khalili, D., & Hosseinpanah, F. (2021). Prediction models for type 2 diabetes risk in the general population: A systematic review of observational studies. *International Journal of Endocrinology and Metabolism*, *19*(3). <https://doi.org/10.5812/ijem.109206.Systematic>
- Barbaresko, J., Neuenschwander, M., Schwingshackl, L., & Schlesinger, S. (2019). Dietary factors and diabetes-related health outcomes in patients with type 2 diabetes: Protocol for a systematic review and meta-analysis of prospective observational studies. *BMJ Open*, *9*(7). <https://doi.org/10.1136/bmjopen-2018-027298>
- Castelyn, G., Laranjo, L., Schreier, G., & Gallego, B. (2021a). Predictive performance and impact of algorithms in remote monitoring of chronic conditions: A systematic review and meta-analysis. *International Journal of Medical Informatics*, *156*, 104620. <https://doi.org/10.1016/j.ijmedinf.2021.104620>
- Castelyn, G., Laranjo, L., Schreier, G., & Gallego, B. (2021b). Predictive performance and impact of algorithms in remote monitoring of chronic conditions: A systematic review and meta-analysis. *International Journal of Medical Informatics*, *156*, 104620. <https://doi.org/10.1016/j.ijmedinf.2021.104620>
- Chaki, J., Thillai Ganesh, S., Cidham, S. K., & Ananda Theertan, S. (2020). Machine learning and artificial intelligence based diabetes mellitus detection and self-management: A systematic review. *Journal of King Saud University - Computer and Information Sciences*, *34*, 3204-3225. <https://doi.org/10.1016/j.jksuci.2020.06.013>

- Cicek, M., Buckley, J., Pearson-Stuttard, J., & Gregg, E. W. (2021). Characterizing multimorbidity from type 2 diabetes: Insights from clustering approaches. *Endocrinology and Metabolism Clinics of North America*, 50(3), 531-558. <https://doi.org/10.1016/j.ecl.2021.05.012>
- De Silva, K., Jönsson, D., & Demmer, R. T. (2020). A combined strategy of feature selection and machine learning to identify predictors of prediabetes. *Journal of the American Medical Informatics Association*, 27(3), 396-406. <https://doi.org/10.1093/jamia/ocz204>
- De Silva, K., Lee, W. K., Forbes, A., Demmer, R. T., Barton, C., & Enticott, J. (2020). Use and performance of machine learning models for type 2 diabetes prediction in community settings: A systematic review and meta-analysis. *International Journal of Medical Informatics*, 143, 104268. <https://doi.org/10.1016/j.ijmedinf.2020.104268>
- De Silva, K., Lim, S., Mousa, A., Teede, H., Forbes, A., Demmer, R. T., Jönsson, D., & Enticott, J. (2021). Nutritional markers of undiagnosed type 2 diabetes in adults: Findings of a machine learning analysis with external validation and benchmarking. *PLOS ONE*, 16(5), e0250832. <https://doi.org/10.1371/journal.pone.0250832>
- Diouri, O., Cigler, M., Vettoretti, M., Mader, J. K., Choudhary, P., & Renard, E. (2021). Hypoglycaemia detection and prediction techniques: A systematic review on the latest developments. *Diabetes/Metabolism Research and Reviews*, 37(7), e3449. <https://doi.org/10.1002/dmrr.3449>
- Domingo-Lopez, D. A., Lattanzi, G., H. J. Schreiber, L., Wallace, E. J., Wylie, R., O'Sullivan, J., Dolan, E. B., & Duffy, G. P. (2022). Medical devices, smart drug delivery, wearables and technology for the treatment of diabetes mellitus. *Advanced Drug Delivery Reviews*, 185, 114280. <https://doi.org/10.1016/j.addr.2022.114280>
- Eberle, C., Loehnert, M., & Stichling, S. (2021). Clinical effectiveness of different technologies for diabetes in pregnancy: Systematic literature review. *Journal of Medical Internet Research*, 23(4). <https://doi.org/10.2196/24982>
- Farmanfarma, K. H. K., Zareban, I., & Adineh, H. A. (2020). Prevalence of type 2 diabetes in Middle-East: Systematic review & meta-analysis. *Primary Care Diabetes*, 14(4), 297-304. <https://doi.org/10.1016/j.pcd.2020.01.003>
- Felizardo, V., Garcia, N. M., Pombo, N., & Megdiche, I. (2021). Data-based algorithms and models using diabetics real data for blood glucose and hypoglycaemia

- prediction – A systematic literature review. *Artificial Intelligence in Medicine*, 118, 102120. <https://doi.org/10.1016/j.artmed.2021.102120>
- Fitriyani, N. L., Syafrudin, M., Alfian, G., & Rhee, J. (2019). Development of disease prediction model based on ensemble learning approach for diabetes and hypertension. *IEEE Access*, 7, 144777-144789. <https://doi.org/10.1109/ACCESS.2019.2945129>
- Fregoso-Aparicio, L. F., Noguez, J., Montesinos, L., & García, J. A. G. (2021). Machine learning and deep learning predictive models for type 2 diabetes: A systematic review. *Diabetology & Metabolic Syndrome*, 13. <https://doi.org/10.1186/s13098-021-00767-9>
- Galbete, A., Tamayo, I., Libroero, J., Enguita-Germán, M., Cambra, K., & Ibáñez-Beroiz, B. (2022). Cardiovascular risk in patients with type 2 diabetes: A systematic review of prediction models. *Diabetes Research and Clinical Practice*, 184. <https://doi.org/10.1016/j.diabres.2021.109089>
- Gandevani, S. B., Amiri, M., Yarandi, R. B., & Tehrani, F. R. (2019). The impact of diagnostic criteria for gestational diabetes on its prevalence: A systematic review and meta-analysis. *Diabetology & Metabolic Syndrome*, 11, 1-18. <https://doi.org/10.1186/s13098-019-0406-1>
- Gautier, T., Ziegler, L. B., Gerber, M. S., Campos-Náñez, E., & Patek, S. D. (2021). Artificial intelligence and diabetes technology: A review. *Metabolism: Clinical and Experimental*, 124. <https://doi.org/10.1016/j.metabol.2021.154872>
- Haghi Kashani, M., Madanipour, M., Nikravan, M., Asghari, P., & Mahdipour, E. (2021). A systematic review of IoT in healthcare: Applications, techniques, and trends. *Journal of Network and Computer Applications*, 192, 103164. <https://doi.org/10.1016/j.jnca.2021.103164>
- Hong, N., Park, H., & Rhee, Y. (2020). Machine learning applications in endocrinology and metabolism research: An overview. *Endocrinology and Metabolism*, 35(1), 71-84. <https://doi.org/10.3803/EnM.2020.35.1.71>
- Jaiswal, V., Negi, A., & Pal, T. (2021). A review on current advances in machine learning based diabetes prediction. *Primary Care Diabetes*, 15(3), 435-443. <https://doi.org/10.1016/j.pcd.2021.02.005>
- Kattini, R., Hummelen, R., & Kelly, L. (2020). Early gestational diabetes mellitus screening with glycated hemoglobin: A systematic review. *Journal of Obstetrics and Gynaecology Canada*, 42(11), 1379-1384. <https://doi.org/10.1016/j.jogc.2019.12.015>

- Khan, F. A., Zeb, K., Al-Rakhami, M., Derhab, A., & Bukhari, S. A. C. (2021). Detection and prediction of diabetes using data mining: A comprehensive review. *IEEE Access*, 9, 43711-43735. <https://doi.org/10.1109/ACCESS.2021.3059343>
- Kodama, S., Fujihara, K., Horikawa, C., Kitazawa, M., Iwanaga, M., Kato, K., Watanabe, K., Nakagawa, Y., Matsuzaka, T., Shimano, H., & Sone, H. (2022). Predictive ability of current machine learning algorithms for type 2 diabetes mellitus: A meta-analysis. *Journal of Diabetes Investigation*, 13(5), 900-908. <https://doi.org/10.1111/jdi.13736>
- Kvitkina, T., Narres, M., Claessen, H., Metzendorf, M. I., Richter, B., & Icks, A. (2020). Incidence of stroke in the diabetic compared with the non-diabetic population: A systematic review protocol. *Diabetes/Metabolism Research and Reviews*, 36(6). <https://doi.org/10.1002/dmrr.3311>
- Li, K., Daniels, J., Liu, C., Herrero, P., & Georgiou, P. (2020). Convolutional recurrent neural networks for glucose prediction. *IEEE Journal of Biomedical and Health Informatics*, 24(2), 603-613. <https://doi.org/10.1109/JBHI.2019.2908488>
- Madhava, P. S., & Verma, S. (2019). A systematic literature review for early detection of type ii diabetes. En *2019 5th International Conference on Advanced Computing and Communication Systems (ICACCS)* (pp. 220-224). <https://doi.org/10.1109/ICACCS.2019.8728377>
- Magliano, D. J., Islam, R. M., Barr, E. L. M., Gregg, E. W., Pavkov, M. E., Harding, J. L., Tabesh, M., Koye, D. N., & Shaw, J. E. (2019). Trends in incidence of total or type 2 diabetes: Systematic review. *BMJ*, 366. <https://doi.org/10.1136/bmj.l5003>
- Nuankaew, P., Chaising, S., & Temdee, P. (2021). Average weighted objective distance-based method for type 2 diabetes prediction. *IEEE Access*, 9, 137015-137028. <https://doi.org/10.1109/ACCESS.2021.3117269>
- Pease, A., Lo, C., Earnest, A., Kiriakova, V., Liew, D., & Zoungas, S. (2019). The efficacy of technology in type 1 diabetes: A systematic review, network meta-analysis, and narrative synthesis. *Diabetes Technology & Therapeutics*, 22(5), 411-421. <https://doi.org/10.1089/dia.2019.0417>
- Peer, N., Balakrishna, Y., & Durao, S. (2020). Screening for type 2 diabetes mellitus. *Cochrane Database of Systematic Reviews*, 5. <https://doi.org/10.1002/14651858.CD005266.pub2>
- Ray, A., & Chaudhuri, A. K. (2021). Smart healthcare disease diagnosis and patient management: Innovation, improvement and skill development. *Machine Learning with Applications*, 3, 100011. <https://doi.org/10.1016/j.mlwa.2020.100011>

- Safaei, M., Sundararajan, E. A., Driss, M., Boulila, W., & Shapi'i, A. (2021). A systematic literature review on obesity: Understanding the causes & consequences of obesity and reviewing various machine learning approaches used to predict obesity. *Computers in Biology and Medicine*, 136, 104754. <https://doi.org/10.1016/j.compbimed.2021.104754>
- Santos, D. S., Regina, C., Batistelli, S., & Marilac, M. (2022). The effectiveness of the use of telehealth programs in the care of individuals with hypertension and, or diabetes mellitus: Systematic review and meta-analysis. *Diabetology & Metabolic Syndrome*, 14, 76. <https://doi.org/10.1186/s13098-022-00846-5>
- Schwartz, J. L., Tseng, E., Maruthur, N. M., & Rouhizadeh, M. (2022). Identification of prediabetes discussions in unstructured clinical documentation: Validation of a natural language processing algorithm. *Journal of Medical Internet Research*, 10(2). <https://doi.org/10.2196/29803>
- Shahid, A. H., & Singh, M. P. (2019). Computational intelligence techniques for medical diagnosis and prognosis: Problems and current developments. *Biocybernetics and Biomedical Engineering*, 39(3), 638-672. <https://doi.org/10.1016/j.bbe.2019.05.010>
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333-339. <https://doi.org/10.1016/j.jbusres.2019.07.039>
- Tuppad, A., & Patil, S. D. (2022). Machine learning for diabetes clinical decision support: A review. *Advances in Computational Intelligence*, 2(2), 1-24. <https://doi.org/10.1007/s43674-022-00034-y>
- Vounzoulaki, E., Khunti, K., Abner, S. C., Tan, B. K., Davies, M. J., & Gillies, C. L. (2020). Progression to type 2 diabetes in women with a known history of gestational diabetes: Systematic review and meta-analysis. *BMJ*, 369, m1361. <https://doi.org/10.1136/bmj.m1361>
- Wang, Q., Cao, W., Guo, J., Ren, J., Cheng, Y., & Davis, D. N. (2019). DMP_MI: An effective diabetes mellitus classification algorithm on imbalanced data with missing values. *IEEE Access*, 7, 102232-102238. <https://doi.org/10.1109/ACCESS.2019.2929866>
- Wang, Y., Liu, D., Li, X., Liu, Y., & Wu, Y. (2021). Antidepressants use and the risk of type 2 diabetes mellitus: A systematic review and meta-analysis. *Journal of Affective Disorders*, 287(45), 41-53. <https://doi.org/10.1016/j.jad.2021.03.023>
- Woldaregay, A. Z., Årsand, E., Walderhaug, S., Albers, D., Mamykina, L., Botsis, T., & Hartvigsen, G. (2019). Data-driven modeling and prediction of blood

- glucose dynamics: Machine learning applications in type 1 diabetes. *Artificial Intelligence in Medicine*, 98, 109-134. <https://doi.org/10.1016/j.artmed.2019.07.007>
- Xie, J., & Wang, Q. (2020). Benchmarking machine learning algorithms on blood glucose prediction for type i diabetes in comparison with classical time-series models. *IEEE Transactions on Biomedical Engineering*, 67(11), 3101-3124. <https://doi.org/10.1109/TBME.2020.2975959>
- Zaitcev, A., Eissa, M. R., Hui, Z., Good, T., Elliott, J., & Benaissa, M. (2020). A deep neural network application for improved prediction of HbA 1c in type 1 diabetes. *IEEE Journal of Biomedical and Health Informatics*, 24(10), 2932-2941. <https://doi.org/10.1109/JBHI.2020.2967546>
- Zhang, H., Shao, J., Chen, D., Zou, P., Ciu, N., Tang, L., Wang, D., & Ye, Z. (2020). Reporting and methods in developing prognostic prediction models for metabolic syndrome: A systematic review and critical appraisal. *Diabetes, Metabolic Syndrome and Obesity*, 13, 4981-4992.
- Zhang, Z., Yang, L., Han, W., Wu, Y., Zhang, L., Gao, C., Jiang, K., Liu, Y., & Wu, H. (2022). Machine learning prediction models for gestational diabetes mellitus: Meta-analysis. *Journal of Medical Internet Research*, 24(3). <https://doi.org/10.2196/26634>
- Zheng, M., Bernardo, C. O., Stocks, N., & Gonzalez-Chica, D. (2022). Diabetes mellitus diagnosis and screening in Australian general practice: A national study. *Journal of Diabetes Research*. DOI: 10.1155/2022/156640
- Zhu, T., Li, K., Herrero, P., & Georgiou, P. (2021). Deep learning for diabetes: A systematic review. *IEEE Journal of Biomedical and Health Informatics*, 25(7), 2744-2757. <https://doi.org/10.1109/JBHI.2020.3040225>
- Zimmerman, J., Soler, R. E., Lavinder, J., Murphy, S., Atkins, C., Hulbert, L., Lusk, R., & Ng, B. P. (2021). Iterative guided machine learning-assisted systematic literature reviews: A diabetes case study. *Systematic Reviews*, 10, 97. <https://doi.org/10.1186/s13643-021-01640-6>

ARTÍCULO DE DIVULGACIÓN

Recibido: 13 de octubre del 2022 / Aceptado: 6 de noviembre del 2022

doi: <https://doi.org/10.26439/interfases2022.n016.6105>

FORMULACIÓN DE REQUISITOS PARA UN SISTEMA DE INFORMACIÓN DE APOYO AL PROCESO DE INDUCCIÓN DE ESTUDIANTES UNIVERSITARIOS: CASO UNAD

PILAR ALEXANDRA MORENO

pilar.moreno@unad.edu.co

<https://orcid.org/0000-0002-6990-6105>

Escuela de Ciencias Básicas, Tecnología e Ingeniería,
Universidad Nacional Abierta y a Distancia, Colombia

JUAN OLEGARIO MONROY VÁSQUEZ

juan.monroy@unad.edu.co

<https://orcid.org/0000-0003-1757-2016>

Escuela de Ciencias Básicas, Tecnología e Ingeniería,
Universidad Nacional Abierta y a Distancia, Colombia

RESUMEN. La Universidad Nacional Abierta y a Distancia (UNAD), institución de educación superior pública en Colombia, ofrece programas en diferentes niveles de formación tecnológica, profesional, especializaciones y maestrías, a través de la modalidad de educación a distancia y en ambientes virtuales de aprendizaje. Por tanto, resulta evidente que la comunidad estudiantil unadista que ingresa por primera vez requiere de sistemas y herramientas de inducción especiales que les ayuden en su ingreso, adaptación y transición a la formación virtual. Así pues, surge la necesidad de crear un sistema de información que apoye el proceso de inducción de estudiantes de primera matrícula, centrado en su transición de una formación presencial hacia una formación virtual. Este sistema de información debe tener dos requisitos básicos: en primer lugar, ser un sistema en línea que sirva de guía a los estudiantes en los primeros pasos en su ruta de formación virtual y, en segundo lugar, ayudar a la universidad en las decisiones sobre las herramientas de inducción y los recursos de adaptación especiales para responder a la población heterogénea que recibe, dado que ofrece un sistema abierto, sin restricciones para el ingreso y aceptación de la población estudiantil. En este sentido, el objetivo de este trabajo es describir el análisis llevado a cabo para el planteamiento de requisitos, como parte del proceso del diseño del sistema de información.

PALABRAS CLAVE: sistema de información, proceso de inducción, caracterización, desarrollo tecnológico, análisis de información

FORMULATION OF REQUIREMENTS FOR AN INFORMATION SYSTEM TO SUPPORT THE INDUCTION PROCESS OF UNIVERSITY STUDENTS: UNAD CASE

ABSTRACT. The Universidad Nacional Abierta y a Distancia (UNAD), a public higher education institution in Colombia, offers distance education programs at different levels of technical and professional training, specializations, and master's degrees in virtual learning environments. Therefore, those entering the unadista student community for the first time require induction systems and tools that help their entry, adaptation, and transition to virtual training. Thus, an information system that supports the induction process of first-time students focused on their transition from face-to-face to virtual training is needed. This information system must comply with two basic requirements: firstly, to be an online system that guides students in the first steps of their virtual training and, secondly, to help the university's decision-making regarding the induction tools and the adaptation resources to respond to the heterogeneous population it receives, since it offers an open system without restrictions for the entrance and acceptance of students. This work aims to describe the analysis determining these requirements as part of the information system design.

KEYWORDS: information system, induction process, characterization, technological development, information analysis

1. INTRODUCCIÓN

Este artículo de divulgación presenta el proceso de análisis de la información del sistema que se propone desarrollar, a través de la caracterización del proceso de inducción en la Universidad Nacional Abierta y a Distancia (UNAD) de Colombia. Así, se relacionan el contexto de la universidad, sus unidades, sistemas y dispositivos encargados de acompañar el inicio de la formación de los estudiantes en los entornos virtuales de aprendizaje que la universidad tiene dispuestos para ellos, lo cual permite empezar a analizar y describir los primeros elementos dentro del diseño del sistema.

Para garantizar el acompañamiento a sus estudiantes, la UNAD cuenta en su estructura organizacional con la Vicerrectoría de Servicios a Aspirantes, Estudiantes y Egresados (VISAE), la cual gestiona y ayuda a asegurar la calidad del servicio unadista. Una de sus funciones es el acompañamiento a los aspirantes y estudiantes de primera matrícula que ingresan a la universidad. Igualmente, ofrece consejería académica que se encarga de, entre otras responsabilidades, disponer estrategias y escenarios de acompañamiento y orientación a los estudiantes, para que puedan superar los obstáculos que se les presenten en su proceso de aprendizaje y apropiarse herramientas de la modalidad abierta y a distancia, y de los ambientes virtuales. Esta consejería académica centra su atención en estudiantes de primera matrícula, que son los que particularmente se encuentran en la transición hacia la formación virtual; y para ello, establece las acciones que se convierten en los aspectos centrales de la caracterización del proceso de inducción y que deben contemplarse en el sistema de información propuesto.

De esta forma, para lograr estructurar este sistema de información, es preciso conocer y caracterizar, de manera detallada, tanto el proceso de inducción de estudiantes de primera matrícula, el cual traza la ruta que debe seguir actualmente un estudiante en el inicio de su formación en la universidad; como las unidades, estamentos y dispositivos encargados de organizar los servicios de atención a los estudiantes. Esto es parte de la etapa de levantamiento de información inicial dentro del análisis de dicho sistema. En este artículo, se describen como resultados los elementos identificados en la revisión del proceso de inducción y que constituyen la base para el correspondiente análisis del sistema de información:

- Caracterización de la transición a la formación virtual de estudiantes de primera matrícula: se enfoca en el estudio del proceso de ciclo de vida del estudiante y su procedimiento de consejería académica.
- Inducción: para identificar las acciones que debe seguir el estudiante cuando ingresa a la universidad.

- Prueba de caracterización a estudiantes de primera matrícula: con el fin de determinar las variables y características de los estudiantes, las cuales se deben considerar para la construcción del sistema de información de apoyo en la transición a una educación virtual.
- Plan de Acción Pedagógica Contextualizada (PAPC): para revisar la dinámica de organización y trabajo para la incorporación de la programación de las acciones de inducción en cada zona dentro del sistema de información.

Este artículo está estructurado de la siguiente manera: la segunda sección explica el modelo del proceso de inducción; la tercera describe la formulación de los requisitos del sistema de información; y, finalmente, se extraen las conclusiones y recomendaciones para trabajos futuros.

2. MODELO DEL PROCESO DE INDUCCIÓN

En esta sección se hace una revisión y análisis de las acciones de consejería académica de la UNAD, unidad encargada de establecer estrategias y escenarios de acompañamiento y orientación a los estudiantes de primera matrícula en su proceso de adaptación a la metodología y a los escenarios virtuales de aprendizaje que se encuentran en la universidad. Esta consejería académica, como se había mencionado, centra su atención en estudiantes de primera matrícula, quienes se hallan en la transición hacia la formación virtual. En ese sentido, Acevedo Velandia (2020) plantea los siguientes lineamientos:

- La caracterización del estudiante, orientada a establecer el perfil sociodemográfico y educativo de los estudiantes. Sobre esta se basa el diseño inicial de las estrategias pedagógicas orientadas a aumentar la permanencia e interés en su proceso de formación.
- La acogida e integración unadista, que contempla tanto jornadas de inducción permanente como el desarrollo del curso Cátedra Unadista. Buscan la apropiación, por parte de los estudiantes, del modelo de educación a distancia, de la metodología de estudio y del manejo de la plataforma virtual.
- El Plan de Acción Pedagógica Contextualizada pretende que se diseñen e implementen estrategias particulares y contextualizadas, de acuerdo con los resultados de la caracterización de estudiantes.
- Acompañamiento desde la consejería, en donde se encuentran acciones de intervención para estudiantes, orientadas a disminuir los factores de riesgo, aumentar su permanencia y favorecer el éxito académico en la UNAD.

Además de la estructura y unidades organizacionales ya mencionadas, la UNAD también cuenta con un sistema integrado de gestión (SIG) como su principal herramienta de gestión universitaria, el cual facilita el desarrollo de los proyectos y el cumplimiento de los objetivos establecidos en el componente misional de la UNAD, incluidos la VISAE y su consejería virtual. Este SIG ha sido certificado por SGS¹ como ente certificador externo, bajo los estándares ISO 9001, ISO 14001, ISO 45001 y NTC 5580. Específicamente, dentro de este sistema de gestión, se encuentra establecido el proceso de ciclo de vida del estudiante, que estandariza el accionar frente a los estudiantes de la UNAD, su atención, su acompañamiento y su formación, así como el procedimiento de consejería virtual, unidad fundamental que planifica los proyectos y procesos frente a los primeros pasos de los estudiantes en la universidad. Todos estos procesos de negocio consolidan el modelo de inducción en la UNAD. De esta forma, se establecieron las siguientes actividades: (1) transición a la formación virtual de estudiantes de primera matrícula, (2) programa de acogida e integración unadista, (3) prueba de caracterización y (4) el Plan de Acción Pedagógica Contextualizada.

2.1 Transición a la formación virtual de estudiantes de primera matrícula

Como se mencionó, para empezar el análisis y caracterización de la transición de los estudiantes de primera matrícula de la UNAD hacia una formación virtual, es necesario revisar el proceso misional denominado *ciclo de vida del estudiante*, cuyo objetivo es consolidar la interacción con aspirantes, estudiantes y egresados, desde la identificación y caracterización de sus necesidades, el acompañamiento y orientación académica, pedagógica, psicosocial y administrativa, para el logro de sus aprendizajes y formación integral, el otorgamiento de su titulación y el apoyo a su inserción en el mercado laboral o productivo, de conformidad con los programas y servicios de los diferentes sistemas de formación de la institución (UNAD, 2017b).

A continuación, se presentan tanto los insumos como las actividades establecidas y que tienen relación con la transición de la formación presencial a la formación virtual de los estudiantes en la UNAD. Se analizaron tanto el proceso como su procedimiento, llamado *P-7-7 Consejería Académica*, que es el que directamente tiene que ver con ese primer acercamiento del estudiante a su proceso de formación, pues su objetivo es:

Establecer las acciones para orientar a los aspirantes para el ingreso a la UNAD; programar, desarrollar y evaluar las actividades de acogida y acompañamiento a los estudiantes de primera matrícula en programas de grado y

1 Como se indica en su página web institucional: <https://sig.unad.edu.co/acerca-del-sig>

posgrado, con el fin de favorecer su integración a la modalidad de educación a distancia, la apropiación de las estrategias de aprendizaje y su promoción académica. (UNAD, 2017b, p. 1)

2.1.1 *Insumos, actividades y salidas del proceso de ciclo de vida del estudiante*

En este proceso, los insumos más importantes son los que se relacionan de manera directa con el proceso de transición del estudiante que viene de una formación presencial y pasa a una virtual. Estos son los siguientes:

- Procesos formativos previos: es ese conjunto de presaberes con los que llega el estudiante y que para la universidad se convierten en la base para configurar actividades, tareas y/o estrategias conducentes a lograr una nivelación del estudiante en los conceptos básicos con los cuales debe iniciar su proceso de formación.
- Infraestructura tecnológica de la UNAD (plataforma tecnológica integrada y sistemas de información): se refiere al conjunto de elementos de *hardware* y *software* que se requieren para dar soporte al desarrollo de los procesos de formación en el cumplimiento de la misión de la universidad, a través de un soporte de conectividad con características como disponibilidad, seguridad, inclusión y estabilidad.
- Recursos bibliográficos y repositorio institucional de contenidos académicos y de investigación: buscan dar soporte a los procesos de formación, poniendo a disposición de la comunidad académica de la universidad material académico, en diversos formatos, propendiendo a una cultura del autoaprendizaje personal y profesional.
- Cursos académicos y recursos educativos digitales: espacios y recursos educativos a través de los cuales los estudiantes inician su proceso de formación en los programas matriculados.
- Servicios ofrecidos a aspirantes, estudiantes y egresados: buscan el fortalecimiento de aspectos relacionados con el desarrollo humano de la comunidad académica y están dispuestos para cada uno de los integrantes desde el Sistema de Bienestar Integral Unadista.

Entre las actividades del proceso de ciclo de vida del estudiante, están poner en marcha el programa de acogida e inducción unadista y realizar acompañamiento psicopedagógico, didáctico y académico a los estudiantes que inician el proceso formativo. Asimismo, implementar estrategias de acompañamiento docente a los estudiantes en los cursos académicos.

Finalmente, las salidas del procedimiento son estudiantes de primera matrícula con acompañamiento psicopedagógico, didáctico y académico. Además, estudiantes con acompañamiento docente.

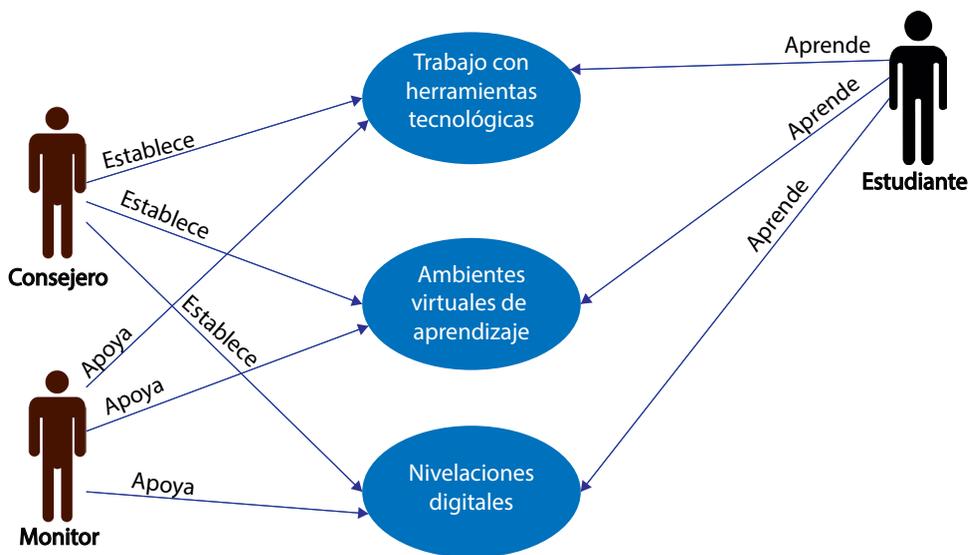
2.1.2 Insumos, actividades y salidas del procedimiento de consejería

Dentro de la descripción del procedimiento, es importante destacar aquellos elementos que cobran gran relevancia en la transición que debe realizar el estudiante hacia la formación virtual, tales como la caracterización del estudiante, la inducción unadista y los planes de acción contextualizados. Entonces, se deben considerar los siguientes aspectos:

- Inducción: conjunto de actividades, de tipo introductorio, orientadas a la apropiación del estudiante de la modalidad de educación a distancia y sus diversas metodologías que la institución ofrece al iniciar el proceso formativo (UNAD, 2013, artículo 15).
- Prueba de caracterización: prueba que se aplica a los estudiantes de primera matrícula con el propósito de identificar el perfil de ingreso, sus competencias básicas y generales, a fin de establecer estrategias pedagógico-didácticas y de acompañamiento pertinentes, que contribuyan al desempeño exitoso del estudiante en la modalidad de educación abierta y a distancia, y en los ambientes virtuales de aprendizaje. La prueba es de carácter obligatorio y no condiciona su ingreso (UNAD, 2013, artículo 45). Para los estudiantes de posgrado, la prueba de caracterización no contempla la evaluación de competencias.
- Plan de Acción Pedagógica Contextualizada: servicio académico de apoyo que la institución le ofrece al estudiante con el propósito de contribuir al logro de las competencias requeridas para el óptimo desarrollo de su proceso de aprendizaje. Este servicio se realiza a través de la oferta de estrategias de carácter pedagógico y didáctico, en las áreas requeridas, de acuerdo con los resultados de la prueba de caracterización (UNAD, 2013, artículo 20).

Figura 1

Proceso de transición



2.2 Programa de acogida e integración unadista

La UNAD cuenta con el programa de acogida e integración a la comunidad unadista, desde el cual se trabajan en dos líneas de acción: acogida e integración unadista.

2.2.1 Acogida

Está enfocada en la atención a aspirantes y estudiantes, inducción a estudiantes de primera matrícula y el Bienestar Unadista, a partir de diferentes estrategias:

- Atención a aspirantes y estudiantes: a partir de elementos como la ventanilla de atención primaria, apoyo al proceso de matrícula, información del programa que se va a cursar.
- Inducción a estudiantes de primera matrícula: la primera sesión de inducción a los estudiantes de la UNAD se realiza antes de iniciar las actividades académicas de cada periodo, incluidos los periodos intersemestrales, según la programación académica establecida y regulada por la circular de programación emitida anualmente. Esta actividad permite contextualizar al estudiante en la UNAD, dándole a conocer los servicios con los que cuenta y el acompañamiento que va a recibir durante su formación como profesional.

El primer encuentro de inducción se organiza en tres momentos que se describen en la Tabla 1.

Tabla 1

Momentos del primer encuentro de inducción

Momento	Objetivo	Actividad
Momento 1. Explorando-ando	Reconocer la comunidad unadista, su modelo y su proyecto institucional	Parte I. Presentación y acogida unadista Parte II. Presentación del contexto institucional Bienvenida de Bienestar Unadista
Momento 2. Indagando-ando	Familiarizar al estudiante con su escuela y el campus virtual	Parte I. Presentación del programa y plan de estudios Parte II. Modelación de la navegación en el campus virtual Bienvenida de Bienestar Unadista
Momento 3. Afianzando-ando	Desarrollar estrategias efectivas que permitan la apropiación del campus virtual por parte del estudiante de la universidad	Uso y conocimiento de la biblioteca virtual y de los laboratorios El trabajo colaborativo como estrategia de aprendizaje en la UNAD Generalidades sobre tecnologías para la gestión del conocimiento Bienvenida de Bienestar Unadista

Nota. Elaborado por los autores con datos de UNAD (2015).

El segundo encuentro de inducción pretende fortalecer el conocimiento de estrategias que le faciliten al estudiante la adaptación al modelo y el desarrollo del aprendizaje autónomo. Este segundo momento se realiza a partir de la programación de encuentros en los centros o con mediación *b-learning*, donde se desarrollan los siguientes temas:

- Manejo de actividades en cursos virtuales, uso de laboratorios y biblioteca virtual,
- El trabajo colaborativo como estrategia de aprendizaje en la UNAD
- Tecnologías para la gestión del conocimiento
- Bienestar Unadista

2.2.2 Integración unadista

Orientada a la reinducción de estudiantes y a la formación de la impronta unadista, se plantea las siguientes acciones:

- Reinducción de estudiantes: dirigida a quienes no asistieron al primer encuentro o que aún tienen dudas.

- Formación de la impronta unadista: para la cual se establecieron la Cátedra Región, la Cátedra Social Solidaria (Sistema de Servicio Social Unadista [SISSU]) y la Cátedra Unadista. Esta última pertenece al núcleo básico común en todos los programas de pregrado y posgrado que ofrece la universidad y está representada a través de un curso académico.

Figura 2

Proceso de acogida e inducción unadista

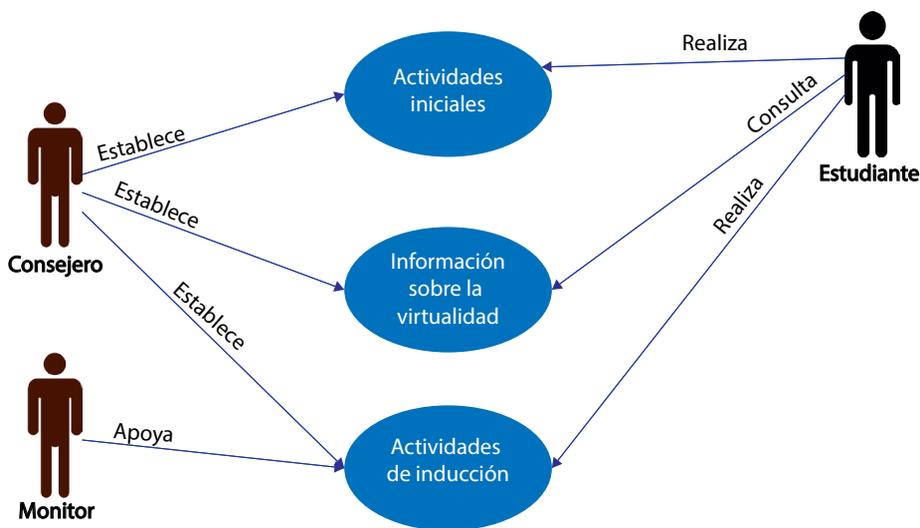
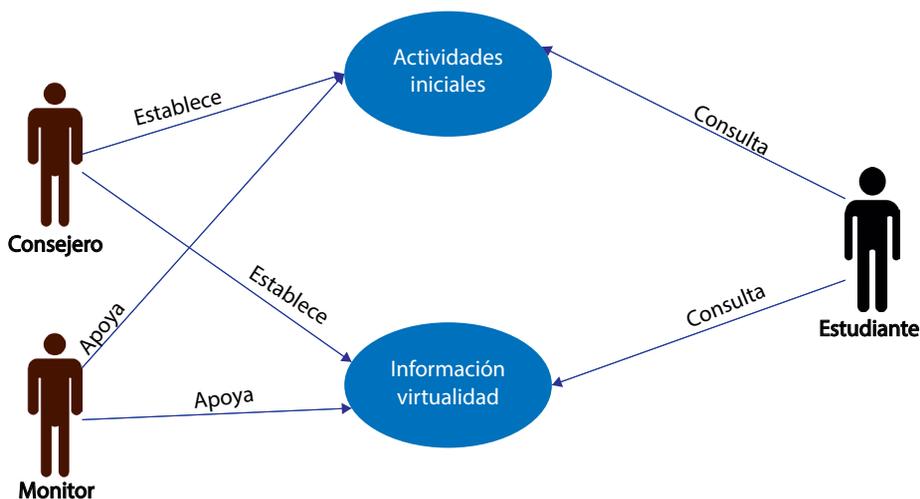


Figura 3

Proceso de atención a aspirantes



2.3 Prueba de caracterización

Todos los estudiantes, en las actividades iniciales, deben pasar una prueba de caracterización que permite a la UNAD conocer sus aspectos sociodemográficos, académicos, familiares, laborales y psicológicos, con el fin de identificar con precisión las potencialidades y necesidades de mejoramiento con las que inician el proceso de formación profesional.

La información derivada de la prueba de caracterización constituye el marco de referencia para la implementación de acciones estratégicas tanto para la nivelación en competencias académicas de los estudiantes como para la vinculación efectiva a las distintas líneas de acción de Bienestar Unadista. Ambos tipos de acciones constituyen el Plan de Acción Pedagógica Contextualizada (PAPC) (UNAD, 2020). Con base en esta información, se les ofrece a los estudiantes de manera oportuna los programas y servicios que más aporten a su adaptación, permanencia y graduación en la universidad. La prueba es de carácter obligatorio y debe ser diligenciada completamente por los estudiantes antes de iniciar sus actividades para continuar con su proceso formativo. Contempla 16 aspectos, tal como se muestra en la Tabla 2.

Tabla 2

Aspectos de la caracterización de estudiantes nuevos

Aspectos solicitados	Número de ítems	
Datos personales	22	TD, documento, estudiante, último acceso al campus, fecha de encuesta, edad, sexo, país, departamento, estrato, zona de residencia, estado civil, contacto, zona, CEAD, escuela, programa, matrícula en convenio.
Grupos poblacionales	12	Raizal, palenquero, afrocolombiano, otra comunidad negra, indígena, víctima desplazado, víctima ACR, funcionario INPEC, recluso INPEC, tiempo de condena, centro de reclusión.
Datos familiares	11	Tipo de vivienda actual, con quién vive actualmente, personas que tiene a su cargo, dependencia económica, escolaridad padre y madre, número de hermanos, familiares con estudios en la UNAD.
Datos académicos	11	Colegio bachillerato, modalidad, estudios anteriores, tiempo sin estudiar, ha tomado cursos virtuales, razón para elegir el programa, razón para estudiar en la UNAD.
Equipos con que cuenta para acceder al campus	11	Computadora de escritorio, portátil, tableta, teléfono inteligente; cuenta con servicio de energía eléctrica, cuenta con servicio de internet; ha usado plataformas virtuales con anterioridad; manejo de paquetes ofimáticos, foros virtuales, uso de correo electrónico.

(continúa)

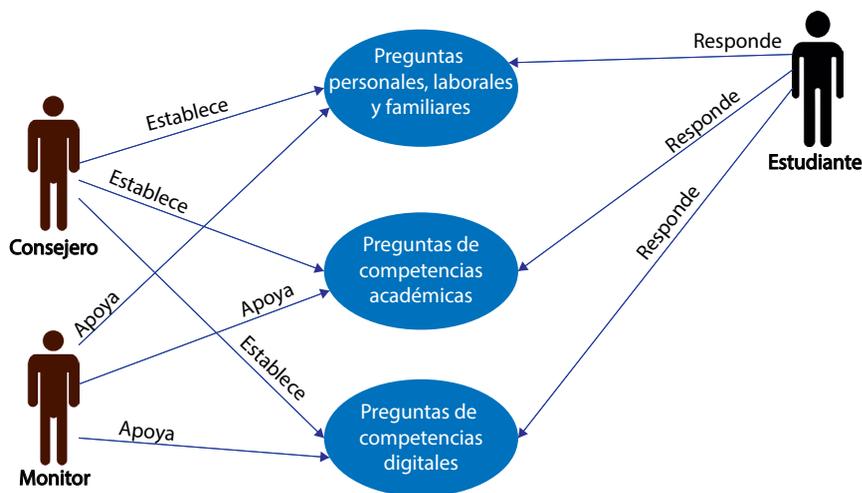
(continuación)

Aspectos solicitados	Número de ítems	
Formas de aprender la información que consulta	5	Texto, video, organizadores gráficos, animaciones, medio para comunicarse con amigos, conocidos, familiares o docentes a través de internet.
Datos laborales	12	Situación laboral actual, sector económico, cargo que ocupa, antigüedad, tipo de contrato, ingresos mensuales, tiempo con el que cuenta para desarrollar las actividades académicas, debe trabajar para estudiar en la UNAD, origen de los recursos económicos con los que estudia en la UNAD.
Actividades deportivas o recreativas	9	Qué deporte practica, está interesado en formar parte de un grupo representativo en deportes.
Actividades artísticas o culturales	9	Qué actividad artística practica, está interesado en formar parte de un grupo representativo en artes y cultura.
Géneros de danza	4	Qué género de danza practica.
Emprendimiento	3	Cuenta con una empresa que dé respuesta a una necesidad social en su comunidad; si es así, qué necesidad cubre. Temas de emprendimiento en los que le gustaría recibir capacitación.
Estilo de vida saludable	2	Cuáles cree que son las causas más frecuentes del estrés. A través de qué estrategias le gustaría conocer el autocuidado.
Proyecto de vida	1	Qué temas le gustaría abordar en la UNAD para su crecimiento personal.
Medioambiente	5	Cómo define la educación ambiental.
Psicosociales	11	Respuesta a cambios, manejo del tiempo para actividades académicas, actitudes y reacciones ante dificultades, reacciones ante discusiones, satisfacción consigo mismo.
Competencias (alto, medio, bajo)	7	Lectura crítica, razonamiento cuantitativo, inglés, biología, física, química.

Nota. Elaborado por los autores sobre la base de UNAD (2020).

Figura 4

Proceso de la prueba de caracterización para estudiantes



2.4 Plan de Acción Pedagógica Contextualizada

Los resultados de la prueba de caracterización estudiantil se convierten en el insumo para la proyección y definición del Plan de Acción Pedagógica Contextualizada (PAPC), de tal manera que, en cada una de las zonas que conforman la UNAD, se planifica una serie de actividades que forman parte del proceso de inducción de los estudiantes. Cada una de estas actividades responden a esas necesidades detectadas con los resultados de la prueba para cada zona, y se programan y atienden en función de los contextos y dinámicas propias de cada región. Esta serie de actividades se consolidan en un recurso denominado PAPC, cuyo propósito es:

Contribuir al logro de las competencias requeridas para el óptimo desarrollo de su proceso de aprendizaje. Este servicio se realiza a través de la oferta de estrategias de carácter pedagógico y didáctico, en las áreas requeridas, acorde a los resultados de la prueba de caracterización. (UNAD, 2013, artículo 20)

A continuación, se refiere con más detalle la dinámica de organización y trabajo que se debe tener en cuenta para la programación de los PAPC en cada zona. Esta información fue tomada del documento de la UNAD, *Hoja de ruta acción y direccionamiento de VISAE (2017a)*.

La VISAE, en articulación con las escuelas, da los lineamientos para la creación, organización e implementación del material de trabajo académico en los módulos

que se requieren para el proceso que van a realizar los aspirantes y los estudiantes; de igual manera, establece los criterios para la asignación de los docentes seleccionados, previo diálogo con decanos y líderes nacionales de programa, quienes en conjunto y según perfiles escogen al personal idóneo para la ejecución de este. Este proceso inicia antes del periodo académico para los aspirantes y se lleva a cabo al inicio de cada periodo ofertado institucionalmente para los estudiantes. De esta manera, se establecen varios momentos de acompañamiento al estudiante, coordinados y programados desde la consejería académica.

- Reconocimiento por parte del aspirante en su nivel de competencias básicas.
- El estudiante identifica sus necesidades y conoce los beneficios que recibe como estudiante de primera matrícula desde la consejería académica.
- Una vez formalizada la matrícula, el estudiante presenta la prueba de caracterización.
- En las jornadas de acogida e integración unadista, el estudiante conoce los espacios pedagógicos planteados y organizados, y proyecta su participación en ellos. De igual forma, se le informa sobre los espacios virtuales donde tiene la posibilidad de nivelarse.

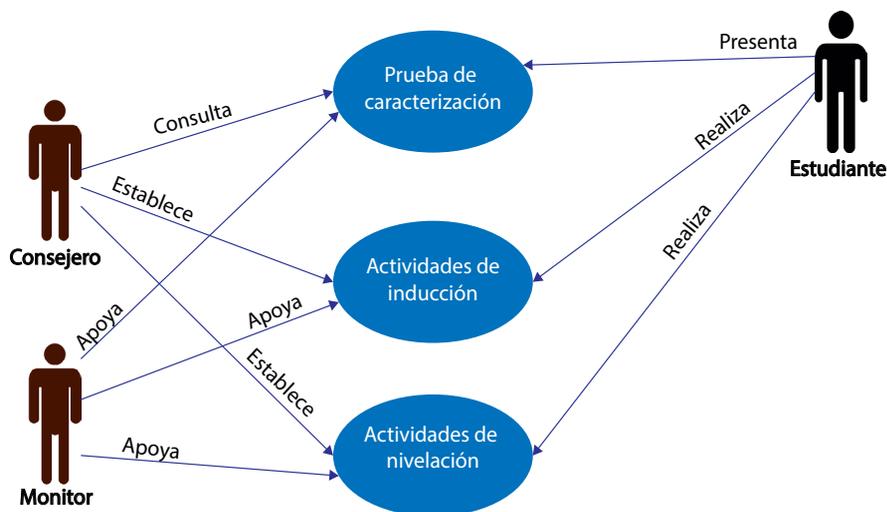
En general, y dentro de los elementos que se comparten con cada una de las zonas, se encuentran estrategias pedagógico-didácticas enfocadas en razonamiento cuantitativo, lectura crítica, inglés, competencias digitales básicas, estrategias de aprendizaje (hábitos de estudio, aprendizaje autónomo y colaborativo, etcétera). Además, de manera puntual y dependiendo de cada zona, se proyectan acciones complementarias.

Este servicio es diseñado, coordinado y evaluado por la consejería académica e implementado de manera articulada con las escuelas para contribuir a un desempeño y permanencia exitosa del estudiante (UNAD, 2017a, p. 4.)

Actualmente, entre las competencias que se cubren a través del PAPC están lectura crítica, pensamiento lógico-matemático, inglés, camino al éxito unadista y competencias digitales. Para cada una de estas competencias, se programan entre una y diez sesiones, cada una con duraciones que van desde 45 minutos hasta de una hora y media, de acuerdo con el contexto y los resultados de las pruebas de caracterización en cada zona y centro. La programación de estas sesiones se envía a los estudiantes a través de diferentes medios de contacto, tales como correo electrónico institucional, foros y chats de comunicación con consejeros, en donde constantemente se les invita a ser partícipes activos.

Figura 5

Proceso del Plan de Acción Pedagógica Contextualizada



2.5 Principales opciones del sistema de información para el estudiante

Una vez hecho el recorrido por el conjunto de acciones, estrategias, recursos y demás actividades planeadas desde la universidad para orientar el proceso de inducción al estudiante nuevo, a manera de resumen, se enuncian los pasos que debe seguir el estudiante en esa transición hacia el mundo de la virtualidad que emprende en la UNAD, no sin antes aclarar que el orden que se presenta no es necesariamente el que se da en la realidad y que tampoco está sujeto a que se haga de esa forma.

- Presentar la prueba de caracterización
- Asistir a las inducciones programadas
- Participar en los eventos programados en los PAPC

En los pasos anteriores, se deben centrar todas las diferentes oportunidades de mejora, propuestas, complementos y herramientas de apoyo que puedan darse con respecto a los procesos de inducción para estudiantes de primera matrícula de la UNAD, los cuales deben necesariamente contemplarse dentro del sistema de información propuesto.

Características del usuario

Tabla 3

Tipos de usuarios para el sistema de información

Tipo de usuario: administrador	
Formación	Profesional, con conocimientos de informática y manejo del proceso de inducción unadista
Actividades	Control, configuración, alimentación y administración del sistema en general
Tipo de usuario: consejero	
Formación	Profesional, con cargo de consejero en la UNAD
Actividades	Configuración de los pasos de inducción Registro y configuración de los eventos del PAPC Administración y atención del chat del sistema
Tipo de usuario: monitor	
Formación	Pregrado en desarrollo, estudiante activo de la UNAD
Actividades	Atención del chat del sistema
Tipo de usuario: estudiante	
Formación	Pregrado en desarrollo, estudiante activo de la UNAD
Actividades	Acceso al sistema para realizar todos los pasos e indicaciones en su proceso de inducción como estudiante de primera matrícula de la UNAD Uso del recurso del chat como recurso de comunicación sincrónica con su consejero y/o monitor asignado para su centro regional Consulta y acceso a los eventos y sesiones programados dentro del cronograma de inducción en su centro regional
Tipo de usuario: aspirante	
Formación	Bachiller y/o tecnólogo, no vinculado con la UNAD
Actividades	Acceso al sistema para conocer la universidad y su modalidad de formación virtual Uso del recurso del chat como recurso de comunicación sincrónica para conocer información específica acerca de la universidad y de sus programas Acceso al espacio de orientación vocacional, del que dispone la universidad

Restricciones

- Interfaz para ser usada con internet.
- Lenguajes y tecnologías en uso: PHP, Visual Studio Code, MySQL.
- Los servidores deben tener capacidad de concurrencia.
- El sistema se diseñará según un modelo cliente/servidor.
- El sistema debe contar con un diseño sencillo, interfaz amigable, independiente de la plataforma, *framework* y lenguajes de codificación.

Suposiciones y dependencias

- Se asume que los requerimientos iniciales identificados son estables.
- Tanto el *hardware* como el *software*, en los que se vaya a ejecutar el sistema, deben cumplir con los anteriores requisitos para garantizar la correcta ejecución del sistema.

3.2 Requisitos comunes de las interfaces

Interfaz de usuario

La interfaz de usuario debe basarse en un conjunto de ventanas con botones, menús desplegados y campos de texto abiertos. Ha de ser construida específicamente para el sistema propuesto y será visualizada desde un navegador de internet y de dispositivos inteligentes. Además, tiene que responder a factores de accesibilidad, inclusión y manejo de parámetros y protocolos de imagen institucional: colores, fuentes, tamaños, formas, logos y estructura.

Interfaces de hardware

El sistema deberá funcionar con equipos con las siguientes características:

- Conexión a internet: wifi, cable, red.
- Computadoras de escritorio o portátiles, con mínimo 2 GB de memoria RAM, almacenamiento disponible: 10 GB, Windows 10 o superior.
- Dispositivos inteligentes: tabletas, celulares, con sistema Android o iOS, 2 GB de memoria RAM.

Interfaces de software

- Sistema operativo: Windows 10 o superior, Android o iOS.
- Explorador: Chrome, Firefox, Edge.

Interfaces de comunicación

Los servidores, clientes y aplicaciones se comunicarán entre sí mediante protocolos estándares en internet, siempre que sea posible. Por ejemplo, para transferir archivos con datos de identificación de estudiantes, se puede establecer un protocolo de comunicación a través de *web service* con un nivel de seguridad determinado por los administradores del sistema de la UNAD.

3.3 Requisitos funcionales

Tabla 4

Requerimientos funcionales para el sistema de información

Identificación del requerimiento	RF-001
Nombre del requerimiento	Autenticación de usuarios
Características	Los usuarios deberán identificarse para acceder y usar el sistema.
Descripción del requerimiento	El <i>software</i> brindará acceso a aquellos usuarios que posean credenciales válidas y previamente se hayan registrado. Como mecanismo de seguridad, se debe implementar el doble factor de autenticación. Los roles que se deben implementar son estudiante, aspirante, consejero, monitor y administrador.
Prioridad del requerimiento	Alta
Identificación del requerimiento	RF-002
Nombre del requerimiento	Registrar usuarios
Características	Los usuarios deberán registrarse en el sistema para acceder a las opciones según su rol.
Descripción del requerimiento	Los usuarios podrán registrarse para obtener las credenciales de acceso y así acceder a funcionalidades exclusivas de cada uno de los roles implementados: estudiante, aspirante, consejero, monitor y administrador.
Prioridad del requerimiento	Alta
Identificación del requerimiento	RF-003
Nombre del requerimiento	Gestor de datos
Características	El sistema debe contemplar un sistema de bases de datos robusto y con capacidad de almacenamiento y procesamiento suficientes para el sistema. Debe ser compatible con los sistemas y bases de datos que se manejan en la plataforma tecnológica de la universidad.
Descripción del requerimiento	Se debe diseñar y modelar una base de datos que permita almacenar los datos e información que manejará el sistema. Dentro de ella se deben contemplar las siguientes entidades: roles, datos básicos, programaciones, pasos, comunicaciones (chat, eventos).
Prioridad del requerimiento	Alta
Identificación del requerimiento	RF-004
Nombre del requerimiento	Menús y opciones del sistema
Características	El sistema debe contar con un menú diferente para cada rol.
Descripción del requerimiento	El sistema debe adaptar el menú, sus opciones y presentación, dependiendo del rol del usuario. Debe configurarse para cada rol: estudiante, aspirante, consejero, monitor y administrador.
Prioridad del requerimiento	Alta

(continúa)

(continuación)

Identificación del requerimiento	RF-005
Nombre del requerimiento	Menú estudiante
Características	La presentación del sistema adaptada para el rol estudiante debe ser completamente parametrizable, flexible e intuitiva.
Descripción del requerimiento	Debe contemplar tres opciones básicas para el estudiante: <ul style="list-style-type: none"> • Información de la inducción: primeros pasos. • Informar sobre el proceso de estudio a través de la virtualidad, soportado en material multimedial. • Acceder a la programación de las actividades establecidas desde consejería, como resultado del análisis de la prueba de caracterización.
Prioridad del requerimiento	Alta
Identificación del requerimiento	RF-006
Nombre del requerimiento	Información para ingreso y activación de estudiantes
Características	Configuración de pasos para ingreso y activación de estudiantes, parametrizables de acuerdo con lo que establece la inducción en el proceso de consejería académica.
Descripción del requerimiento	El sistema debe contar con una primera opción para informar al estudiante sobre los pasos y tareas iniciales en la activación de su ingreso a la plataforma de aprendizaje de la universidad, los cuales se encuentran en el protocolo del proceso de inducción establecido desde consejería académica. Los consejeros deben tener la opción para alimentar el sistema con la información del protocolo de proceso de inducción, a través de pasos, materiales, accesos e instrucciones claramente identificados.
Prioridad del requerimiento	Alta
Identificación del requerimiento	RF-007
Nombre del requerimiento	Admin del sistema
Características:	<ul style="list-style-type: none"> • Parametrización de datos básicos del sistema. • Categorización de la información básica. • Flexibilidad en la consulta y edición de cada tabla básica. • Criterios de búsqueda que satisfagan las necesidades de los usuarios.
Descripción del requerimiento	Los datos básicos que se deben configurar para el sistema son zonas, centros regionales, periodos académicos, usuarios (estudiante, aspirante, consejero, monitor, administrador), programas académicos, pasos iniciales, eventos, preguntas, frases. Los únicos usuarios autorizados para alimentar y configurar los datos básicos son aquellos que cuenten con el rol de administrador. Todo usuario con rol de consejero puede editar y alimentar únicamente los datos de "eventos" del sistema. Los demás roles no tienen acceso a la configuración de datos básicos.
Prioridad del requerimiento	Alta
Identificación del requerimiento	RF-008

(continúa)

(continuación)

Nombre del requerimiento	Notificaciones
Características	Los estudiantes serán informados dentro de la aplicación que tienen eventos próximos programados dentro del proceso de inducción, en su CEAD. La opción debe ser fácil, accesible, segura y parametrizable.
Descripción del requerimiento	Los usuarios con rol de estudiante contarán con una opción visible que les alerta de notificaciones sobre los eventos que tienen programados en el mismo día que ingresan al sistema. Al acceder a esas notificaciones, el estudiante puede ver la información de fecha, hora, objetivo y tipo del evento que debe tomar. Los consejeros deben tener la opción para alimentar los eventos que tienen programados en su CEAD, de acuerdo con las acciones del PAPC organizado para su zona, desde la consejería académica.
Prioridad del requerimiento:	Media
Identificación del requerimiento	RF-009
Nombre del requerimiento	Chat
Características	El sistema debe contar con un recurso de comunicación sincrónica, con características como sistema chat, recurso incluido y propio del sistema, rápido y de fácil acceso.
Descripción del requerimiento	Establecer dentro del sistema un recurso tipo chat, propio, que puedan utilizar los estudiantes para su comunicación y contacto directo con sus consejeros y/o monitores asignados para la atención virtual en el CEAD.
Prioridad del requerimiento	Media

3.4 Requisitos no funcionales

Tabla 5

Requerimientos no funcionales para el sistema de información

Identificación del requerimiento	RNF-001
Nombre del requerimiento	Rendimiento
Características	Garantizar que el <i>software</i> tenga la capacidad de cómputo necesaria para responder en forma eficiente a cada uno de los procesos implementados.
Descripción del requerimiento	El sistema de información debe poder dar acceso a una población aproximada de estudiantes de primera matrícula, quienes usarán los recursos del sistema dispuestos para el desarrollo de actividades y el conocimiento de la virtualidad. Garantizar que los servicios de consulta de la información respondan en los tiempos apropiados para las infraestructuras de tipo web. Evitar bloqueos de la interfaz de usuario, comunicación con sistemas externos o demoras en el acceso a la base de datos.

(continúa)

(continuación)

Prioridad del requerimiento	Alta
Identificación del requerimiento	RNF-002
Nombre del requerimiento	Seguridad
Características	Implementación de mecanismos orientados a la mitigación de riesgos y protección de la información de ingreso y de contenido del <i>software web</i> .
Descripción del requerimiento	El ingreso al sistema se debe validar a partir del número de identificación del usuario, el cual será el dato de conexión con los servidores de la universidad. El sistema debe contar con un sistema de doble factor de autenticación que permita confirmar, de manera segura, la autenticación del usuario que usará el sistema. Los únicos datos sensibles que se deben traer de los servidores de la universidad son código del estudiante y correo electrónico. Construir un código de calidad orientado a evitar filtración de datos y técnicas de <i>hacking</i> que amenacen la seguridad de los datos. Implementar mecanismos de autenticación alternos que, combinados, permitan identificar que el usuario es quien dice ser.
Prioridad del requerimiento	Alta
Identificación del requerimiento	RNF-003
Nombre del requerimiento	Disponibilidad
Características	Garantizar que la operación del sistema de información tenga un nivel de servicio 24/7.
Descripción del requerimiento	El <i>software</i> debe permitir la atención sincrónica de los estudiantes, vía chat, en cada una de las seis zonas en las que se encuentran distribuidos los diferentes centros en el país. El <i>software web</i> debe garantizar que la operación no será interrumpida a pesar de que uno o varios de sus componentes falle. Se debe asegurar que el usuario pueda acceder a la aplicación y que esté informado de forma oportuna, clara y veraz sobre las tareas de mantenimiento en caso de que tengan lugar estas actividades.
Prioridad del requerimiento	Alta
Identificación del requerimiento	RNF-004
Nombre del requerimiento	Portabilidad
Características	Garantizar características que debe presentar el <i>software</i> para facilitar su traslado a otras plataformas o entornos.
Descripción del requerimiento	El sistema debe tener un diseño web <i>responsive</i> o adaptativo, dado que debe orientarse hacia un desarrollo que busque su correcta visualización en distintos dispositivos, especialmente celulares y dispositivos inteligentes, pues la mayoría de los usuarios serán estudiantes de primera matrícula que cuentan con estos dispositivos como mayor recurso de acceso a los espacios virtuales.
Prioridad del requerimiento	Alta

(continúa)

(continuación)

Identificación del requerimiento	RNF-005
Nombre del requerimiento	Mantenibilidad
Características	Garantizar que la operación técnica y funcional se encuentre documentada.
Descripción del requerimiento	El <i>software</i> web debe contar con la documentación técnica y funcional necesaria para facilitar y satisfacer la configuración en los aspectos técnico y funcional.
Prioridad del requerimiento	Alta

4. CONCLUSIONES

Desde el punto de vista de la universidad, se ha logrado establecer una estructura que permite identificar las necesidades de sus estudiantes nuevos, así como su situación académica y psicosocial, para que se puedan dar respuestas y soporte para que los estudiantes logren las competencias requeridas para el inicio de su proceso de formación. Es importante destacar que esta respuesta obedece al contexto del estudiante y se busca atenderle de acuerdo con las características propias del centro, del recurso humano disponible y de asistencias que históricamente se han dado.

Desde el punto de vista del estudiante, en sus inicios en el ambiente de formación virtual, el hecho de tener que abordar los diferentes elementos que se le ofrecen, tales como el proceso de inducción, la prueba de caracterización y la asistencia a encuentros programados en el PAPC, puede hacer que requiera de mayor control para el agendamiento de todas esas actividades y así lograr una programación efectiva para su asistencia y atención requeridas, con el fin de sacar el mayor provecho de cada una de ellas.

Como en cualquier situación nueva, el estudiante se ve envuelto en una serie de tareas que debe asumir para lograr, antes que nada, una mayor asistencia para sacar provecho, lograr respuestas a sus inquietudes, calmar su posible ansiedad, buscar soporte y orientación efectiva de acuerdo con sus características propias de tiempo, presaberes, necesidades particulares y ajustes de tiempos y espacios.

De acuerdo con los puntos anteriores, el sistema de información debe ser escalable, dada la cobertura nacional que tiene la universidad y el proceso permanente de creación de nuevos programas y sedes, incluyendo las internacionales. Además, el sistema de información tiene que cubrir tres aspectos esenciales:

- Ofrecer al estudiante la posibilidad de desarrollar de manera intuitiva la ruta inicial que establece la universidad, la cual constantemente se revisa y se

ajusta. En esta ruta se le tiene que dar la posibilidad de recibir una explicación, con apoyo multimedial, de la actividad que va a realizar, así como direccionarlo al enlace del espacio virtual en donde debe realizarla.

- Disponer de un espacio que permita orientar al estudiante con respecto a las características de la formación virtual, presentando un contexto en el que pueda comparar la formación tradicional a la que está acostumbrado y la formación virtual que inicia con la UNAD. Este espacio debe también ser atractivo, ofrecer recursos y herramientas actuales y multimediales que aumenten su motivación y atención.
- Brindar al estudiante un recurso tecnológico tipo agenda de actividades, en el cual pueda consultar la programación de todos los eventos que se programan como resultado de la prueba de caracterización y que forman parte del PAPC específico de la sede y zona a la que pertenece.

REFERENCIAS

- Acevedo Velandia, S. J. (2020). La Consejería Académica, apuesta organizacional y académica por la permanencia exitosa en la UNAD. *Congresos CLABES*, 1272-1278. <https://revistas.utp.ac.pa/index.php/clabes/article/view/2774>
- Kendall, K. E., & Kendall, J. E. (2005). *Análisis y diseño de sistemas* (6.ª ed.). Pearson Educación.
- Monferrer Agut, R. (2001). *Especificación de requisitos software según el estándar de IEEE 830*. Universitat Jaume I, Departament d'Informàtica. <http://textos.pucp.edu.pe/pdf/3134.pdf>
- Paulk, M. C., Weber, C. V., & Curtis, B. (1997). *The capability maturity model: Guidelines for improving the software process*. Addison-Wesley.
- Pressman, R. (2012). *Ingeniería del software. Un enfoque práctico* (7.ª ed.). McGraw-Hill.
- Senn, J. A. (2008). *Análisis y diseño de sistemas de información* (6.ª ed.). McGraw-Hill.
- Sommerville, I. (2010). *Ingeniería del software* (9.ª ed.). Addison-Wesley.
- Universidad Nacional Abierta y a Distancia. (2013). *Acuerdo 0029 del 13 de diciembre de 2013 por el cual se expide el Reglamento Estudiantil de la Universidad Nacional Abierta y a Distancia (UNAD) y se dictan otras disposiciones*. https://sgeneral.unad.edu.co/images/documentos/consejoAcademico/acuerdos/2013/COAC_ACUE_029_20131213.pdf

Universidad Nacional Abierta y a Distancia. (2015). *Programa de acogida e integración a la Comunidad Unadista*. Vicerrectoría de Servicios a Aspirantes, Estudiantes y Egresados.

Universidad Nacional Abierta y a Distancia. (2017a). *Hoja de ruta acción y direccionamiento de VISAE*.

Universidad Nacional Abierta y a Distancia. (2017b). *Proceso: Ciclo de vida del estudiante. Procedimiento: Consejería académica*. <https://sig.unad.edu.co/documentos/sgc/procedimientos/P-7-7.pdf>

Universidad Nacional Abierta y a Distancia. (2020). *Documento Plan de Acción Pedagógico Contextualizado*.

Weitzenfeld, A. (2005). *Ingeniería de software orientada a objetos con UML, Java e internet*. Thomson.

DATOS DE LOS AUTORES

MÓNICA CALVA

Magíster en Ciencias y Tecnologías de la Computación por la Universidad Técnica Particular de Loja (UTPL), Ecuador. Ingeniera de Sistemas Informáticos y Computación por la misma casa de estudios. Sus intereses de investigación comprenden informática médica, ontologías, base de datos y sistemas distribuidos de soporte de decisiones clínicas. Actualmente, se desempeña como coordinadora del área de Sistemas del Hospital UTPL, Ecuador.

RUTH MARÍA REÁTEGUI ROJAS

Doctora por la École de technologie supérieure (ÉTS), Canadá. Actualmente, es profesora en el Departamento de Ciencias de la Computación y Electrónica de la Universidad Técnica Particular de Loja (UTPL), Ecuador. Desde el 2019, es coordinadora del grupo de investigación Inteligencia Artificial Aplicada de la UTPL. Sus intereses de investigación incluyen minería de texto, minería de datos, aprendizaje automático, aprendizaje profundo con aplicaciones en salud y educación.

GRACIELE TONIAL

Doctora en Ingeniería y Gestión del Conocimiento por la Universidade Federal de Santa Catarina (UFSC), Brasil. Es profesora de la Universidade do Oeste de Santa Catarina (UNOESC) e investigadora en la línea de estrategia organizacional, emprendimiento e innovación.

NATHALIA BERGER WERLANG

Doctora en Administración por la Universidade Federal de Santa Catarina (UFSC), Brasil. Es profesora del Departamento de Ciencias de la Información de la UFSC e investigadora en el campo de la gestión de la información, la capacidad de absorción y la innovación.

ALESSANDRA CASSOL

Doctora en Administración de Empresas, área de concentración en Estrategia y Organizaciones, por la Universidade Nove de Julho (UNINOVE), Brasil. Es profesora adjunta de la Universidade Federal Rural do Rio de Janeiro (UFRRJ), e investigadora en el área de aprendizaje organizacional, innovación y emprendimiento.

DANISSON LUIZ DOS SANTOS REIS

Doctor del Programa de Posgrado en Ingeniería y Gestión del Conocimiento por la Universidade Federal de Santa Catarina (UFSC), Brasil. MBA en Ingeniería e Innovación. Su línea de investigación corresponde al emprendimiento, innovación y sostenibilidad. Actualmente, forma parte del grupo de investigación y extensión VIA-ESTACIÓN CONOCIMIENTO del Departamento de Ingeniería y Gestión del Conocimiento de la UFSC.

GREGÓRIO VARVAKIS

Doctor en Ingeniería de Manufactura por la Loughborough University of Technology, Reino Unido. Magíster en Ingeniería de Producción por la Universidade Federal de Santa Catarina (UFSC), Brasil. Licenciado en Ingeniería Mecánica por la Universidade Federal do Rio Grande do Sul (UFRGS), Brasil. Actualmente, es profesor de la UFSC en el Departamento de Ingeniería del Conocimiento, y en los Programas de Posgrado de Ingeniería y Gestión del Conocimiento y Ciencias de la Información. Tiene experiencia en el área de gestión, con énfasis en gestión de procesos, gestión del conocimiento y gestión de organización de servicios; trabaja principalmente en los siguientes temas: innovación, gestión del conocimiento, productividad, mejora continua, tecnologías de la información y flujo de información.

PAULO ROBERTO DE MOURA

Especialista en dirección de personas, ponente, consultor de empresas, enfocado en intraemprendimiento. Máster universitario en Ingeniería de Gestión del Conocimiento. Magíster en Dirección de Personas por la Fundación de Administradores del Estado de Santa Catarina. Tiene un posgrado en Gestión de Tutoría por UNIASSELVI, Brasil.

NERI DOS SANTOS

Doctor en Ergonomía de la Ingeniería por el Conservatoire National des Arts et Métiers (CNAM), Francia. Magíster en Ergonomía por la Universidad de París XIII, Francia. Es profesor sénior del Programa de Posgrado en Ingeniería y Gestión del Conocimiento de la UFSC y CEO del Instituto STELA. Fue presidente de ABEPRO (1992-1993 y 1994-1995). Sus áreas de investigación y enseñanza se centran en la

transformación digital en la ingeniería y la gestión del conocimiento. Fue clasificado por el Ranking for Scientist del Scientific Index como uno de los investigadores de la UFSC más citados en el mundo (duodécimo lugar).

RENATO CORONADO ÁLVAREZ

Ingeniero de Sistemas por la Universidad de Lima, Perú, especializado en Ingeniería de *Software*. Actualmente, es *junior full stack developer* en Grupo Intercorp.

ADRIAN ERNESTO LINO VILLARAN

Bachiller de la Carrera de Ingeniería de Sistemas por la Universidad de Lima, Perú. Actualmente, se desempeña como programador móvil de iOS y Android en una empresa multinacional. Apasionado por el mundo del emprendimiento e innovación, logró obtener el segundo puesto en el Concurso Primer Paso 2020 y ConectaRSE para Crecer 2021 en la categoría de reactivación económica. Posee habilidad técnica e interés de investigación en el desarrollo de *software*, dispositivos móviles, internet de las cosas y electrónica.

OLDA BUSTILLOS ORTEGA

Magíster en Auditoría en Sistemas de la Información. Directora y docente de la Escuela de Ingeniería Informática, Facultad de Ingeniería y Arquitectura, Universidad Internacional de las Américas, Costa Rica.

JAVIER ROJAS SEGURA

Magíster en Business Research por el Tecnológico de Costa Rica. Es docente investigador en la Escuela de Ingeniería Informática en la Universidad Internacional de las Américas, Costa Rica.

GANÍMEDES T. ROSALES REYES

Magíster en Administración de Negocios, ingeniero industrial, consultor independiente y docente universitario. Tiene más de catorce años de experiencia en posiciones ejecutivas y de dirección, en las que ha realizado y liderado actividades vinculadas a la formulación y ejecución de estrategias, mejora de rendimiento organizacional e implementación de tecnologías de la información y comunicaciones (TIC), tanto en el sector público como privado. Su experiencia en los campos de la dirección y mejora de rendimiento organizacional incluye el diseño de modelos de negocios para gestionar Oficinas de Proyectos (PMO) y Oficinas de Gestión de la Excelencia (*Business Process Office*), así como el diseño e implementación de arquitecturas empresariales y de

procesos, y sistemas de gestión de mejora continua bajo estándares internacionales. Cuenta con certificaciones PMP (*Project Management Professional*), ITIL v3, Lean Six Sigma Black Belt - ASQ y especializaciones en las normas ISO 9001.

AUGUSTO ENRIQUE HAYASHIDA MARCHINARES

Candidato a doctor en Ingeniería de Sistemas e Informática por la Universidad Nacional Mayor de San Marcos, Perú. Magíster en Telecomunicaciones y en Ingeniería de Sistemas e Informática por la misma universidad. Ingeniero de sistemas e informática por la Universidad Alas Peruanas. Jefe de Proyectos Informáticos, certificado como Project Management Professional - PMP®, PM4R Expert Certificate - BID, Certified Information Security Management Professional ISO 27001, Professional Agile Leadership I - PAL I, Professional Scrum Master I - PSM I, Agile Scrum Master, SCRUM Study Certified Trainer (SCT™), Lean Management Professional, OKR Certified Professional, OKR Advance Professional, OKR Certified Associate, Innovation Management Certified Professional (IMCP). En el ámbito académico, su área de investigación es la gestión de proyectos informáticos y machine learning aplicado a la gestión de proyectos informáticos.

XAVIER ALBERTO GUTIERREZ CORAL

Candidato a doctor en Ingeniería de Sistemas e Informática por la Universidad Nacional Mayor de San Marcos, Perú. Magíster en Dirección de Tecnologías de Información por la Universidad ESAN, Perú, con honores *summa cum laude*. Máster en Gestión de las Tecnologías de la Información por la Universitat Ramon Llull, España. Bachiller en Administración y Gerencia por la Universidad Ricardo Palma, Perú. En el ámbito académico, su área de investigación se ha enfocado en la analítica avanzada, incluyendo minería de datos, textos e imágenes, así como en las redes *blockchain*, entre las que se encuentran mecanismos de gobierno descentralizado, criptografía, tokenización de activos e integración de actividades *on-chain* y *off-chain* para procesos transaccionales complejos.

JULIO VLADIMIR QUISPE SOTA

Candidato a doctor en Ingeniería de Sistemas e Informática por la Universidad Nacional Mayor de San Marcos, Perú. Posgrado en Ingeniería Computacional y Sistemas Inteligentes en la Universidad del País Vasco, España. Graduado en Ingeniería Informática y de Sistemas por la Universidad Nacional de San Antonio Abad del Cusco (UNSAAC), Perú. Actualmente, es docente en esta universidad, y asistente administrativo en la Corte Superior de Justicia de Cusco, en la Unidad de Servicios Judiciales, donde fue reconocido en varias oportunidades por los proyectos de desarrollo de *software*. Obtuvo el tercer lugar en la feria de innovación tecnológica

organizada por el Poder Judicial en el 2022 con el proyecto Sistema de Gestión de Expedientes Judiciales para el Pool de Peritos. Sus áreas de investigación son el *machine learning*, los algoritmos paralelos, la criptografía y la ingeniería de *software*.

ANGELO RODRIGO TACO JIMENEZ

Estudiante de la Carrera de Ingeniería de Sistemas en la Universidad de Lima, Perú. Se inclina por la especialización en Ingeniería de *Software*. Se interesa en la investigación en desarrollo web y aplicaciones digitales.

ROCIO ISABEL BENITES LOJA

Estudiante del Programa de Estudios de Ingeniería de Sistemas de la Facultad de Ingeniería de la Universidad Católica Sedes Sapientiae (UCSS), Perú. Ha sido parte del equipo de la Corporación VG Architec S. A. C., donde se desempeñó en la Oficina de Desarrollo e Investigación.

MARCO ANTONIO CORAL YGNACIO

Candidato a doctor en Ingeniería de Sistemas por la Universidad Nacional Mayor de San Marcos (UNMSM), Perú. Magíster en Ingeniería de Sistemas y Computación por la misma casa de estudios. Experto en temas de transformación digital en universidades. Ha sido responsable técnico del proyecto Cero Papel en la UNMSM, trabajando con sistemas de gestión documentaria con firma digital; implementación de documentos digitales, tales como grados, títulos y otros; e integración de sistemas y generación de servicios para la universidad. Ha sido jefe de la Unidad de Tecnología Educativa, jefe de la Unidad de Servidores y Sistemas de Información de la Red Telemática de la UNMSM y jefe de la Oficina de Calidad y Acreditación Académica de la FISI-UNMSM, responsable de la acreditación de los programas de posgrado e implementación de la norma ISO 9001:2015. Es investigador en temas de transformación digital, ha publicado en revistas científicas y es evaluador de proyectos de innovación tecnológica para CONCYTEC.

PILAR ALEXANDRA MORENO

Magíster en *e-learning* por la UNAD Florida, Estados Unidos. Ingeniera de sistemas por la Universidad de Boyacá, Colombia. Especialista en Pedagogía para el Desarrollo del Aprendizaje Autónomo. Es docente asistente de la UNAD en el Programa de Ingeniería de Sistemas. Sus líneas de investigación son, entre otras, la ingeniería del *software*, la programación de sistemas, ambientes virtuales de aprendizaje, interacción humano-computadora, gamificación, consejería virtual, *big data* y agentes inteligentes. Ha sido reconocida por Colciencias, y es par evaluador

Colciencias - Programa Nacional de Electrónica, Telecomunicaciones e Informática. Cuenta con experiencia en análisis, diseño y desarrollo de sistemas de información en áreas empresariales, administrativas y académicas.

JUAN OLEGARIO MONROY VÁSQUEZ

Magíster en Tecnologías de la Información Aplicadas a la Educación. Ingeniero electrónico con especialización en Multimedia Educativa por la Universidad Tecnológica y Pedagógica de Colombia. Es docente asistente de la UNAD, en el programa de Ingeniería Electrónica de la cadena ETR, en las áreas de control, programación, *software* avanzado para ingeniería, *software* para ingeniería, teoría general de sistemas, microprocesadores y microcontroladores, algoritmos e introducción a la programación, electrónica, circuitos digitales, arquitectura de computadores, lógica matemática, programación G y Matlab. Cuenta con experiencia en programación a bajo nivel y trabajo con multimedia.

DOSIER

Retos y oportunidades en la accesibilidad de datos

Modelamiento de tópicos utilizando mensajes de Twitter relacionados con el cáncer cervical

O papel da capacidade absorptiva na inovação: evidências de práticas desenvolvidas por empresas de base tecnológica inseridas em redes colaborativas

Capital humano: a chave da gestão do conhecimento em organizações públicas - Uma abordagem integrativa

ARTÍCULOS DE INVESTIGACIÓN

Reforzamiento de habilidades aritméticas de fracciones y divisiones en niños de primaria mediante un juego serio educativo

Sistema de monitoreo de ambiente para el control de las condiciones ambientales de las granjas acuícolas de la selva del Perú basado en internet de las cosas

Protocolo básico de ciberseguridad para pymes

Aplicación de *machine learning* para campañas de marketing en la banca comercial

ARTÍCULOS DE REVISIÓN

Diseño de interfaces de sistemas interactivos utilizando técnicas de machine learning: una revisión del diseño y la usabilidad

Dispositivos *wearables* y los riesgos a la privacidad: una revisión de la literatura

Una revisión de las implementaciones de sistemas para la identificación de tendencias de la diabetes

ARTÍCULO DE DIVULGACIÓN

Formulación de requisitos para un sistema de información de apoyo al proceso de inducción de estudiantes universitarios: caso UNAD