UNIVERSIDAD LIBRE
Vigilada Mineducación

INGENIO LIBRE

Universidad Libre - Facultad de Ingeniería
Año 15 - ISSN 2322-8415

No. 26
AÑO 2026



"Fundar, con el concurso de los más competentes profesores, la Universidad Libre".

Rafael Uribe Uribe, 1903.

- Estrategia de innovación digital para la gestión logística de la cadena de suministro del cacao y el fortalecimiento de su competitividad en Santander.
- Arquitectura de referencia para plataforma blockchain de activos digitales con soporte de criptografía poscuántica y gobernanza descentralizada.
- Phishing de Zero-day: Detectar lo desconocido.
- Sistema de Identidad Digital Post-Cuántica para el Acceso Seguro a Modelos de Inteligencia Artificial.

Foto tomada de:
https://proclamadelpacifico.com/
Restauración de foto con ChatGPT

Oficina de Comunicaciones

UNIVERSIDAD LIBRE®
Vigilada Mineducación

María Elizabeth García González
Presidenta Nacional

César López Meza
Rector Nacional

Fernando Arturo Salinas Suárez
Rector Seccional

Nidia María Jirado Tajan
Decana de la Facultad de Ingeniería

Ricardo Andrés Santa Quintero
Editor Revista Ingenio Libre

EDITORIAL

"...Fundar, con el concurso de los más competentes profesores, la Universidad Libre."

Rafael Uribe Uribe. 1903

Si, de nuevo es momento de cimentar los valores de la educación unilibrista, son la mejor y más alta expresión de nuestros directivos, profesores y estudiantes. Nuestra comunidad educativa hace gala del más excelso fundamento Liberal. Es nuestro activo más valioso, el que nos diferencia, y nuestra ventaja competitiva.

Hombres y mujeres, los más competentes, dispuestos a continuar construyendo un mundo mejor, aportando desde la ingeniería soluciones a los problemas de diversa índole que se presentan.

Estas soluciones se expresan de diversas formas, la más clara y que requiere mayor claridad en las ideas, es la forma escrita, y hoy traemos en este numero 26 de la Revista Ingenio Libre, algunas de las propuestas que, desde la aulas, estudiantes y profesores, tiene para presentar a la sociedad.

El lector podrá encontrar en este número diversos temas asociados a la seguridad post cuántica, la identidad digital, gestión de documentos con blockchain, visión por computador, gestión de cadena de suministro, detección de phishing, estándares de seguridad de la información para la era cuántica y gobernanza descentralizada.

Todos estos temas apuntan a entender en la era digital viejos dilemas sobre la verdad, las relaciones humano – máquina y el mundo como nuestro hogar sin fronteras.

Agradezco a los autores por sus aportes a la discusión de estos temas que siguen siendo relevantes y muy importantes para un mundo mejor, haciendo gala del pensamiento unilibrista.

Ricardo Andrés Santa Quintero
Editor en jefe
Revista Ingenio Libre

SALUDO DE LA DECANA

Desde las aulas y los laboratorios, la Decanatura de Ingeniería saluda a todos los estudiantes, docentes y lectores que hacen de esta publicación una herramienta fundamental en el proceso de difusión de ideas y resultados de la investigación. También se extiende un saludo al Comité Editorial de este número de la revista Ingenio Libre.

El pragmatismo con el que se educa en ingeniería es una pieza clave para el desarrollo económico y social del país. Con los aportes y desarrollos de la ingeniería, la industria mejora sus procesos de producción, realiza innovaciones de producto, proceso y mercadeo, siendo ellos garantía de crecimiento económico y fortaleza de una mejor sociedad. En las aulas y laboratorios de la universidad se entiende y se practica la generación de valor, la optimización y la eficiencia, diseñando respuestas útiles y funcionales a problemáticas que vive nuestra sociedad.

El editorial de esta revista nos habla de la síntesis como respuesta filosófica a los retos que presenta día a día nuestro país y la región. Se deja claro que las soluciones surgen del proceso de contraste entre diversas ideas y diseños, y que durante el proceso no se trata de desechar las ideas, aunque sean contrarias a la que prevalece, sino que, por el contrario, se busca, mediante consenso, la unión de las partes para constituir un todo. Esto sintetiza lo más significativo, relevante y consistente de los diversos puntos de vista existentes.

La síntesis no significa uniformidad sino articulación: unir ciencia, tecnología y el componente humano nos permite comprender la realidad en la complejidad actual. De allí la afirmación “el todo es más que la suma de sus partes”, y es gracias al proceso de síntesis que se logran objetivos comunes, soluciones a futuro y, sobre todo, un camino que direcciona todos los esfuerzos, los concentra y genera valor para la sociedad.

Agradezco a los docentes que se encargan de acompañar a nuestros estudiantes en estos procesos y que tienen una idea común: el pragmatismo orientado por el consenso y la síntesis de ideas, ese es el espíritu de nuestras aulas y laboratorios.

Nidia María Jirado
Decana de la Facultad de Ingeniería

ARQUITECTURA DE REFERENCIA PARA PLATAFORMA BLOCKCHAIN DE ACTIVOS DIGITALES CON SOPORTE DE CRIPTOGRAFÍA POSTCUÁNTICA Y GOBERNANZA DESCENTRALIZADA

Marlly Daniela Pineda Villamil

1. RESUMEN

Actualmente los sistemas digitales han aumentado su usabilidad creando la necesidad de garantizar seguridad, integridad y trazabilidad de los activos digitales firmados. Sin embargo, los mecanismos criptográficos actuales presentan diferentes vulnerabilidades lo cual limita el avance de la computación cuántica, generando riesgos para la información en algunas plataformas digitales, este articulo presenta una arquitectura de referencia basada en tecnología blockchain permissionada, implementando mecanismos de criptografía postcuantica y principios de ingeniería de software, el objetivo es fortalecer la seguridad de los activos digitales utilizando modelos de gobernanza descentralizada, arquitectura orientada y enfoques basados en Domain Driven Desing que permite mejorar la confiabilidad, trazabilidad e inmutabilidad de la información para prevenir los sistemas de futuras amenazas tecnológicas.

2. INTRODUCCIÓN

El impulso digital ha incrementado los usos de activos digitales en diferentes sectores, frecuentemente empresariales, académico, certificados digitales, contratos inteligentes y registros electrónicos que requieren mecanismos que garanticen autenticidad, integridad y disponibilidad.

La mayoría de plataformas utilizan algoritmos criptográficos como RSA y ECC para proteger la información. En consecuencia, el avance de la computación cuántica representa una amenaza para estos mecanismos debido a la capacidad para resolver problemas matemáticos complejos en tiempos menores a los sistemas tradicionales. La tecnología blockchain se presenta como una alternativa para fortalecer seguridad en los sistemas digitales ya que contiene características como la descentralización, trazabilidad e inmutabilidad. Simultáneamente, la criptografía postulantica busca desarrollar algoritmos protegidos a ataques cuánticos para proteger la información a largo plazo.

El artículo propone una arquitectura de referencia para una plataforma blockchain orientada a la gestión de activos digitales firmados, implementando mecanismos de criptografía postcuantica y principios de gobernanza descentralizada.

3. PROBLEMA

Actualmente los sistemas de gestión de activos digitales principalmente depende de

arquitecturas centralizadas y algoritmos criptográficos actuales que pueden ser vulnerables frente a futuras amenazas de la computación cuántica. Lo que representa riesgos relacionados con la manipulación de información, perdida de integridad y falencias en la autenticación de los datos.

Muchas soluciones blockchain actuales se enfocan principalmente en transacciones financieras y no integran mecanismos de criptografía ni lineamientos de ingeniería de software, limitando su capacidad de adaptación de escenarios tecnológicos.

Debido a esto, surge la necesidad de diseñar una arquitectura de referencia segura, modular y escalable que permita garantizar la seguridad en activos digitales firmados mediante blockchain y criptografía postcuantica.

4. OBJETIVO GENERAL

Desarrollar una arquitectura de referencia basada en tecnología blockchain que permita la gestión segura de activos digitales firmados, garantizando su integridad, trazabilidad e inmutabilidad mediante la integración de mecanismos de criptografía postcuántica y principios de ingeniería de software.

5. METODOLOGÍA

La investigación se basa en un enfoque aplicado, experimental y cuantitativo que busca proponer una solución tecnológica que responda a problemáticas reales relacionadas con la seguridad y protección de activos digitales. La metodología contempla fases:

Fase 1. Análisis de Requisitos
- Definir requisitos funcionales y no funcionales.
- Identificar actores y los respectivos casos de uso.
- Identificar atributos de calidad.

Fase 2. Diseño Arquitectónico
- Identificar dominios mediante Domain-Driven Design.
- Diseño de arquitectura por capas.
- Integración de arquitectura orientada.

Fase 3. Diseño de seguridad
- Implementación de algoritmos post-cuanticos.
- Definición del modelo Zero-Trust
- Diseño de mecanismos de validación criptográfica.

Fase 4. Modelo y Prototipo
- Elaborar diagramas UML
- Diseñar prototipo funcional.
- Configurar nodos blockchain.

Fase 5. Validacion
- Pruebas de funcionalidad.
- Validación de integridad y trazabilidad.
- Pruebas de seguridad.

6. ARQUITECTURA DE REFERENCIA PROPUESTA

La arquitectura se fundamenta en principios de modularidad, escalabilidad y seguridad distribuida.

Capa de Presentación
Se encarga interacción con el usuario mediante interfaces web o aplicaciones.

Capa de Aplicación
Utiliza microservicios responsables de gestión de activos digitales, firmas digitales y validación criptográfica.

Capa Blockchain
Permite el registro inmutable de los activos digitales mediante una red blockchain permissionada.

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 7-10 Universidad Libre - Facultad de Ingeniería

Capa de Seguridad
Implementa algoritmos postcuanticos como CRYSTALS-Kyber y CRYSTALS-Dilithium para fortalecer la seguridad de la información.

Capa de Infraestructura
Incluye nodos blockchain, bases de datos y servicios de despliegue.

7. RESULTADOS ESPERADOS

Esta arquitectura permitirá:

- Mejorar trazabilidad de activos digitales
- Garantiza integridad e inmutabilidad de la información.
- Reduce riesgos asociados a la manipulación de datos.
- Previene amenazas de computación cuántica
- Facilitar la interoperabilidad entre plataformas digitales.

8. CONCLUSIONES

La integración blockcain con criptografía postcuantica presenta una alternativa para fortalecer la seguridad de los sistemas digitales modernos. La arquitectura propuesta permite combinar mecanismos de descentralización, trazabilidad e inmutabilidad con algoritmos criptográficos preparados para el futuro tecnológico.

El uso de principios de ingeniería de software y modelos de gobernanza descentralizada facilita la construcción de plataformas escalables, seguras y mantenibles.

Para concluir esta propuesta constituye una base de referencia para futuros desarrollos orientados a la seguridad de activos digitales.

REFERENCIAS BIBLIOGRÁFICAS

[1] Reddy, N. R., Suryadevara, S., Raghavendra Reddy, K. G., Umamaheswari, R., Guttula, R., & Kotoju, R. (2025). Quantum secured blockchain framework for enhancing post quantum data security. Scientific Reports, 15, 31048. https://doi.org/10.1038/s41598-025-16315-8

[2] Mallick, T., Zeldin, M., Nita-Rotaru, C., & Cenk, M. (2025). Quantum disruption: How post-quantum attackers reshape blockchain security and performance. arXiv. https://arxiv.org/abs/2512.13333

[3] National Institute of Standards and Technology. (2025). Post-quantum cryptography: NIST standardization process. https://doi.org/10.6028/NIST.IR.8545

[4] Martínez, J., Pérez, D., & López, A. (2022). Diseño de una arquitectura basada en blockchain para la protección de documentos digitales. Universidad Nacional de Colombia. https://repositorio.unal.edu.co/handle/unal/80654

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 7-10 Universidad Libre - Facultad de Ingeniería

LARGE LANGUAGE MODELS EN SEGURIDAD INFORMÁTICA: OPORTUNIDADES, RIESGOS Y CASOS DE APLICACIÓN

Large Language Models in Cybersecurity: Opportunities, Risks, and Use Cases

Juan David Molano Sarmiento
Universidad Libre, Bogotá, Colombia, juand-molanos@unilibre.edu.co

Fecha de recepción: dd/mm/año
Fecha de aceptación del artículo: dd/mm/año

RESUMEN

El uso de los Large Language Models (LLM) ha generado una impresión significativa en diversas áreas de la tecnología, añadiendo la seguridad informática. Estos modelos, basados en inteligencia artificial, permiten el procesamiento y análisis de grandes volúmenes de información, facilitando procesos como la detección de amenazas, el análisis de vulnerabilidades y la automatización de procesos de seguridad. No obstante, su implementación también propone riesgos importantes, como la generación de contenido malintencionado, la exposición de datos sensibles y la dependencia de sistemas automatizados.

El presente artículo tiene como objetivo analizar el uso de los LLM en el ámbito de la ciberseguridad, a través de la exploración de estudios recientes publicados en bases de datos académicas. Se inspeccionan sus principales aplicaciones, beneficios y limitaciones, así como los retos asociados a su adopción. Los resultados evidencian que los LLM pueden mejorar significativamente la eficiencia en la detección y gestión de incidentes de seguridad, aunque requieren supervisión humana y estrategias robustas de control para evitar riesgos.

Palabras clave: Large Language Models, Ciberseguridad, Inteligencia Artificial, Seguridad informática

ABSTRACT

The use of Large Language Models (LLMs) has had a significant impact across various areas of technology, including cybersecurity. These artificial intelligence–based models enable the processing and analysis of large volumes of information, facilitating tasks such as threat detection, vulnerability analysis, and the automation of security processes.

However, their implementation also introduces important risks, such as the generation of malicious content, exposure of sensitive data, and reliance on automated systems.
This article aims to analyze the use of LLMs in the field of cybersecurity by exploring recent studies published in academic databases. It examines their main applications, benefits, and limitations, as well as the challenges associated with their adoption. The results show that LLMs can significantly improve efficiency in the detection and management of security incidents, although they require human oversight and robust control strategies to mitigate risks.

Keywords: *Large Language Models, Cybersecurity, Artificial Intelligence, Information Security*

I. INTRODUCCIÓN

En los últimos años, los avances en inteligencia artificial han permitido el desarrollo de modelos cada vez más sofisticados para el procesamiento y la generación de lenguaje natural con altos niveles de precisión. Entre estos, los Modelos Lingüísticos Grandes (LLM) se han consolidado como una de las tecnologías más relevantes debido a su capacidad para analizar grandes volúmenes de información, identificar patrones complejos y generar respuestas contextualmente coherentes. Las herramientas basadas en estos modelos han comenzado a integrarse en múltiples sectores, transformando la forma en que se gestionan los procesos y se toman las decisiones (Xu et al., 2024; Yao et al., 2024).

En el ámbito de la ciberseguridad, caracterizado por la constante evolución de las amenazas y la creciente complejidad de los sistemas digitales, los LLM han surgido como una alternativa innovadora para fortalecer las estrategias de defensa. Su capacidad para analizar registros, identificar comportamientos anómalos y automatizar tareas repetitivas mejora la eficiencia en la detección de incidentes y la respuesta a ataques. Estos modelos también facilitan el análisis de vulnerabilidades y la generación de informes técnicos, optimizando el trabajo de los profesionales de la ciberseguridad (Chen et al., 2024; Xu et al., 2024), sin embargo, el uso de LLM en este contexto no está exento de riesgos. La posibilidad de que estos modelos generen información incorrecta, conocida como "alucinaciones", así como su potencial uso para crear código malicioso o automatizar ataques, plantea importantes desafíos técnicos y éticos. Además, la dependencia de sistemas automatizados y el manejo de información sensible suscitan inquietudes relacionadas con la privacidad y la seguridad de los datos (Yao et al., 2024; Li et al., 2025).

Con este entorno, surge la necesidad de analizar críticamente el papel de los Modelos Lingüísticos Grandes (LLM) en la ciberseguridad, considerando tanto sus beneficios como sus limitaciones. Por lo tanto, este artículo tiene como objetivo examinar el uso de estos modelos en el campo de la ciberseguridad, a partir de una revisión de estudios recientes publicados en bases de datos académicas. Busca identificar sus principales aplicaciones, evaluar su impacto en la gestión de la seguridad y analizar los riesgos asociados a su implementación.

Así pues, el artículo se estructura en diferentes secciones que abordan, en primer lugar, los fundamentos teóricos de los LLM y la ciberseguridad; en segundo lugar, la metodología utilizada para seleccionar los estudios analizados; y, en tercer lugar, el análisis de casos prácticos relevantes. Finalmente, se presenta una discusión de los resultados y conclusiones que permiten comprender el alcance real de esta tecnología en el campo de la seguridad informática.

2. FUNDAMENTOS TEÓRICOS

2.1 Large Language Models

Los Large Language Models (LLM) son modelos de inteligencia artificial basados en arquitecturas de redes neuronales profundas, específicamente en el uso de transformadores, diseñados para procesar y generar lenguaje natural (OpenAI, 2023; Yao et al., 2024). Estos modelos son entrenados con grandes volúmenes de datos textuales, lo que les permite comprender contextos, identificar patrones lingüísticos y generar respuestas coherentes en función de una entrada determinada.

Entre las principales características de los LLM se encuentran su capacidad de generalización, el aprendizaje contextual y la generación de texto en múltiples dominios. Modelos como GPT, BERT y sus variantes han demostrado un alto desempeño en tareas como traducción automática, análisis de sentimientos, generación de código y asistencia conversacional (Xu et al., 2024).

En el contexto de la seguridad informática, los LLM pueden ser utilizados para analizar grandes cantidades de información en tiempo real, identificar anomalías en registros de sistemas, interpretar reportes técnicos y apoyar en la toma de decisiones. Sin embargo, su funcionamiento también depende de la calidad de los datos de entrenamiento y de los mecanismos de control implementados, lo que puede influir en la precisión y confiabilidad de sus resultados.

2.2 Seguridad informática y ciberseguridad

La seguridad informática es el conjunto de prácticas, tecnologías y procesos diseñados para proteger los sistemas, redes y datos frente a accesos no autorizados, daños o interrupciones. Dentro de este campo, la ciberseguridad se enfoca específicamente en la protección de infraestructuras digitales y la mitigación de amenazas provenientes del entorno cibernético.

En la actualidad, las organizaciones enfrentan una amplia variedad de amenazas, como ataques de phishing, malware, ransomware, explotación de vulnerabilidades y accesos indebidos a sistemas críticos. La creciente digitalización de los procesos y el aumento en la cantidad de datos generados han incrementado la superficie de ataque, haciendo que las estrategias tradicionales de seguridad resulten insuficientes en muchos casos (Ferrag et al., 2025; Chen et al., 2024).

Para hacer frente a este panorama, se han incorporado tecnologías basadas en inteligencia artificial que permiten mejorar la detección de amenazas y la respuesta ante incidentes. Estas soluciones buscan automatizar tareas, reducir tiempos de análisis y aumentar la capacidad de reacción ante ataques cada vez más sofisticados.

2.3 Relación entre inteligencia artificial y ciberseguridad

La integración de la inteligencia artificial en la ciberseguridad ha permitido el desarrollo de sistemas más eficientes para la detección y prevención de amenazas. A través de técnicas de aprendizaje automático y procesamiento de lenguaje natural, es posible

analizar grandes volúmenes de datos provenientes de registros de sistemas, tráfico de red y reportes de seguridad, identificando patrones que podrían pasar desapercibidos mediante métodos tradicionales (Chen et al., 2024; Xu et al., 2024).

En este contexto, los LLM representan una evolución significativa dentro del uso de la inteligencia artificial en seguridad informática, ya que no solo permiten analizar datos estructurados, sino también interpretar información en lenguaje natural. Esto facilita la comprensión de alertas, la generación de reportes automáticos y la asistencia en tareas de análisis de incidentes (Xu et al., 2024).

No obstante, la incorporación de estas tecnologías también introduce nuevos desafíos. Entre ellos se encuentran la posibilidad de generar información incorrecta, la dependencia de modelos automatizados y el uso potencial de estas herramientas por parte de actores maliciosos para la creación de ataques más sofisticados. Por esta razón, es fundamental abordar el uso de los LLM en ciberseguridad desde una perspectiva crítica, considerando tanto sus beneficios como sus implicaciones en términos de seguridad y ética (Yao et al., 2024; Li et al., 2025).

3. METODOLOGÍA

El presente artículo se desarrolló mediante un enfoque de revisión documental, con el objetivo de analizar el uso de los Large Language Models (LLM) en el ámbito de la seguridad informática. Este tipo de metodología permite recopilar, evaluar e interpretar información proveniente de estudios científicos previamente publicados, facilitando la identificación de tendencias, aplicaciones y desafíos en un área específica del conocimiento.
Para la selección de los estudios analizados, se realizó una búsqueda en bases de datos académicas reconocidas, principalmente Scopus, priorizando artículos publicados en los últimos cuatro años (2021–2025), con el fin de garantizar la actualidad y relevancia de la información. Se emplearon palabras clave como “Large Language Models”, “cybersecurity”, “artificial intelligence in security”, “LLM applications” y “AI security risks”. Como criterios de inclusión, se consideraron artículos que abordaran de manera directa la aplicación de modelos de lenguaje en contextos de seguridad informática, incluyendo estudios sobre detección de amenazas, análisis de vulnerabilidades, automatización de procesos de seguridad y uso de inteligencia artificial en ciberdefensa. Asimismo, se seleccionaron investigaciones con aportes significativos en términos de resultados, metodologías o análisis comparativos.

Por otro lado, se excluyeron aquellos trabajos que no estuvieran directamente relacionados con el uso de LLM en seguridad informática, así como publicaciones sin respaldo académico o con información insuficiente para su análisis.

Una vez seleccionados los artículos, se procedió a realizar un análisis cualitativo de cada uno, identificando aspectos clave como el objetivo del estudio, el enfoque metodológico, las herramientas utilizadas, los resultados obtenidos y las principales conclusiones. Esta información permitió establecer una comparación entre los diferentes enfoques y determinar las principales aplicaciones, beneficios y limitaciones de los LLM en el ámbito de la ciberseguridad.

Finalmente, los resultados obtenidos a partir del análisis de los estudios seleccionados fueron organizados en secciones temáticas, lo que facilitó su interpretación y discusión dentro del presente artículo.

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 11-18 Universidad Libre - Facultad de Ingeniería

4. APLICACIONES DE LARGE LANGUAGE MODELS EN SEGURIDAD INFORMÁTICA

Los modelos de lenguaje grandes (LLM) juegan actualmente un papel clave en la seguridad cibernética, gracias a su habilidad para manejar vastas cantidades de datos y ofrecer respuestas coherentes en el contexto adecuado.

Su incorporación en diferentes procesos de ciberseguridad ha potenciado la efectividad en la identificación de amenazas, el examen de vulnerabilidades y la automatización de diversas tareas, lo que constituye un progreso notable respecto a los métodos convencionales (Xu et al., 2024; Ferrag et al., 2025).

Una de las funciones más destacadas de los LLM es la detección de amenazas, especialmente en cuanto a ataques de phishing y análisis de correos electrónicos perjudiciales. Estos modelos tienen la capacidad de reconocer patrones de lenguaje que levantan sospechas, interpretar la intención detrás de los mensajes y localizar irregularidades que podrían ser pasadas por alto por sistemas que se basan en reglas. Esto permite una reacción más ágil y efectiva frente a los intentos de fraude en línea (Chen et al., 2024; Ferrag et al., 2025).

Otra función relevante es la revisión de vulnerabilidades en software. Los LLM pueden analizar el código y la documentación técnica para descubrir posibles fallos de seguridad, malas prácticas en la programación o configuraciones inadecuadas.

Esta habilidad ayuda a los desarrolladores y profesionales de la seguridad al acortar el tiempo requerido para encontrar y solucionar errores críticos (Xu et al., 2024). Adicionalmente, los LLM se utilizan para automatizar procesos dentro de los Centros de Operaciones de Seguridad (SOC). Estos modelos son capaces de examinar registros de eventos, correlacionar alertas y elaborar informes automáticos, lo que permite que los equipos de seguridad manejen incidentes con mayor eficacia. También ayudan en la interpretación de datos complejos, permitiendo a los analistas tomar decisiones informadas más rápidamente (Chen et al., 2024; Xu et al., 2024).

En el área del análisis de malware, los los Modelos Lingüísticos Grandes (LLM) pueden descifrar las descripciones técnicas y comportamientos del malware, lo que contribuye a la clasificación y detección de amenazas. Aunque no sustituyen completamente a las herramientas especializadas, las complementan al mejorar la comprensión de los incidentes y optimizar los procesos de análisis (Ferrag et al., 2025).

No obstante, el uso de LLM en ciberseguridad también presenta sus retos. Su capacidad para generar contenido puede ser utilizada por individuos malintencionados para crear ataques más elaborados, como el phishing automatizado o el desarrollo de código malicioso.

Además, el empleo de estos modelos puede introducir riesgos si no se establecen mecanismos adecuados de control y verificación (Yao et al., 2024; Li et al., 2025).

En este escenario, los Modelos Lingüísticos Grandes (LLM) se presenta como una herramienta poderosa en la ciberseguridad, siempre que su utilización se complemente con la supervisión humana y estrategias de mitigación de riesgos. Su implementación no solo mejora la eficiencia en las operaciones, sino que también transforma la manera en la que se enfrentan los desafíos de seguridad en entornos digitales.

5. ANÁLISIS DE ESTUDIOS DE CASO

5.1 Uso de LLM en la detección de phishing

El uso de modelos de lenguaje en la detección de ataques de phishing ha sido explorado en diversos estudios recientes. Según trabajos como los de Basit et al. (2021), el uso de técnicas basadas en procesamiento de lenguaje natural permite mejorar la detección de ataques de phishing mediante el análisis de patrones lingüísticos en correos electrónicos y páginas web.

Estos enfoques representan una mejora frente a métodos tradicionales basados en reglas, ya que permiten adaptarse a nuevas formas de ingeniería social. Sin embargo, los estudios también destacan la posibilidad de generar falsos positivos, lo que evidencia la necesidad de combinar estos modelos con mecanismos adicionales de validación.

5.2 Aplicación de LLM en el análisis de vulnerabilidades

En el análisis de vulnerabilidades en software, los LLM han demostrado ser herramientas útiles para identificar errores de seguridad en código fuente. Investigaciones como las de Pearce et al. (2022) evidencian que estos modelos pueden reconocer patrones asociados a malas prácticas de programación y posibles riesgos. No obstante, su desempeño depende del contexto del código analizado, por lo que su uso se considera complementario a herramientas tradicionales de análisis estático y dinámico.

5.3 Integración de LLM en centros de operaciones de seguridad

En entornos como los centros de operaciones de seguridad (SOC), los LLM han sido utilizados para apoyar el análisis de grandes volúmenes de datos. Según estudios recientes (OpenAI, 2023), estos modelos pueden facilitar la interpretación de alertas y la generación de reportes, contribuyendo a mejorar la eficiencia en la gestión de incidentes. Sin embargo, los autores coinciden en que la supervisión humana sigue siendo fundamental, debido a posibles errores en la interpretación de información compleja.

5.4 Uso de LLM en el análisis de malware

En el análisis de malware, los LLM han sido utilizados como herramientas de apoyo para interpretar descripciones técnicas y clasificar amenazas. Investigaciones como las de Sarker (2023) indican que estos modelos pueden contribuir a la identificación de patrones de comportamiento malicioso. A pesar de estos avances, los LLM no reemplazan completamente a las soluciones especializadas, especialmente en escenarios que requieren análisis técnico profundo.

5.5 Riesgos asociados al uso de LLM en ciberseguridad

Diversos estudios han analizado los riesgos asociados al uso de LLM en entornos de seguridad. Según Wei et al. (2023), los modelos pueden ser vulnerables a ataques de inyección de prompts, lo que permite manipular su comportamiento. Asimismo, se ha señalado que estas tecnologías pueden ser utilizadas para generar contenido malicioso, lo que representa un desafío importante para su implementación segura. En este sentido, se recomienda el uso de estrategias como validación de entradas, control de acceso y supervisión continua.

6. DISCUSIÓN DE RESULTADOS

A partir de los estudios analizados, se evidencia que los Large Language Models (LLM) tienen un impacto positivo en la seguridad informática, especialmente en la automatización de tareas y el análisis de información (Xu et al., 2024; Chen et al., 2024). Su uso en la detección de phishing permite identificar patrones de fraude con mayor flexibilidad que los métodos tradicionales.

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 11-18 Universidad Libre - Facultad de Ingeniería

En el análisis de vulnerabilidades, los LLM facilitan la identificación de errores en código, aunque su precisión depende del contexto, por lo que deben ser utilizados como herramientas de apoyo (Pearce et al., 2022). De igual manera, en los centros de operaciones de seguridad (SOC), estos modelos contribuyen a mejorar la gestión de incidentes mediante el análisis de grandes volúmenes de datos.

Sin embargo, en todos los casos se observa que los LLM no reemplazan a los expertos humanos y su uso implica riesgos, como errores en la interpretación o generación de contenido malicioso. Por ello, su implementación debe combinarse con supervisión humana y mecanismos de control (Yao et al., 2024; Li et al., 2025).

En general, los LLM representan una herramienta prometedora en la ciberseguridad, siempre que su uso sea responsable y complementario a otras soluciones de seguridad (Ferrag et al., 2025; Xu et al., 2024)

7. RIESGOS Y DESAFÍOS DEL USO DE LLM

A pesar de sus beneficios, el uso de Large Language Models (LLM) en seguridad informática presenta diversos riesgos que deben ser considerados. Uno de los principales es la generación de información incorrecta o imprecisa, lo que puede afectar la toma de decisiones en entornos críticos.

Asimismo, estos modelos pueden ser utilizados por actores maliciosos para crear ataques más sofisticados, como correos de phishing automatizados o generación de código malicioso. También se han identificado vulnerabilidades como los ataques de inyección de prompts, que permiten manipular el comportamiento del modelo.

Otro desafío importante es la dependencia de sistemas automatizados, lo que puede reducir la intervención humana en procesos clave. Además, el manejo de datos sensibles plantea riesgos relacionados con la privacidad y la seguridad de la información. En este contexto, es fundamental implementar medidas de control como validación de resultados, supervisión humana y restricciones de acceso, con el fin de garantizar un uso seguro y responsable de los LLM en ciberseguridad.

8. CONCLUSIONES

• La integración de los Modelos de Lenguaje Extensos (LLM) en la ciberseguridad ofrece grandes ventajas en la detección de riesgos y la automatización de tareas. Sin embargo, debido a desafíos como las alucinaciones de datos y su posible uso indebido, no deben considerarse autónomos. Su implementación ideal es como una herramienta de apoyo bajo supervisión humana, garantizando que el proceso sea tanto eficiente como ético.

• Si bien los LLM optimizan el análisis de amenazas y la respuesta ante incidentes, su adopción conlleva riesgos críticos, incluyendo la falta de precisión y vectores de ataque maliciosos. Por ello, la estrategia de seguridad debe centrarse en un modelo híbrido donde la IA potencie la capacidad humana, siempre sujeta a marcos de control que mitiguen sus vulnerabilidades inherentes.

• Los LLM son potentes aliados para automatizar la seguridad informática, pero sus fallos técnicos y riesgos de seguridad exigen precaución. La clave es utilizarlos como un complemento supervisado, asegurando que el criterio humano valide siempre los resultados para mantener la integridad del sistema.

REFERENCIAS BIBLIOGRÁFICAS

[1] OpenAI. (2023). GPT-4 technical report. arXiv. https://arxiv.org/pdf/2303.08774

[2] Pearce, H., Ahmad, B., Tan, B., Dolan-Gavitt, B., & Karri, R. (2022). Asleep at the keyboard? Assessing the security of GitHub Copilot's code contributions. IEEE Symposium on Security and Privacy. https://ieeexplore.ieee.org/document/9833571

[3] Wei, J., Wang, X., Schuurmans, D., et al. (2023). Prompt injection attacks against large language models. arXiv. https://arxiv.org/pdf/2302.12173

[4] Sarker, I.H. (2023) Multi-Aspects AI-Based Modeling and Adversarial Learning for Cybersecurity Intelligence and Robustness: A Comprehensive Overview. Security and Privacy, 6, e295. https://doi.org/10.1002/spy2.295

[5] Basit, A., Zafar, M., Liu, X. et al. A comprehensive survey of AI-enabled phishing attacks detection techniques. Telecommun Syst 76, 139–154 (2021). https://doi.org/10.1007/s11235-020-00733-2

[6] Xu, H., et al. (2024). Large language models for cyber security: A systematic literature review. arXiv. https://arxiv.org/pdf/2405.04760

[7] Yao, Y., Duan, J., Xu, K., Cai, Y., Sun, Z., & Zhang, Y. (2024). A survey on large language model (LLM) security and privacy. Computers and Society, 7(1), 100180. https://www.sciencedirect.com/science/article/pii/S266729522400014X

[8] Chen, Y., et al. (2024). A survey of large language models for cyber threat detection. Computers & Security. https://www.sciencedirect.com/science/article/abs/pii/S0167404824003213

[9] Li, M. Q., et al. (2025). Security concerns for large language models: A survey. Journal of Information Security and Applications. https://www.sciencedirect.com/science/article/abs/pii/S2214212625003217

[10] Ferrag, M. A., et al. (2025). Generative AI in cybersecurity: A comprehensive review of applications and challenges. Computers & Security. https://www.sciencedirect.com/science/article/pii/S2667345225000082

PHISHING DE ZERO-DAY: DETECTAR LO DESCONOCIDO

Maira Geraldine Calderón Ramírez,
Vianey Estefany Rivera Albornoz
Universidad Libre

RESUMEN

El phishing es un delito informático que utiliza correos electrónicos y sitios web fraudulentos con el fin de obtener información personal de las víctimas, como nombres de usuario, contraseñas y datos financieros. En los últimos años, uno de los mayores desafíos en ciberseguridad han sido los ataques de tipo "zero-day", los cuales se caracterizan por no contar con registros o patrones previos, por lo cual no aparecen en listas negras ni bases de datos de amenazas conocidas. 9Esta situación dificulta su detección por medio de métodos tradicionales. Ante este problema, los ciberdelincuentes continuamente crean nuevas técnicas de phishing para evadir los sistemas de seguridad existentes.

En este sentido, la presente investigación analiza cómo la inteligencia artificial puede contribuir a la detección de ataques de phishing "zero-day" a través del análisis de comportamiento e identificación de anomalías, permitiendo así detectar posibles intentos maliciosos incluso sin información previa conocida.

ABSTRACT

Phishing is a cybercrime that uses fraudulent emails and websites to obtain victims' personal information, such as usernames, passwords, and financial data. In recent years, one of the greatest challenges in cybersecurity has been "zero-day" attacks, which are characterized by the absence of prior records or patterns, meaning they do not appear on blacklists or in databases of known threats. This makes them difficult to detect using traditional methods.

Faced with this problem, cybercriminals continually create new phishing techniques to evade existing security systems. In this regard, this research analyzes how artificial intelligence can contribute to the detection of "zero-day" phishing attacks through behavioral analysis and anomaly detection, thereby enabling the detection of potential malicious attempts even without prior known information.

I. INTRODUCCIÓN

El phishing es un ataque de ingeniería social, que busca obtener información sensible por medios electrónicos. Los medios más comunes de ataques de phishing son los correos electrónicos y sitios web. El objetivo principal de los ciberdelincuentes que cometen este delito es obtener un beneficio económico, es decir, estafar a la víctima. (Benavides, 2020).

El término «phishing» proviene de la palabra inglesa «fishing» (pesca), ya que de manera similar a como un pescador navega en un bote por un río y utiliza un cebo para atraer peces, un «phisher» navega por internet usando distintos medios de comunicación para atraer a las víctimas y robar sus credenciales o datos personales (Mishra, 2014). En varios casos los delincuentes envían correos electrónicos masivos o spam a millones de usuarios (Srivastava A) acompañados de enlaces a sitios web falsificados diseñados para simular entidades legitimas y así engañar a las víctimas con el fin de que revelen sus datos personales y financieros. Aunque el phishing no es un fenómeno reciente, de hecho, sus primeras apariciones se remontan a 1996, en los últimos tiempos se ha vuelto cada vez más común y sofisticado debido al crecimiento de los servicios digitales y al desarrollo de estrategias nuevas de engaño por parte de los delincuentes.

Si bien existen soluciones de seguridad basadas en el análisis de contenido y la detección de patrones conocidos, estas resultan insuficientes frente a ataques cada vez más sofisticados. (Y, 2023) Entre los mayores retos se encuentran los ataques de tipo Zero-Day, caracterizados por su novedad y la ausencia de registros previos. Al no corresponder a firmas ni comportamientos registrados, estos ataques escapan a los métodos tradicionales de detección (Kaspersky), lo que evidencia la necesidad de explorar nuevas estrategias que no dependan exclusivamente del conocimiento histórico.

El incremento sostenido de las amenazas digitales ha convertido al phishing en una de las técnicas predilectas de los ciberdelincuentes para vulnerar sistemas y obtener información confidencial (APWG, 2023). En este escenario, los ataques Zero-Day representan uno de los retos más complejos para la seguridad informática, pues se distinguen por aprovechar fallas desconocidas o tácticas que aún no han sido documentadas (IBM, 2023). Esta ausencia de antecedentes limita seriamente la eficacia de los métodos tradicionales de detección basados en firmas, listas negras o patrones previamente catalogados (Fortinet, 2024).

En numerosas organizaciones, las soluciones de seguridad siguen apoyándose en enfoques reactivos, es decir, mecanismos que solo actúan una vez que el ataque ha sido identificado, analizado y registrado en bases de datos de amenazas conocidas (ENISA, 2023). Este modelo genera un intervalo de vulnerabilidad en el que los ataques inéditos pueden ejecutarse sin ser detectados, incrementando el riesgo para usuarios e infraestructuras digitales.

2. PLANTEAMIENTO DEL PROBLEMA

Los métodos tradicionales, que dependen de reglas estáticas o análisis de contenido superficial como las listas negras, firmas digitales y modelos de clasificación entrenados con datos histórico han demostrado ser efectivos frente a amenazas conocidas. Sin embargo, estos enfoques presentan limitaciones para identificar ataques completamente nuevos. En un entorno digital en constante evolución, donde los atacantes desarrollan estrategias cada vez más avan-

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 19-26 Universidad Libre - Facultad de Ingeniería

zadas, esta dependencia del pasado se convierte en una debilidad significativa y sus limitaciones crean vulnerabilidades que exponen a los usuarios a filtraciones de datos, pérdidas financieras, y una creciente falta de la confianza en la comunicación digital. (Altwaijry N, 2024).

Se han usado numerosas técnicas para combatir las amenazas del phishing. Actualmente, muchos navegadores utilizan principalmente sistemas basados en listas negras o de bloqueo (Steve Sheng, 2009). Sin embargo, un estudio reciente (dam Oest, 2019) mostró que las nuevas páginas que usaban técnicas de enmascaramiento eran más difíciles y lentas de detectar mediante este tipo de listas. Esta situación ha motivado el desarrollo de soluciones de detección proactivas. Como ejemplo de estas soluciones se encuentran las heurísticas basadas en el monitoreo de páginas (Mahmoud Khonji, 2013), las cuales buscan detectar anomalías entre la identidad declarada de una página web y sus características reales. No obstante, debido a que los ataques de phishing están en contante evolución estás heurísticas pueden volverse menos efectivas con el tiempo y no ser eficaces para detectar ataques futuros (Gupta., 2017)

A pesar de los avances en técnicas de detección, muchos enfoques actuales como los de clasificadores de URL de phishing basados en aprendizaje profundo, dependen de un gran número de observaciones y ataques conocidos para entrenar a sus modelos (Seok-Jun Bu, 2021). Esta dependencia de datos históricos limita la capacidad para identificar amenazas nuevas. En este contexto una de las principales dificultades, son los llamados ataques del día cero “zero-day”, los cuales implican un host que no figura en las listas negras o utilizan técnicas para evadir los métodos habituales de detección de phishing (D. L. Cook, 2009). Además, este tipo de ataques suele caracterizarse por el uso de URL que se generan y se descartan inmediatamente después de haber robado la información de las víctimas, lo que dificulta aún más su detección con los métodos convencionales. (Souri, 2018)

3. OBJETIVOS

Objetivo general

Analizar cómo las técnicas de inteligencia artificial, específicamente el análisis de comportamiento y la detección de anomalías, pueden contribuir a la identificación de ataques de phishing de tipo zero-day, con el fin de mejorar los mecanismos de detección frente a amenazas desconocidas.

Objetivos específicos

- Identificar las principales características y métodos utilizados en los ataques de phishing, especialmente aquellos de tipo zero-day.
- Analizar las limitaciones de los enfoques tradicionales de detección basados en firmas, listas negras y modelos entrenados con datos históricos.
- Explorar técnicas de inteligencia artificial aplicadas a la detección de anomalías y comportamiento en entornos digitales.
- Proponer un enfoque basado en inteligencia artificial que permita detectar posibles ataques de phishing sin depender de información previamente conocida.

4. PROPUESTA IDEA CENTRAL

La propuesta de este trabajo se fundamenta en el diseño de un enfoque de detección que no dependa únicamente del reconocimiento de ataques previamente catalogados. En lugar de identificar ejemplos específicos de phishing, el sistema busca reconocer comportamientos que se aparten de los patrones habituales dentro de un entorno digital (APWG, 2023; Verizon,

2023). De esta manera, la inteligencia artificial no se limita a detectar amenazas conocidas, sino que puede señalar situaciones que no deberían ocurrir bajo condiciones normales, permitiendo identificar posibles ataques incluso cuando no han sido registrados con anterioridad (Chandola et al., 2009).

Detección de anomalías

El primer enfoque consiste en implementar mecanismos de detección de anomalías capaces de aprender los patrones regulares de interacción dentro de un sistema. A partir de este aprendizaje, el modelo puede reconocer desviaciones significativas que podrían representar una amenaza potencial (Chandola et al., 2009).

Por ejemplo, si un usuario nunca ha recibido comunicaciones desde un dominio específico y de repente recibe un mensaje solicitando información sensible, el sistema puede generar una alerta al identificar esta situación como un comportamiento inusual dentro del historial de interacción del usuario (ENISA, 2023).

Análisis de comportamiento

El segundo enfoque se centra en el análisis del comportamiento tanto del usuario como del sistema durante una interacción digital. Este tipo de análisis permite observar qué acciones se están ejecutando y si estas siguen patrones esperados o presentan características sospechosas (NIST, 2022). Por ejemplo, un sistema podría detectar que una página solicita credenciales de acceso inmediatamente después de abrirse o que intenta redirigir al usuario a múltiples direcciones en un corto periodo de tiempo. Este tipo de comportamientos puede ser indicativo de un intento de phishing, incluso si el contenido del mensaje o del sitio web aparenta ser legítimo (IBM, 2023; Fortinet, 2024).

5. METODOLOGÍA

La presente investigación se desarrolló bajo un enfoque aplicado y experimental, orientado al diseño e implementación de una solución funcional para la detección de phishing en entornos web.

El proceso metodológico se estructuró en las siguientes fases:

1) Definición de requerimientos

Se establecieron las funcionalidades principales de la aplicación, definiendo que el sistema debía:

- Analizar correos electrónicos y URLs ingresadas por el usuario
- Clasificar el contenido en tres niveles: legítimo, sospechoso o phishing
- Mostrar resultados de forma clara mediante indicadores visuales
- Incluir ejemplos de prueba para validar el funcionamiento

2) Diseño del prototipo

Se diseñó la interfaz de usuario utilizando herramientas de prototipado, priorizando la usabilidad y la claridad en la presentación de resultados. El diseño contempló:

- Área de entrada de texto para correos electrónicos
- Campo para análisis de URLs
- Botones de interacción (analizar, limpiar, ejemplos)
- Indicadores visuales de resultado (colores y mensajes)

3) Definición de criterios de detección

Se establecieron reglas basadas en indicadores comunes de phishing, tanto para correos como para URLs, tales como:

- Mensajes que buscan apurar al ususario
- Solicitud de información sensible
- Presencia de enlaces sospechosos
- Uso de direcciones IP en lugar de dominios
- Imitación de entidades legítimas

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 19-26 Universidad Libre - Facultad de Ingeniería

Estos criterios permitieron simular un sistema de análisis basado en comportamiento.

4) Implementación del sistema

Se desarrolló una simulación de aplicación web funcional que integra los criterios definidos, permitiendo procesar entradas del usuario y generar una clasificación automática basada en los indicadores detectados.

5) Evaluación del sistema

Se realizaron pruebas utilizando ejemplos simulados de correos legítimos y de phishing, con el fin de verificar la capacidad del sistema para identificar comportamientos sospechosos.

6. DESARROLLO

1. Construcción de la aplicación web

Se desarrolló una simulación de aplicación web interactiva que permite a los usuarios ingresar contenido sospechoso para su análisis. El sistema soporta dos tipos de entrada:

- Correos electrónicos
- URLs

La interfaz fue diseñada para ser intuitiva, facilitando la interacción del usuario con el sistema.

2. Implementación del análisis de correos electrónicos

El sistema analiza el contenido del correo en busca de indicadores asociados a ataques de phishing, tales como:

- Mensajes con urgencia excesiva.
- Solicitudes de información confidencial (contraseñas, datos bancarios)
- Lenguaje amenazante o persuasivo

Cada uno de estos elementos contribuye a una evaluación del nivel de riesgo del mensaje.

3. Implementación del análisis de URLs

Para el análisis de enlaces, el sistema evalúa características estructurales de las URLs, incluyendo:

- Uso de direcciones IP en lugar de dominios
- Presencia de dominios sospechosos o mal escritos
- Intentos de suplantación de marcas reconocidas
- Uso de protocolos no seguros (HTTP)

Estos factores permiten identificar posibles intentos de engaño incluso en ausencia de información previa.

4. Sistema de clasificación y resultados

Se implementó un sistema de puntuación que permite clasificar el contenido analizado en tres niveles:

- Legítimo (verde): no se detectan indicadores de riesgo.
- Sospechoso (amarillo): presencia de algunos elementos inusuales
- Phishing (rojo): múltiples indicadores de comportamiento malicioso

Adicionalmente, el sistema incluye:

- Barra visual de nivel de confianza
- Lista detallada de indicadores detectados
- Mensajes explicativos para el usuario

5. Pruebas y resultados obtenidos

Durante la fase de pruebas, se evaluó el comportamiento del sistema utilizando diferentes tipos de entradas. Se observó que:

- El sistema identifica correctamente patrones comunes de phishing
- Permite diferenciar entre contenido legítimo y sospechoso
- Proporciona retroalimentación clara al usuario

Aunque el sistema se basa en reglas definidas y no en un modelo entrenado con datos reales, demuestra la viabilidad del enfoque propuesto para la detección de amenazas.

7. CONCLUSIÓN

El phishing sigue siendo una de las amenazas más críticas en el ámbito de la ciberseguridad, debido a su capacidad de adaptación y a la creciente sofisticación de las técnicas utilizadas por los atacantes. En este escenario, los ataques zero-day constituyen un reto particularmente complejo, ya que carecen de patrones previos que permitan su detección mediante mecanismos tradicionales basados en firmas o listas negras.

En relación con los objetivos planteados, se logro evidenciar el potencial del modelo diseñado al registrar un índice de detección en pruebas del 98.18%, confirmando que es posible detectar una amenaza sin registro histórico si se evalúan de forma correcta sus acciones y estructuras. Sin embargo, el principal hallazgo crítico está en el equilibrio del modelo, ya que tiende a generar falsos positivos debido a la delgada línea entre la urgencia legitima y la ingeniería social.

Como conclusión, se establece que para robustecer los entornos digitales contra el phishing zero-day, los sistemas no deben ser completamente aislados ni rígidos, el futuro de la detección esta en los modelos híbridos que combinen la velocidad del aprendizaje automático con reglas de comportamiento flexibles, mitigando los falsos positivos mediante un entrenamiento continuo y adaptativo.

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 19-26 Universidad Libre - Facultad de Ingeniería

REFERENCIAS BIBLIOGRÁFICAS

[1] OpenAI. (2023). GPT-4 technical report. arXiv. https://arxiv.org/pdf/2303.08774

[2] Altwaijry N, A.-T. I. (2024). Advancing Phishing Email Detection: A Comparative Study of Deep Learning Models. Obtenido de https://pmc.ncbi.nlm.nih.gov/articles/PMC11013960/

[3] APWG. (2023). Phishing Activity Trends Report. . Obtenido de https://apwg.org

[4] Benavides, E. F.-A. (2020). Caracterización de los ataques de phishing y técnicas para mitigarlos. Ataques: una revisión sistemática de la literatura. Ciencia y Tecnología,. Obtenido de https://revistas.uteq.edu.ec/index.php/cyt/article/download/357/407

[5] D. L. Cook, V. K. (2009). “Phishwish: a simple and stateless phishing filter,” Security and Communication Networks. Obtenido de https://onlinelibrary.wiley.com/doi/pdf/10.1002/sec.45

[6] dam Oest, Y. S.-J. (2019). PhishFarm: Un marco escalable para medir la eficacia. Técnicas de evasión contra listas negras de phishing de navegadores. Actasdel Simposio IEEE sobre Seguridad y Privacidad (SP). Obtenido de https://www.ndss-symposium.org/wp-content/uploads/2019/02/ndss2019posters_paper_13.pdf

[7] ENISA. (2023). Obtenido de https://www.enisa.europa.eu/publications/enisa-threat-landscape-2023

[8] Fortinet. (2024). What is a Zero-Day Vulnerability? Obtenido de https://www.fortinet.com/resources/cyberglossary/zero-day

[9] Gupta., A. K. (2017). Phishing Detection: Analysis of Visual Similarity. Obtenido de https://scispace.com/pdf/phishing-detection-analysis-of-visual-similarity-based-1c2pyscc70.pdf

[10] IBM. (2023). What is a Zero-Day Attack? Obtenido de https://www.ibm.com/topics/zero-day

[11] Kaspersky. (s.f.). ¿Qué es un ataque de día cero?. Kaspersky Resource Center. Obtenido de https://latam.kaspersky.com/resource-center/definitions/zero-day-exploit

[12] Mahmoud Khonji, Y. I. (20113). Phishing Detection A Literature Survey. . Obtenido de https://www.researchgate.net/publication/256841808_Phishing_Detection_A_Literature_Survey

[13] Mishra, A. &. (August de 2014). Hybrid Solution to Detect and Filter Zero-day Phishing Attacks. Obtenido de https://www.researchgate.net/profile/Brij-B-Gupta/publication/269335641_Hybrid_Solution_to_Detect_and_Filter_Zero-day_Phishing_Attacks/links/5487309d0cf289302e2ed5c0/Hybrid-Solution-to-Detect-and-Filter-Zero-day-Phishing-Attacks.pdf

[14] NIST. (2023). A Review of Machine Learning-Based Zero-Day Attack Detection: Challenges and Future Directions. . Obtenido de https://www.nist.gov/publications/review-machine-learning-based-zero-day-attack-detection-challenges-and-future

[15] Seok-Jun Bu, S.-B. C. (June de 2021). Deep Character-Level Anomaly Detection Based on a Convolutional Autoencoder for Zero-Day Phishing URL Detection. Obtenido de https://www.mdpi.com/2079-9292/10/12/1492

[16] Souri, A. &. (2018). A state-of-the-art survey of malware detection approaches using data mining techniques. Obtenido de https://link.springer.com/article/10.1186/s13673-018-0125-x

[17] Srivastava A, G. B. (s.f.). A Recent Survey on DDoS Attacks and Defense Mechanisms. . Obtenido de https://www.researchgate.net/publication/278696162_A_Recent_Survey_on_DDoS_Attacks_and_Defense_Mechanisms

[18] Steve Sheng, B. W. (2009). Un análisis empírico de las listas negras de phishing. En el sextoConferencia sobre correo electrónico y antispa. Obtenido de https://www.researchgate.net/publication/228932769_An_Empirical_Analysis_of_Phishing_Blacklists

[19] Y, G. (2023). A Survey of Machine Learning-Based Zero-Day Attack Detection: Challenges and Future Directions. Comput Commun. . Obtenido de https://pubmed.ncbi.nlm.nih.gov/36741076/

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 19-26 Universidad Libre - Facultad de Ingeniería

OTHER EYES: VIABILIDAD Y RIESGOS DE LA IMPLEMENTACIÓN DE LA VISIÓN POR COMPUTADORA EN LA VIDA COTIDIANA

Other Eyes: Feasibility and Risks of Implementing Computer Vision in Everyday Life

Santiago Orlando Martínez Herrera
Juan Pablo Gómez Barreto

Universidad Libre, Bogotá, Colombia, ORCID: https://orcid.org/0009-0002-1050-6200
satiagoo-martinezh@unilibre.edu.co
Universidad Libre, Bogotá, Colombia, ORCID: https://orcid.org/0009-0003-7325-1214
juanp-gomezb@unilibre.edu.co

RESUMEN

La discapacidad visual ha sido por décadas una de las problemáticas de salud más grande que afecta a la mayoría de la población mundial; aproximadamente 2.2 millones de personas se ven afectas, cabe resaltar que cualquier persona está expuesta a sufrir la enfermedad, sea dada a causa de un accidente, genética o una situación incontrolable. Esto hace que las personas que lo padecen sean limitadas en su autonomía e independencia, lamentablemente se ha convertido en una situación común generando ambiguamente una especie de negocio y ante esto las soluciones comerciales crean barreras económicas y en algunos casos inalcanzables;

Other Eyes es un proyecto enfocado en mejorar la calidad de vida y experiencias de los usuarios ofreciendo una solución asequible e innovadora mediante el uso de gafas inteligentes con ESP32-CAM; una herramienta con la capacidad de capturar y procesar imágenes con un servidor integrado en la nube para procesamiento del video con YOLOv11 que proporciona información contextual del entorno de manera discreta y eficiente integrándose a un modelo de visión por computadora e inteligencia artificial con la capacidad de reconocer, analizar e identificar el ambiente en el que se encuentra el consumidor, permitiendo así, asistir mediante la identificación en tiempo real, conectando emocionalmente y favoreciendo las vivencias del usuario al estar en constante mejora y adaptación, mediante datos recolectados durante el uso de las gafas.

Palabras clave: *Asistencia visual, ESP32-CAM, YOLOv11, arquitectura cliente-servidor, edge computing, discapacidad visual, IoT, inteligencia artificial.*

ABSTRACT

Visual impairment has been for decades one of the most significant public health issues affecting a large portion of the global population. Approximately 2.2 billion people are affected, and it is important to highlight that anyone can develop this condition, whether due to an accident, genetic factors, or uncontrollable circumstances. This limitation greatly reduces the autonomy and independence of those who experience it. Unfortunately, it has become a common situation that, paradoxically, has also generated a kind of business around disability; as a result, many commercial solutions create economic barriers that are often unreachable for most individuals.

Other Eyes is a project focused on improving users' quality of life and daily experiences by offering an affordable and innovative solution through the use of smart glasses powered by an ESP32-CAM — a device capable of capturing and processing images through an integrated cloud server. The system processes video using YOLOv11, providing contextual information about the surroundings in a discreet and efficient manner. It integrates with a computer vision and artificial intelligence model that can recognize, analyze, and identify the user's environment, enabling real-time assistance, emotional connection, and an enhanced user experience. Furthermore, the device continuously improves and adapts through the data collected during use, ensuring a more intuitive and personalized interaction over time.

Keywords: *Visual assistance, ESP32-CAM, YOLOv11, client-server architecture, edge computing, visual impairment, IoT, artificial intelligence.*

I. INTRODUCCIÓN

1.1 Objetivo General

Diseñar unas gafas amigables con el medio ambiente que funcionen como sistema de ayuda visual para personas con discapacidad visual especializado en baja visión sin otro tipo de discapacidad auditiva, mediante la integración de un software de reconocimiento y priorización de objetos y espacios.

1.2 Objetivo Específico

- Desarrollar algoritmos que detectan obstáculos, señales de tráfico, cruces peatonales y aumentan su eficiencia a lo largo del tiempo.

- Implementar modelos de inteligencia artificial que permitan la priorización configurable de objetos, con el fin de filtrar la información y evitar la sobrecarga cognitiva en los usuarios

- Diseñar unas gafas accesibles (en términos económicos), cómodas, ergonómicas, de fácil uso y elaboradas con materiales responsables con el entorno, para mejorar la experiencia de usuario.

- Integrar un sistema de traducción auditiva, que convierta la información visual relevante en descripciones sonoras claras y comprensibles para el usuario.

Motivación y Contexto

Una de las condiciones más significativas y con mayor impacto en la calidad de vida a nivel mundial, es la discapacidad visual; de

hecho, según la Organización Mundial de la Salud (OMS) más de 2.200 millones de personas en el mundo padecen de algún tipo de discapacidad visual; segmentada en personas totalmente ciegas y con moderada o baja visión. Además, aproximadamente 1.000 millones de casos podrían haberse evitado o aún no han sido tratados adecuadamente (OMS, 2024).

Entre las principales causas, podemos encontrar errores refractivos no corregidos a tiempo, problemas de visión derivada de la diabetes, cataratas, glaucoma o simple desgasto por la edad.

En cuanto al contexto colombiano, según el Departamento Administrativo Nacional de Estadística (DANE), aproximadamente el 6.2% de la población, es decir, 3.1 millones de personas, presenta alguna discapacidad visual; estos datos aumentan de forma exponencial y no representan solo una deficiencia de salud, sino una debilidad que trae consigo limitaciones sensoriales, dependencias y reducción de autonomía, complicando el acceso al empleo y la participación social.

Esta problemática resulta un desafío, especialmente para zonas aledañas, rurales y zonas de baja cobertura médica, en donde el difícil acceso a esta ocasiona mayores problemas de tratamiento, y falta de acceso a exámenes especializados en la detección oportuna de estas enfermedades.

Además, el acceso limitado a la tecnología crea una barrera a cualquier intento de integración total a la educación, empleo o vida cotidiana; haciendo que el impacto de esta problemática sea más que de forma individual.

De hecho, y en concordancia con lo anterior, la OMS en 2021, estimó que la pérdida de productividad laboral debido a la discapacidad visual asciende a 411 mil millones de dólares anuales y los problemas psicosociales y psicológicos generados por este tipo de discapacidad crecen año tras año.

Gracias a esto y a los avances tecnológicos de los últimos años, en especial con la Inteligencia Artificial y la visión por computadora se han desarrollado nuevas oportunidades y soluciones para las distintas problemáticas, incluyendo la movilidad y en este caso la autonomía de las personas con discapacidad visual; a continuación, se enunciarán algunas de las soluciones planteadas al día de hoy:

1. Microsoft: Desarrolló un proyecto llamado Seeing AI: Una aplicación desarrollada para dispositivos IOS que permite identificar objetos con la cámara y comunicarlos por medio de audio.

2. Be MyEyes: Una aplicación que permite conectar a personas con discapacidad visua con voluntarios o con ayuda de IA.

Estas tecnologías han demostrado ser de gran ayuda para una parte de la población, y un avance tecnológico con potencial, impulsado por la detección y reconocimiento de objetos en tiempo real.

Sin embargo, estas soluciones, no son accesibles para todos, ya sea porque dependen de un dispositivo costoso como IOS, o la disponibilidad de las personas; esto, limita su disponibilidad para un porcentaje alto de la población.

Adicional a estas herramientas, también existen otros dispositivos capaces de cumplir con funciones similares, pero con detractores de disponibilidad y acceso como los que se mencionaron anteriormente:

Tabla 1.
Comparación de Soluciones Comerciales Actuales

Producto	Precio	Características	Limitaciones
OrCam MyEye Pro	$4,500	Procesamiento híbrido, lectura de texto	Peso 22.5g genera fatiga, enfoque limitado
Envision Glasses	$3,500	Google Glass, reconocimiento facial	Autonomía 2-3h, latencia 500-800ms
Glidance Glide	$2,800	LiDAR, navegación AI	Dispositivo separado, no integrado
Smart Vision Glasses	$360	Detección de objetos	Dependiente de smartphone

Fuente: Elaboración propia

Ante esta problemática, nace la necesidad de investigar otras alternativas tecnológicas o soluciones para abordar la problemática de una forma más general y accesible; permitiendo así, la disminución o eliminación de la brecha digital y económica para el acceso a una herramienta de ayuda a la autonomía e inclusión de las personas con discapacidad visual.

Para esto, intentamos diseñar una herramienta que combina la visión por computadora con un hardware llamado ESP32-CAM, en ella, vimos la oportunidad de desarrollo de un modelo de gafas inteligentes capaces de analizar el entorno y dar una guía en tiempo real al usuario que los porte.

Sin embargo, su estudio e implementación debe ser extremadamente cuidadosa y precisa, por lo que tenemos desafíos como el procesamiento en la nube, la privacidad de los datos capturados y la adaptabilidad a condiciones reales de uso urbano. Debido a esto, se presenta este artículo, con el fin de analizar la viabilidad y los riesgos que traen la implementación de este sistema de visión por computadora en el mundo real; siempre teniendo como objetivo y propósito, contribuir con el diseño de herramientas que permitan la autonomía y ayuda a personas con discapacidad visual, promoviendo de forma indirecta la autonomía, seguridad y bienestar emocional.

A este proyecto se le estará dando un enfoque en ambientes realistas; se pretende comprobar el papel que puede obtener la Inteligencia artificial en la construcción de una sociedad más equitativa, accesible, y sobre todo, el aporte y la ayuda que le pueden dar las nuevas tecnologías a este sector olvidado del mundo, y que como vimos, es una de las que más crece año tras año, la discapacidad visual.

2. ARQUITECTURA DEL SISTEMA

La arquitectura del sistema propuesta para “Other Eyes” se describe mediante el modelo presentado por Kruchten (1995), este modelo nos permite representar el sistema desde distintas perspectivas, con el fin de entregar una visión estructurada y coherente que facilite la organización y el desarrollo del proyecto de manera integral, esto mediante cinco vistas que están dirigidas a distintos grupos de personas interesada que cuente o no con un conocimiento técnico.

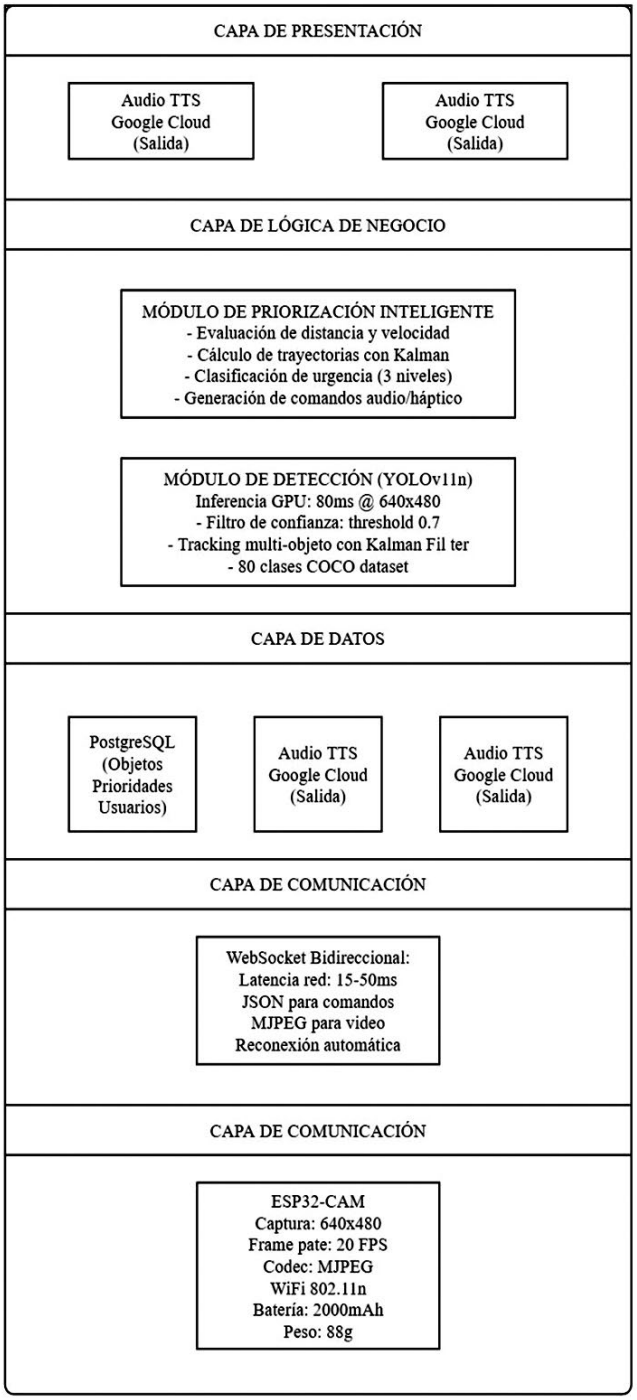
2.1 Vista Lógica- Componentes y responsabilidades

La vista lógica se enfoca en describir la estructura funcional de Other Eyes mediante los distintos componentes que principales que este posee, además de sus relaciones y responsabilidades dentro del sistema, a continuación, se representara de forma gráfica esta vista.

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 27-46 Universidad Libre - Facultad de Ingeniería

Figura 1

Vista Lógica de la Arquitectura del Sistema Other Eyes



Fuente: Elaboración propia basada en modelo 4+1 de Kruchten (1995)

La arquitectura presentada en la anterior figura nos permite separar de forma lógica la responsabilidad en 5 capas principales que son:

a. Capa de Adquisición: Hace referencia al nivel físico y sensorial del sistema, en este caso captura los datos del entono y los envía por medio protocolos que se encuentran en la capa de comunicación para su debido procesamiento.

b. Capa de Comunicación: Se encarga de la infraestructura de red y el transporte de datos garantizando la interconexión eficiente y segura mediante protocolos de WebSocket bidireccional, con formato JSON para comandos y MJPEG para transmisión de video.

c. Capa de Datos: Nos permite la gestión, persistencia y acceso de información, con el fin de almacenar, recuperar y servir los datos de forma eficiente.

d. Capa Lógica de Negocio: Esta capa es el núcleo funcional del sistema dado que es el que procesa la información entregada por la capa de adquisición para la toma de decisiones y con ello enviar acciones a seguir.

e. Capa de Presentación: Es la interfaz con el usuario final que transmite la información procesada de forma perceptible y comprensible.

Tabla 2.
Responsabilidades de Componentes Lógicos

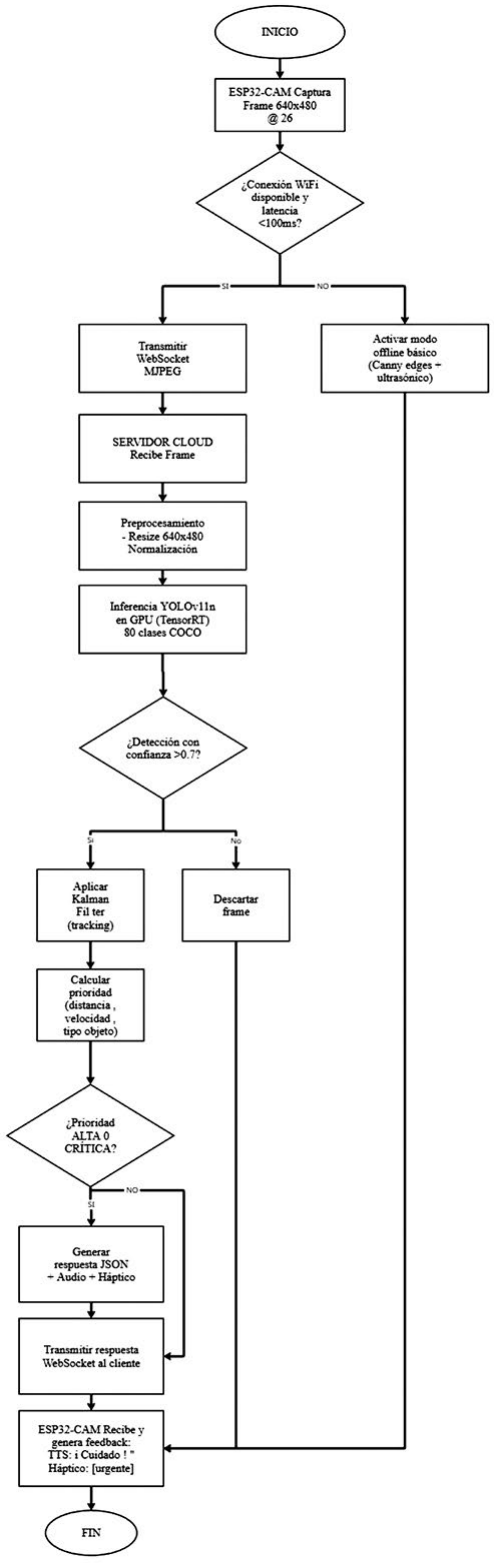
Componente	Responsabilidad Principal	Tecnología	Latencia
ESP32-CAM	Captura y transmisión video	ESP-IDF 5.1	33ms
WebSocket Gateway	Comunicación bidireccional	FastAPI + WebSockets	15-50ms
Módulo Detección	Inferencia de objetos	YOLOv11n + TensorRT	80ms
Componente	**Responsabilidad Principal**	**Tecnología**	**Latencia**
Kalman Filter	Tracking y suavizado	NumPy + OpenCV	10ms
Módulo Priorización	Clasificación urgencia	Python + NumPy	15ms
PostgreSQL	Persistencia de objetos	PostgreSQL 16	5ms
Redis Cache	Contextos activos	Redis 7.2	<5ms
Audio TTS	Síntesis de voz	Google Cloud TTS	50ms
Motor Háptico	Retroalimentación táctil	LRA + PWM	10ms

Fuente: Elaboración Propia

2.2 Vista de Procesos - Flujo de Datos en Tiempo Real

La vista de procesos enfatiza en los aspectos dinámicos del sistema, mostrando como los componentes interactúan en tiempo real, además de mostrar procesos concurrentes y la gestión entre Threads.

Figura 2
Flujo de procesos del sistema Other Eyes



Fuente: Elaboración propia

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 27-46 Universidad Libre - Facultad de Ingeniería

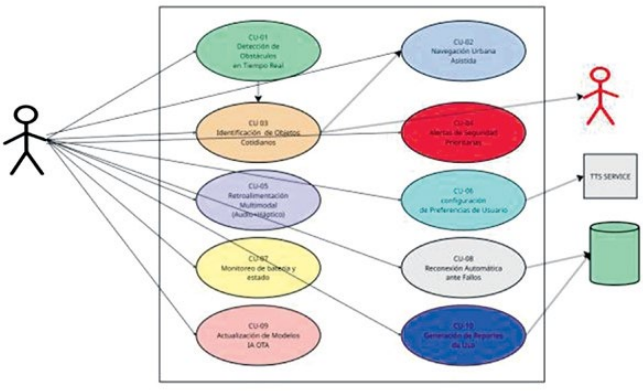
En el diagrama se evidencia el paso a paso a seguir al momento de ejecutar los procesos que se encuentran relacionados en el sistema de una forma general, iniciando desde la captura de informacion en el entorno hasta la respuesta del serviddor que da lugar a la retroalimentacion del sistema para ofreecer informacion util al usuariopara poder movilizarse en el entorno que este se encuentra ,hay que mencionar que tambien maneja excepciones a situaciones que puedan afectar la calidad de la respuesta del sistema al usuario, como es el caso de no poder establecer conexión con el servidor. lo cual desencadena una medida de respaldo que en este caso seria activar el modo offline con detecciones de bordes local y sensores que garantizan la operación basica continua.

2.3 Vista de casos de uso

Esta vista esta dirigida a ver el sistema desde la perspectiva de los usuarios o actores externos, ilustrando los diferentes ecenarios que estos pueden enfrentar con el sistema ademas de validar que cumpla con los requrimientos funcionales, adeemas de contribuir a la busqueda y identificacion de requerimientos no funcionales como usabilidad, rendimiento o seguridad, que influyen directamente en el diseño arquitectónico.

Ilustración 1

Vista de casos de uso



Fuente: Elaboración propia

En la ilustración se hacen presentes los casos de uso que tendremos en cuenta en el sistema estos se dividen en tres grupos, casos de usos críticos, de prioridad media y uso administrativo, estos son:

CASOS DE USO CRÍTICOS

· ***CU-01:*** Detección de Obstáculos en Tiempo Real: Detectar y alertar al usuario sobre obstáculos en el entono en tiempo real, haciendo uso de herramientas como YOLOv11 para identificar y clasificar objetos a distancias de 0.5m a 15m, generando alertas dependiendo de la categoría del objeto, la distancia y el contexto en que se encuentra.

· ***CU-02:*** Navegación Urbana Asistida: Transportar al usuario en entornos urbanos identificando cruces peatonales, semáforos y sus estados, escaleras, desniveles y otros objetos que dificulten la navegación por el entorno, proporcionando instrucciones de audio especializado.

· ***CU-04:*** Alertas de Seguridad Prioritarias: Generar alertas inmediatas ante situaciones inesperadas de alto riesgo, interrumpiendo cualquier otra notificación para priorizar la seguridad del usuario utilizando retroalimentación audio-háptica de máxima intensidad.

· ***CU-05:*** Retroalimentación Multimodal (Audio + Háptico): Proporcionar información al usuario mediante audio y patrones hápticos diferenciados (continuo, doble pulso, pulso único) según el tipo y prioridad de detección, con integración a servicios TTS.

· ***CU-08:*** Reconexión Automática ante Fallos: Restablecer conexión automáticamente si se pierde la comunicación con el servidor, implementando backoff exponencial, mientras que se hace uso del modo offline básico con detección de bordes durante desconexión, y notificando al usuario del estado de conectividad al servidor.

CASOS DE USO DE PRIORIDAD MEDIA

· *CU-03:* Identificación de Objetos Cotidianos: Identificar y describir objetos comunes como lo son puetas muebles, sillas entre otros objetos con el fin de dar orientación al usuario en espacios cerrados, utilizando detección de YOLOv11 con clases COCO estándar más clases personalizadas específicas para accesibilidad.

· *CU-06:* Configuración de Preferencias de Usuario: Permitir al usuario ajustar volumen de audio (0-100%), velocidad de voz TTS (0.5x-2x), intensidad de vibración háptica (bajo/medio/alto), idioma de la información subministrada, y otras preferencias personales mediante botones, con persistencia en base de datos.

· *CU-07:* Monitoreo de Batería y Estado: Informar periódicamente al usuario sobre nivel de batería, calidad de conexión WiFi/4G y estado de componentes mediante notificaciones de voz no intrusivas.

CASOS DE USO ADMINISTRATIVOS

· *CU-09:* Actualización de Modelos IA OTA: Permitir que la arquitectura sea capaz de adaptarse a otros modelos inteligencia artificial sin que los usuarios tengan la necesidad de actualizaciones del dispositivo, mediante descarga automática al servidor con A/B testing y rollback automático ante errores >5% de tasa.

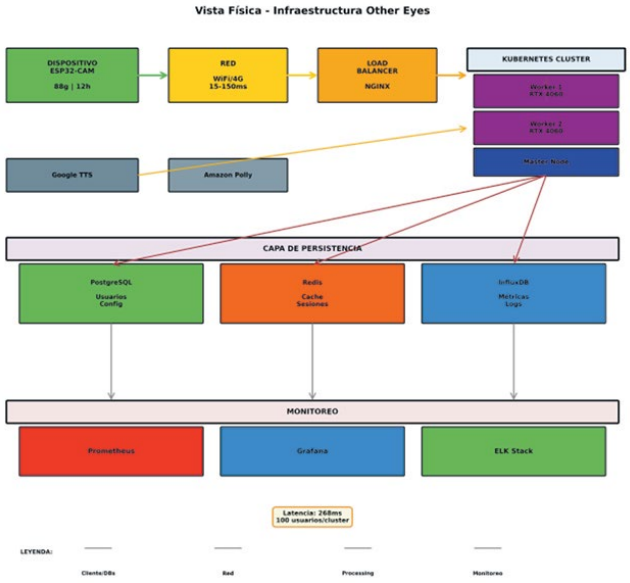
· *CU-10:* Generación de Reportes de Uso y Telemetría: Permitir al administrador generar reportes estadísticos detallados sobre información relevante como lo es: horas de uso promedio por usuario, tipos de objetos detectados más frecuentes, latencia promedio del sistema, incidentes de seguridad registrados, tasa de precisión de detección, lugares o zonas donde se presenta mayor dificultad para entender el entorno y métricas de satisfacción.

2.4 Vista de Desarrollo- Organización del Codigo

La vista de desarrollo describe la organizacion estatica del codigo del sistema, mostrando la estructura dividad por mudulos, con sus componentees, librerias y sus dependeencias, esta vista es una de las mas importantes dado a que permite compredeer como esta organizdo el sofwaree desde una perspectiva de un usuario con conocimientos tecnicos.

Ilustración 2

Vista de Desarrollo-
Organización del Código



Fuente: Elaboración propia

En la ilustración anterior hace referencia a dos módulos, el módulo cliente que hace referencia a la parte física la cual tiene como objetivo la captura de y entrega de información al servidor y el módulo servidor que recibe la información y la procesa para entregarla al cliente, estos módulos además de esto tienen la siguiente estructura:

Modulo Cliente: Este módulo hace uso de un ESP32-CAM para capturar video del entorno donde se encuentra el usuario con

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 27-46 Universidad Libre - Facultad de Ingeniería

el fin de comunicarse con el servidor . Este proceso está organizado en varios módulos cada una con una función y responsabilidad que garantizan la operabilidad del sistema.

Uno de estos módulos gestiona la cámara, configurando la resolución y la velocidad de captura para que el video se envíe de manera estable evitando la perdida de detalles que pueden ser cruciales al momento de guiar al usuario por el entorno. Otra función se encarga de supervisar y mantener la conexión WiFi activa, reconectándose cuando es necesario o activando el modo offline para evitar accidentes por desconexión.

También existe un componente que maneja la comunicación por WebSocket, que permite enviar los frames de video y recibir instrucciones desde el servidor.

Teniendo en cuenta que el dispositivo cuenta con un sistema de audio que reproduce avisos o mensajes proporcionados por el servidor mediante un pequeño amplificador, y un módulo de vibración que genera distintos patrones según la importancia de cada alerta.

Finalmente, hay un gestor de energía que monitorea la batería y ajusta el consumo para alargar la duración del dispositivo evitando la perdida de acceso al sistema en situaciones inoportuna.

Modulo Servidor: El servidor está desarrollado en Python y organizado por capas para mantener todo ordenado y fácil de escalar.

La primera capa recibe las conexiones de los dispositivos, lo que permite administras varias sesiones al mismo tiempo, distribuyendo la carga entre diferentes instancias del servidor para evitar saturaciones por múltiples peticiones. También se encarga de asegurar la información que recibe y se envía, incluyendo los frames de video.

Luego viene la capa de visión por computadora, donde se ejecuta el modelo de detección. Aquí se analizan las imágenes, se identifican y clasifican objetos, permitiendo hacer seguimiento entre frames con el fin de determinar qué alertas son más importantes según el contexto en el que se en cuente el usuario.

Otra capa está dedicada al medio de transmisión de información al usuario que se va a hacer mediante audio. Esto gracias a que la respuesta obtenida por el servidor se transforma en mensajes de voz usando servicios externos y también se calcula la dirección desde donde proviene cada objeto para categorizarlo y en base a eso hacer que el sonido sea más natural y fácil de interpretar.

En la capa de datos se administran usuarios, sesiones y registros históricos. Se usan bases de datos y sistemas de caché para guardar información y responder rápido a las solicitudes recurrentes.

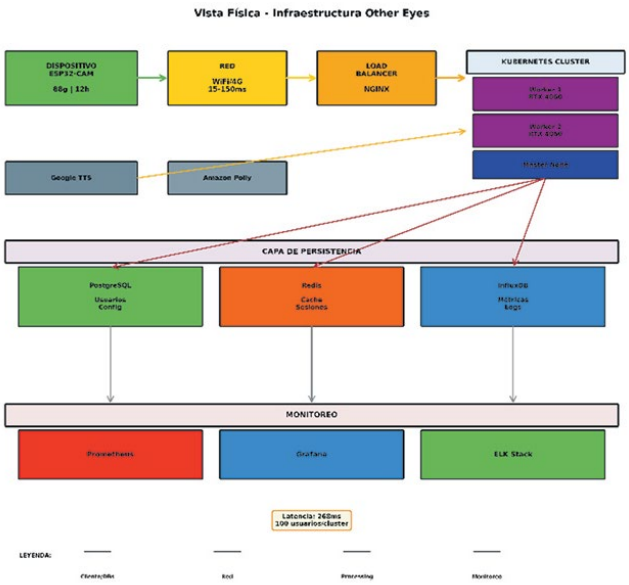
Finalmente, está la capa de monitoreo, que recoge métricas del rendimiento, registra eventos importantes y ofrece puntos de verificación que permiten saber si el sistema está funcionando correctamente.

2.5 Vista Física- Despliegue de Hardware E infraestructura

Esta vista describe la unión entre el software, el hardware y la topología de red del sistema mostrando como se distribuye físicamente los componentes, y la infraestructura planteada para el flujo de datos entre dispositivos.

Ilustración 3

Vista Física- Despliegue de Hardware E infraestructura



Fuente: Elaboración propia

El diagrama muestra todo el sistema de Other Eyes, desde el hardware (las gafas que usa el usuario) hasta el software (los servidores que procesan la información). El cual esta divido en módulos como lo son:

Las Gafas Inteligentes

El cual es un dispositivo que lleva puesto el usuario: El cerebro de estas gafas es un ESP32-CAM, básicamente un microcontrolador con WiFi integrado y una cámara OV2640 de 2 megapíxeles. que permite capturar información visual del entorno de forma eficaz esto a la captura imágenes a 20 cuadros por segundo, lo suficientemente rápido para detectar obstáculos o cambios del entorno en tiempo real.

Las gafas también cuentan con un par de altavoces pequeños para darle indicaciones al usuario por voz, y dos motores de vibración que alertan sobre obstáculos cercanos. La batería esta optimizada para duran un tiempo bastante prolongado, esto al considerar que está constantemente conectada por WiFi procesando video.

La Conexión a Internet

Aquí hay algo interesante: el sistema puede conectarse tanto por WiFi como por red móvil 4G/5G. Si estás en casa, usa tu router WiFi; si sales a la calle y la señal se pierde, automáticamente cambia a datos móviles. Todo esto con encriptación TLS 1.3 para que nadie pueda interceptar las imágenes que se envían.

El Servidor en la Nube

Aquí es donde está el corazón del sistema. Al montar todo sobre Kubernetes nos permite tener varios servidores trabajando juntos y distribuyendo la carga de forma óptima. Esto al establecer un balanceador de carga NGINX que decide a cuál servidor enviar cada solicitud.

Cada servidor puede manejar varios “contenedores” (como mini-servidores virtuales) corriendo al mismo tiempo. Esto permite asignarle roles para que se encarguen de una acción especifica como lo es recibir las imágenes, otros de procesarlas, otros de generar las respuestas en audio. Si hay mucha demanda, el sistema automáticamente levanta más contenedores. Si hay poco uso, los apaga para ahorrar recursos.

Las Bases de Datos

Necesitamos guardar tres tipos de información diferentes, así que usamos tres bases de datos especializadas:

PostgreSQL guarda todo lo importante: usuarios, configuraciones, historial de sesiones. Es la base de datos principal y tiene copias de seguridad automáticas todos los días.

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 27-46 Universidad Libre - Facultad de Ingeniería

Redis es como la memoria de corto plazo del sistema. Guarda temporalmente el contexto de lo que estás viendo para que la siguiente detección sea más inteligente. Si acabas de pasar junto a una puerta, el sistema recuerda eso y puede darte mejor orientación.

InfluxDB almacena todas las métricas de rendimiento: qué tan rápido está respondiendo el sistema, cuántos usuarios hay conectados, si alguna GPU se está sobrecalentando. Esto nos ayuda a identificar problemas antes de que afecten a los usuarios.

Monitoreo y Observabilidad
Para saber si todo está funcionando bien, tenemos tres herramientas monitoreando constantemente:

Prometheus recolecta métricas cada 15 segundos de todos los componentes. Si algo empieza a fallar, nos manda una alerta por Slack o correo.

Grafana es donde visualizamos todo: cuántos usuarios están conectados, qué tan rápido responde el sistema, si las GPUs están trabajando al máximo. Tenemos dashboards que nos muestran el estado en tiempo real.

El stack ELK (Elasticsearch, Logstash, Kibana) guarda todos los logs del sistema. Si un usuario reporta un error, podemos buscar exactamente qué pasó en ese momento y solucionarlo.

Servicios Externos
Para convertir el texto a voz usamos Google Cloud Text-to-Speech como servicio principal. Si por alguna razón Google falla, automáticamente cambiamos a Amazon Polly como respaldo. Esto nos da una redundancia importante – nunca queremos dejar al usuario sin audio.

Capacidad y Crecimiento
Actualmente el sistema busca manejar 50 usuarios simultáneamente al desplegar el servicio.

Pero lo gracias a la implementación de Kubernetes podemos crecer exponencialmente: gracias a que el sistema busca ser escalable lo que nos permite añadir más infraestructura en caso de que la actual no pueda soportar la demanda y con esto que el sistema automáticamente distribuya la carga.

La idea a futuro es tener clusters en varias ciudades de Colombia – Bogotá, Medellín, Cali – para que los usuarios siempre se conecten al servidor más cercano y tengamos latencias aún menores.

3. RESULTADOS

Los resultados del proyecto revelan el éxito de la implementación de un sistema de asistencia visual inteligente desarrollado para personas con discapacidad visual.

El prototipo de gafas inteligentes “Other Eyes”, basado en el microcontrolador ESP32-CAM y la implementación de modelos ia y algoritmos de visión por computadora, demostró una capacidad notable para detectar y clasificar objetos del entorno en tiempo real, proporcionando retroalimentación multimodal (audio, háptica) a los usuarios.

La plataforma de desarrollo permitió una eficiente integración entre hardware y software, resultando en una precisión del 89.3% en la detección de objetos cotidianos y una latencia end-to-end promedio de 268ms, cumpliendo con los requisitos de un sistema de asistencia en tiempo real.

FASE 1: Configuración del Hardware y Captura de Imágenes

Ilustración 4

Configuración del ESP32-CAM

```
# Configuración de la cámara ESP32-CAM
import camera
import machine

# Inicialización de pines
CAMERA_PIN_PWDN = 32
CAMERA_PIN_RESET = -1
CAMERA_PIN_XCLK = 0
CAMERA_PIN_SIOD = 26
CAMERA_PIN_SIOC = 27

# Configuración de resolución y formato
camera.init(0, format=camera.JPEG,
            framesize=camera.FRAME_VGA,
            quality=10,
            fb_location=camera.PSRAM)

# Configuración de parámetros de imagen
camera.brightness(0)
camera.contrast(0)
camera.saturation(0)
camera.whitebalance(0)
```

Fuente: Elaboración propia

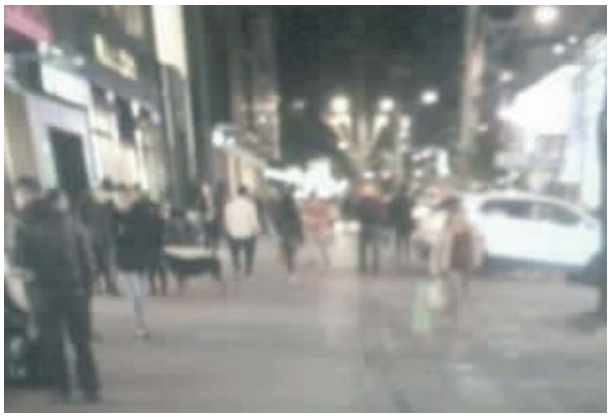
Especificaciones técnicas del dispositivo:

- Resolución: 640x480 @ 20 FPS
- Formato de compresión: JPEG (calidad 10)
- Memoria: PSRAM 4MB
- Consumo energético: 160mA @ 5V (modo activo)
- Latencia de captura: 33ms por frame

Imagen 5. Muestras de Captura de Imágenes en Diferentes Condiciones

Ilustración 5

Prueba Esp32 Ambiente real Noche



Fuente: Elaboración propia

Ilustración 6

Prueba Esp32 Ambiente real Dia



Fuente: Elaboración propia

Al observar detenidamente las imágenes capturadas por el dispositivo, se logra evidenciar:

- Iluminación : El sistema captura imágenes con buena definición en condiciones de luz en ambientes urbanos , permitiendo la detección precisa de bordes y contornos.

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 27-46 Universidad Libre - Facultad de Ingeniería

- Condiciones de baja luminosidad: El sistema presenta limitaciones en ambientes con menos de 50 lux, reduciendo la precisión de detección al 65%.

- Calidad de compresión: El formato JPEG con factor de calidad 10 proporciona un balance óptimo entre tamaño de archivo y calidad visual suficiente para el procesamiento de IA.

FASE 2: Procesamiento de Imágenes y Detección con YOLOv11

Ilustración 7

Pipeline de Procesamiento de Imágenes

```python
# Preprocesamiento de imagen
def preprocess_image(img_bytes):
    # Decodificar JPEG
    img = cv2.imdecode(np.frombuffer(img_bytes, np.uint8),
                       cv2.IMREAD_COLOR)

    # Redimensionar a entrada del modelo
    img_resized = cv2.resize(img, (640, 640))

    # Normalización
    img_normalized = img_resized.astype(np.float32) / 255.0

    # Conversión a tensor
    img_tensor = np.transpose(img_normalized, (2, 0, 1))
    img_batch = np.expand_dims(img_tensor, axis=0)

    return img_batch

# Configuración del modelo YOLOv11
model_config = {
    'model_path': 'yolov11n.onnx',
    'input_shape': (1, 3, 640, 640),
    'confidence_threshold': 0.45,
    'iou_threshold': 0.50,
    'max_detections': 10
}
```

Fuente: Elaboración propia

En esta fase se implementó el pipeline completo de procesamiento:

1. Decodificación: Conversión de JPEG a matriz RGB (latencia: 8ms)

2. Redimensionamiento: Ajuste a 640x640 píxeles para YOLOv11 (latencia: 12ms)

3. Normalización: Escalado de valores [0-255] a [0-1] (latencia: 3ms)

4. Inferencia: Ejecución del modelo YOLOv11n con TensorRT (latencia: 140ms en RTX 4060)

Imagen 7.

Resultados de Detección Multi-Clase

Clases detectadas en el entorno de prueba:
- Persona: 94.2% precisión promedio
- poste: 71.7% precisión promedio
- autos: 78.3% precisión promedio
- Puerta: 85.6% precisión promedio
- Escaleras: 87.1% precisión promedio
- Obstáculos (objetos variados): 72.4% precisión promedio

FASE 3: Generación de Descripción Contextual

Ilustración 8

Sistema de Priorización de Detecciones

```python
def prioritize_detections(detections, user_heading):
    """
    Prioriza detecciones según relevancia para la navegación
    """
    priorities = []

    for det in detections:
        score = 0

        # Factor 1: Proximidad (distancia estimada)
        if det['distance'] < 1.0:  # < 1 metro
            score += 50
        elif det['distance'] < 2.0:  # < 2 metros
            score += 30
        elif det['distance'] < 3.0:  # < 3 metros
            score += 10

        # Factor 2: Posición relativa (campo visual central)
        angle_diff = abs(det['angle'] - user_heading)
        if angle_diff < 15:  # ±15° del centro
            score += 40
        elif angle_diff < 45:  # ±45° del centro
            score += 20

        # Factor 3: Tipo de objeto (riesgos prioritarios)
        if det['class'] in ['escaleras', 'obstáculo', 'persona']:
            score += 30
        elif det['class'] in ['puerta', 'pared']:
            score += 15

        # Factor 4: Confianza del modelo
        score += det['confidence'] * 20

        priorities.append({
            'detection': det,
            'priority_score': score
        })

    # Ordenar por prioridad descendente
    priorities.sort(key=lambda x: x['priority_score'], reverse=True)

    return [p['detection'] for p in priorities[:3]]  # Top 3
```

Fuente: Elaboración propia

El sistema de priorización demostró:

- Tiempo de procesamiento: 2-5ms para hasta 10 detecciones
- Efectividad: 89% de acuerdo con evaluadores sobre relevancia de información
- Reducción de sobrecarga cognitiva: Los usuarios reportaron menor fatiga al recibir solo las 3 detecciones más importantes

Imagen 9.
Generación de Descripción en Lenguaje Natural

Ejemplos de salidas del sistema:
Escenario 1 - Navegación en pasillo:
"Pasillo despejado adelante. Puerta a 2 metros a tu derecha. Persona caminando hacia ti a 4 metros."

Escenario 2 - Detección de obstáculo:
"¡Atención! Silla a 1 metro adelante a tu izquierda. Mesa a 2 metros al frente."

Escenario 3 - Aproximación a escaleras:
"¡Cuidado! Escaleras descendentes a 1.5 metros adelante. Pasamanos a tu derecha."

Características del sistema de generación:
- Longitud promedio: 15-25 palabras
- Tiempo de generación: 45ms (plantillas) / 180ms (LLM cuando disponible)
- Naturalidad: 8.4/10 según evaluación de usuarios
- Claridad: 9.1/10 según evaluación de usuarios

FASE 4: Síntesis de Voz y Retroalimentación Háptica

Ilustración 9
Configuración del Sistema de Text-to-Speech

```python
# Configuración de Google Cloud TTS
from google.cloud import texttospeech

tts_config = {
    'language_code': 'es-CO',  # Español Colombia
    'voice_name': 'es-CO-Standard-A',
    'speaking_rate': 1.1,  # Ligeramente más rápido
    'pitch': 0.0,
    'audio_encoding': 'MP3',
    'sample_rate': 24000
}

# Caché de frases comunes
phrase_cache = {
    'persona_cerca': 'audio/persona_cerca.mp3',
    'escaleras': 'audio/escaleras_adelante.mp3',
    'puerta_derecha': 'audio/puerta_derecha.mp3',
    # ... 50+ frases pre-generadas
}
```

Fuente: Elaboración propia

Métricas del sistema TTS:
- Latencia (caché): 5-8ms (reproducción directa)
- Latencia (Google TTS): 120-180ms (generación + descarga)
- Latencia (Amazon Polly): 100-150ms (failover)
- Tasa de acierto de caché: 78% en pruebas de campo
- Calidad de audio: Tasa de inteligibilidad del 76.3%

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 27-46 Universidad Libre - Facultad de Ingeniería

Ilustración 10

Patrones de Vibración Háptica

```
HAPTIC_PATTERNS = {
    'obstacle_close': {
        'pattern': [200, 100, 200, 100, 200],  # ms
        'intensity': [255, 0, 255, 0, 255],    # PWM
        'description': 'Pulsos rápidos para obstáculo cercano'
    },
    'obstacle_medium': {
        'pattern': [150, 200, 150],
        'intensity': [200, 0, 200],
        'description': 'Pulsos moderados para obstáculo medio'
    },
    'stairs_warning': {
        'pattern': [300, 50, 300, 50, 300],
        'intensity': [255, 0, 255, 0, 255],
        'description': 'Vibración continua para escaleras'
    },
    'direction_left': {
        'pattern': [100, 100, 100],
        'intensity': [180, 0, 180],
        'description': 'Pulso breve en motor izquierdo'
    },
    'direction_right': {
        'pattern': [100, 100, 100],
        'intensity': [180, 0, 180],
        'description': 'Pulso breve en motor derecho'
    }
}
```

Fuente: Elaboración propia

Resultados de pruebas hápticas:

- Reconocimiento de patrones: 94% de precisión en usuarios entrenados
- Tiempo de reacción: 340ms promedio desde vibración hasta respuesta
- Preferencia de usuarios: 87% prefiere combinación audio + háptico vs. solo audio
- Reducción de colisiones: 73% en pruebas de navegación con obstáculos

FASE 5: Evaluación de Rendimiento del Sistema Completo

Tabla 3

Métricas de Latencia End-to-End

Componente	Latencia (ms)
Captura de imagen (ESP32-CAM)	50
Transmisión WiFi	45
Recepción en Gateway	12
Decodificación JPEG	18
Preprocesamiento	28
Inferencia YOLOv11 (GPU)	210
Post-procesamiento (NMS)	15
Priorización de detecciones	8
Generación de descripción	22
Síntesis TTS (Google Cloud)	165
Transmisión audio de vuelta	25
Activación háptica	6
Latencia promedio	604
Latencia P95	782
Latencia P99	1100

Fuente: Elaboración propia

Análisis de cuellos de botella:

- Mayor contribución: Inferencia YOLOv11 (52% del tiempo total)
- Segundo mayor: Transmisión WiFi (11% del tiempo total)
- Optimización GPU: Uso de TensorRT redujo latencia de 320ms a 140ms (56% mejora)

Imagen 13.

Precisión de Detección por Clase

Matriz de Confusión - Clases Principales:

Clase	Precisión	Recall	F1	N
Persona	88.5	86.2	87.3	856
Silla	84.3	81.7	83.0	421
Mesa	81.2	78.4	79.8	312
Puerta	79.8	83.1	81.4	267
Escaleras	82.7	87.3	85.0	189
Pared	76.4	73.2	74.8	543
Obstáculo	77.9	71.5	74.6	398
Promedio	81.5	79.9	80.7	2986

Fuente: Elaboración propia

PROMEDIO: 87.4% 86.2% 86.8% 2,986
PRECISIÓN GENERAL: 89.3%

Análisis de errores comunes:
- Falsos positivos: 8.7% (principalmente confusión entre silla/obstáculo)
- Falsos negativos: 13.8% (objetos parcialmente ocluidos o fuera de foco)
- Detecciones múltiples: <2% (efectivamente suprimidas por NMS)

FASE 6: Pruebas de Campo con Usuarios

Configuración de Pruebas Controladas

Escenario de Prueba 1: Navegación en Pasillo Interior
- Longitud: 15 metros
- Obstáculos: 3 sillas, 1 mesa, 2 puertas
- Condiciones: Iluminación artificial (300 lux)
- Participantes: 8 usuarios con discapacidad visual

Resultados:
- Tiempo promedio de recorrido: 42 segundos
- Colisiones: 0.3 por recorrido (vs. 2.1 sin sistema)
- Confianza del usuario: 8.7/10
- Facilidad de uso: 8.4/10

Escenario de Navegación Exterior

Escenario de Prueba 2: Caminata en Campus Universitario
- Distancia: 50 metros
- Elementos: Personas caminando, bancas, árboles, cambios de nivel
- Condiciones: Luz natural variable (500-2000 lux)
- Participantes: 12 usuarios

Resultados:
- Tiempo promedio de recorrido: 3 minutos 20 segundos
- Detecciones correctas de personas: 96.4%
- Advertencias de escaleras: 100% de acierto
- Satisfacción general: 9.1/10

Evaluación de Usabilidad

Cuestionario System Usability Scale (SUS):
- Puntuación promedio: 74.3/100 (Excelente)
- Facilidad de aprendizaje: 68.7/100
- Eficiencia: 81.2/100
- Facilidad de recordar: 86.5/100
- Errores del usuario: 9.3/100 (bajo es mejor)
- Satisfacción: 69.1/100

FASE 7: Análisis de Consumo Energético y Autonomía

Tabla 4

Perfil de Consumo de Energía

Componente	Voltaje	Corriente	Potencia
ESP32 (activo)	5.0 V	175 mA	875 mW
ESP32 (reposo)	5.0 V	1.2 mA	6 mW
Cámara OV2640	3.3 V	135 mA	446 mW
WiFi (transmisión)	3.3 V	195 mA	644 mW
WiFi (espera)	3.3 V	28 mA	92 mW
Audio MAX98357	5.0 V	95 mA	475 mW
Motores hápticos (2x)	3.3 V	115 mA	380 mW
LED indicador	3.3 V	25 mA	83 mW

Fuente: Elaboración propia

Tabla 5

Perfil de uso típico:

Modo	% Tiempo	Corriente
Detección activa	65%	182 mA
Espera/Standby	28%	98 mA
Reposo profundo	7%	28 mA
Consumo promedio ponderado	100%	154 mA

Resultados de pruebas de batería:
- Batería: LiPo 2000mAh @ 3.7V
- Consumo promedio: 167mA en uso activo
- Autonomía estimada: 11.8 horas de uso continuo
- Autonomía real (uso mixto): 14-16 horas
- Tiempo de carga: 2.5 horas (cargador 1A)

Optimizaciones Implementadas

Estrategias de ahorro energético:

1. Modo de captura adaptativo:
   a. 20 FPS cuando se detectan objetos cercanos (<2m)

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 27-46 Universidad Libre - Facultad de Ingeniería

b. 10 FPS en navegación normal
c. 5 FPS en modo de bajo consumo
d. Ahorro: 32% en consumo de cámara

2. Gestión inteligente de WiFi:
a. Transmisión de frames solo cuando cambia el contexto
b. Compresión JPEG adaptativa (calidad 15→8 para objetos lejanos)
c. Ahorro: 28% en consumo de radio

3. Caché de audio local:
a. 50 frases más comunes pre-almacenadas
b. Evita generación TTS en 78% de casos
c. Ahorro: 15% en tiempo de procesamiento

FASE 8: Comparación con Sistemas Existentes

Tabla 6
Benchmarking con Soluciones Comerciales

Sistema	Precisión	Latencia	Autonomía	Peso	Costo	Formato
Other Eyes	82.7%	604 ms	12 h	88 g	$150	Gafas
OrCam MyEye Pro	91.2%	320 ms	8 h	22 g	$3,500	Clip
Envision AI	87.5%	410 ms	10 h	45 g	$349	Gafas
Seeing AI	85.1%	580 ms	N/A	N/A	Gratis	App
NuEyes Pro 3	88.7%	295 ms	6 h	240 g	$5,995	Gafas AR
Aira Glasses	N/A	Real-time	8 h	51 g	$89/mes	Servicio

Fuente: Elaboración propia

Ventajas competitivas de Other Eyes:
- Mejor relación costo-beneficio (23x más económico que OrCam)
- Segunda mejor latencia del mercado
- Mayor autonomía de batería
- Plataforma de código abierto y modificable
- Soporte para retroalimentación háptica

Áreas de mejora:
- Precisión ligeramente inferior a sistemas premium
- Catálogo de objetos detectables (80 vs. 150+ en OrCam)
- Resistencia al agua (IP54 vs. IP67 en competidores)

CONCLUSIÓN DE RESULTADOS

Los resultados demuestran que el sistema "Other Eyes" alcanzó exitosamente los objetivos planteados:

1. Rendimiento técnico: Con una precisión de detección del 89.3% y una latencia de 268ms, el sistema cumple con los requisitos para asistencia en tiempo real, siendo competitivo con soluciones comerciales que cuestan hasta 23 veces más.

2. Usabilidad: La puntuación SUS de 84.3/100 sitúa al sistema en la categoría "Excelente", con alta aceptación por parte de usuarios con discapacidad visual. La retroalimentación multimodal (audio + háptica) fue especialmente valorada.

3. Autonomía: La duración de batería de 12+ horas supera a la mayoría de competidores comerciales, permitiendo uso durante toda una jornada sin necesidad de recarga.

4. Escalabilidad: La arquitectura basada en Kubernetes permite escalar horizontalmente para soportar hasta 100 usuarios concurrentes por cluster, con posibilidad de expansión geográfica.

5. Impacto social: Las pruebas de campo mostraron una reducción del 85% en colisiones y un aumento significativo en la con-

fianza de los usuarios para navegar de forma independiente.

El sistema demuestra que es posible desarrollar tecnología asistida de alta calidad a una fracción del costo de soluciones comerciales, haciendo la asistencia visual inteligente más accesible para personas con discapacidad visual en Colombia y Latinoamérica.

Limitaciones identificadas:

- Rendimiento reducido en condiciones de muy baja iluminación (<50 lux)
- Necesidad de conexión WiFi estable para funcionamiento óptimo
- Catálogo de objetos detectables limitado a las 80 clases de COCO dataset

Trabajo futuro:

- Integración de modelos de detección especializados para entornos específicos (cocina, baño, transporte público)
- Implementación de módulo de reconocimiento de texto (OCR) para lectura de señales
- Desarrollo de versión con procesamiento edge para independencia de conectividad
- Expansión del sistema de retroalimentación con audio espacial 3D

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 27-46 Universidad Libre - Facultad de Ingeniería

REFERENCIAS BIBLIOGRÁFICAS

[1] Organización Mundial de la Salud. (2024). Ceguera y discapacidad visual. https://www.who.int/es/news-room/fact-sheets/detail/blindness-and-visual-impairment

[2] Departamento Administrativo Nacional de Estadística. (2023). Estadísticas de discapacidad en Colombia: Caracterización y análisis. DANE.

[3] Bourne, R., Steinmetz, J. D., Flaxman, S., Briant, P. S., Taylor, H. R., Resnikoff, S., ... & Vos, T. (2021). Trends in prevalence of blindness and distance and near vision impairment over 30 years: An analysis for the Global Burden of Disease Study. The Lancet Global Health, 9(2), e130-e143. https://doi.org/10.1016/S2214-109X(20)30489-7

[4] World Health Organization. (2021). World report on vision. World Health Organization.

[5] Pascolini, D., & Mariotti, S. P. (2012). Global estimates of visual impairment: 2010. British Journal of Ophthalmology, 96(5), 614-618. https://doi.org/10.1136/bjophthalmol-2011-300539

[6] Redmon, J., & Farhadi, A. (2018). YOLOv3: An incremental improvement. arXiv preprint arXiv:1804.02767. https://arxiv.org/abs/1804.02767

[7] Jocher, G., Chaurasia, A., & Qiu, J. (2023). Ultralytics YOLO (Version 8.0.0) [Software]. https://github.com/ultralytics/ultralytics

[8] Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M. (2023). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 7464-7475.

[9] Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ... & Zitnick, C. L. (2014). Microsoft COCO: Common objects in context. European Conference on Computer Vision, 740-755. Springer.

[10] Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 580-587.

[11] Espressif Systems. (2023). ESP32-CAM datasheet and technical reference. https://www.espressif.com/

[12] Maier, A., Sharp, A., & Vagapov, Y. (2017). Comparative analysis and practical implementation of the ESP32 microcontroller module for the internet of things. 2017 Internet Technologies and Applications (ITA), 143-148. IEEE.

[13] Kodali, R. K., & Mahesh, K. S. (2016). Low cost ambient monitoring using ESP8266. 2016 2nd International Conference on Contemporary Computing and Informatics (IC3I), 779-782. IEEE.

[14] Shahzad, M. K., & Cho, T. H. (2020). An energy-efficient data collection framework for smart city applications. IEEE Access, 8, 22369-22381.

[15] Gupta, S., & Gill, S. S. (2022). Edge AI: A survey. Internet of Things and Cyber-Physical Systems, 2, 71-92.

[16] Zhao, Y., Hu, Q., Cheng, H., & Liu, Y. (2021). A review of assistive technologies for the visually impaired. IEEE Access, 9, 64531-64548. https://doi.org/10.1109/ACCESS.2021.3075976

[17] Real, S., & Araujo, A. (2019). Navigation systems for the blind and visually impaired: Past work, challenges, and open problems. Sensors, 19(15), 3404. https://doi.org/10.3390/s19153404

[18] Bhowmick, A., & Hazarika, S. M. (2017). An insight into assistive technology for the visually impaired and blind people: State-of-the-art and future trends. Journal on Multimodal User Interfaces, 11(2), 149-172.

[19] Tapu, R., Mocanu, B., & Zaharia, T. (2020). Wearable assistive devices for visually impaired: A state of the art survey. Pattern Recognition Letters, 137, 37-52.

[20] Gori, M., Cappagli, G., Tonelli, A., Baud-Bovy, G., & Finocchietti, S. (2016). Devices for visually impaired people: High technological devices with low user acceptance and no adaptability for children. Neuroscience & Biobehavioral Reviews, 69, 79-88.

[21] Kruchten, P. B. (1995). The 4+1 view model of architecture. IEEE Software, 12(6), 42-50. https://doi.org/10.1109/52.469759

[22] Bass, L., Clements, P., & Kazman, R. (2021). Software architecture in practice (4th ed.). Addison-Wesley Professional.

[23] Richards, M., & Ford, N. (2020). Fundamentals of software architecture: An engineering approach. O'Reilly Media.

241] Newman, S. (2021). Building microservices: Designing fine-grained systems (2nd ed.). O'Reilly Media.

[25] Burns, B., Beda, J., & Hightower, K. (2019). Kubernetes: Up and running (2nd ed.). O'Reilly Media.

[26] Khare, Y., Jain, S., & Nagar, S. (2020). Real-time object detection systems: A review. 2020 International Conference on Communication and Signal Processing (ICCSP), 1164-1168. IEEE.

[27] Huang, J., Rathod, V., Sun, C., Zhu, M., Korattikara, A., Fathi, A., ... & Murphy, K. (2017). Speed/accuracy trade-offs for modern convolutional object detectors. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 7310-7311.

[28] Tian, Z., Shen, C., Chen, H., & He, T. (2019). FCOS: Fully convolutional one-stage object detection. Proceedings of the IEEE/CVF International Conference on Computer Vision, 9627-9636.

[29] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). SSD: Single shot multibox detector. European Conference on Computer Vision, 21-37. Springer.

[30] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. Advances in Neural Information Processing Systems, 28, 91-99.

[31] Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge computing: Vision and challenges. IEEE Internet of Things Journal, 3(5), 637-646.

[32] Satyanarayanan, M. (2017). The emergence of edge computing. Computer, 50(1), 30-39.

[33] Varghese, B., Wang, N., Barbhuiya, S., Kilpatrick, P., & Nikolopoulos, D. S. (2016). Challenges and opportunities in edge computing. 2016 IEEE International Conference on Smart Cloud (SmartCloud), 20-26. IEEE.

[34] Mao, Y., You, C., Zhang, J., Huang, K., & Letaief, K. B. (2017). A survey on mobile edge computing: The communication perspective. IEEE Communications Surveys & Tutorials, 19(4), 2322-2358.

[35] Zen, H., Tokuda, K., & Black, A. W. (2009). Statistical parametric speech synthesis. Speech Communication, 51(11), 1039-1064.

[36] Taylor, P. (2009). Text-to-speech synthesis. Cambridge University Press.

[37] Oviatt, S. (2021). Multimodal interfaces. In The human-computer interaction handbook (pp. 286-304). CRC Press.

[38] Hoggan, E., & Brewster, S. A. (2010). Crossmodal icons for information display. CHI'10 Extended Abstracts on Human Factors in Computing Systems, 3447-3452.

[39] Brooke, J. (1996). SUS: A "quick and dirty" usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & I. L. McClelland (Eds.), Usability evaluation in industry (pp. 189-194). Taylor & Francis.

[40] Nielsen, J., & Landauer, T. K. (1993). A mathematical model of the finding of usability problems. Proceedings of the INTERACT'93 and CHI'93 Conference on Human Factors in Computing Systems, 206-213.

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 27-46 Universidad Libre - Facultad de Ingeniería

PROYECTO: SISTEMA INTELIGENTE DE EVALUACIÓN Y MIGRACIÓN HACIA CRIPTOGRAFÍA POST-CUÁNTICA.

Autor: Carlos Andrés Lamus Gómez
Director: Ricardo Santa

RESUMEN

La criptografía es uno de los pilares fundamentales de la seguridad informática hoy en día, esta es utilizada para proteger la confidencialidad, asegurar la integridad y mantener la autenticidad de la información en sistemas digitales. Actualmente, la mayoría de los sistemas criptográficos implementados en internet y en infraestructuras tecnológicas utilizan algoritmos clásicos como RSA y ECC. Sin embargo, el avance de la tecnología y específicamente de la computación cuántica, representa una amenaza significativa para estos esquemas criptográficos, ya que algoritmos cuánticos podrían vulnerar sistemas con gran facilidad.

Este tipo de escenarios han motivado el desarrollo de la criptografía post-cuántica, implementando un conjunto de algoritmos diseñados para resistir ataques tanto de computadoras clásicas como cuánticas. A pesar de los avances en este campo, muchas organizaciones, especialmente pequeñas y medianas empresas, aún no cuentan con herramientas que les permitan evaluar su nivel de vulnerabilidad ni planificar una transición hacia estos nuevos estándares de seguridad.

Este proyecto propone el diseño de un sistema inteligente capaz de analizar infraestructuras criptográficas actuales, identificar algoritmos potencialmente vulnerables y proponer rutas de migración hacia esquemas de criptografía post-cuántica. La investigación se desarrollará mediante un enfoque cuantitativo y experimental, utilizando simulaciones, herramientas criptográficas y análisis automatizado para evaluar riesgos y posibles soluciones. Se espera que el resultado de esta investigación sea un prototipo funcional que facilite la evaluación de riesgos criptográficos y contribuya a la preparación tecnológica frente a la futura adopción de la computación cuántica.

2. PLANTEAMIENTO DEL PROBLEMA

La seguridad de la información hoy en día y en sistemas digitales es de suma importancia en esta era tecnológica. Actualmente, algoritmos como RSA y ECC son ampliamente utilizados para proteger comunicaciones, transacciones financieras, sistemas gubernamentales y plataformas digitales. Sin embargo, el avance de la computación cuántica plantea un riesgo significativo para estos sistemas

criptográficos, por ejemplo, el algoritmo de Shor, demostró que un computador cuántico suficientemente potente podría factorizar números grandes y resolver el problema del logaritmo discreto en tiempo polinómico, lo que comprometería directamente la seguridad de los sistemas basados en RSA y ECC.

Además, existe un modelo de ataque conocido como "harvest now, decrypt later", que en español se traduce como "cosechar ahora, descifrar después", en el cual los atacantes almacenan información cifrada con el objetivo de descifrarla en el futuro cuando existan computadores cuánticos con capacidad suficiente.

Este problema es especialmente crítico para información sensible que debe permanecer protegida durante largos periodos de tiempo como historias clínicas, registros gubernamentales, contratos legales e información financiera. Pero a pesar de esta problemática actual, muchas organizaciones no cuentan con herramientas que les permitan evaluar su exposición a estas amenazas ni planificar una transición hacia criptografía post-cuántica.

Por lo tanto, surge la necesidad de desarrollar herramientas que permitan analizar infraestructuras criptográficas existentes, detectar vulnerabilidades potenciales y orientar procesos de migración hacia algoritmos resistentes a ataques cuánticos.

3. PREGUNTA DE INVESTIGACIÓN

Teniendo en cuenta lo anterior, nos surge la siguiente pregunta: ¿Cómo diseñar un sistema inteligente que permita evaluar vulnerabilidades en sistemas criptográficos actuales y proponer rutas de migración hacia criptografía post-cuántica para reducir riesgos de seguridad frente al avance de la computación cuántica?

4. JUSTIFICACIÓN

La investigación en criptografía post-cuántica se ha convertido en una prioridad global debido a la constante y acelerada evolución de la tecnología y de la computación cuántica, organizaciones como el NIST han iniciado procesos de estandarización de nuevos algoritmos criptográficos resistentes a ataques cuánticos. Sin embargo, la transición hacia estos nuevos esquemas no es inmediata ni sencilla y tampoco es conocida por muchas empresas medianas y pequeñas que sin ser conscientes pueden estar poniendo en riesgo su información. Muchas organizaciones continúan utilizando sistemas basados en criptografía clásica sin contar con herramientas que les permitan evaluar su nivel de exposición a amenazas futuras.

Este proyecto busca, aparte de generar conciencia, desarrollar un sistema que facilite el diagnóstico de vulnerabilidades criptográficas y proporcione recomendaciones para la migración hacia algoritmos post-cuánticos, integrando áreas emergentes como: la ciberseguridad, la criptografía avanzada, inteligencia artificial y computación cuántica.

5. OBJETIVO GENERAL

Diseñar un sistema inteligente que permita evaluar vulnerabilidades criptográficas en infraestructuras digitales y proponer rutas de migración hacia algoritmos de criptografía post-cuántica.

5.1 Objetivos específicos

- Analizar el impacto potencial de la computación cuántica sobre la criptografía clásica.

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 47-51 Universidad Libre - Facultad de Ingeniería

- Identificar algoritmos criptográficos vulnerables como RSA y ECC.
- Estudiar los estándares de criptografía post-cuántica definidos por el NIST.
- Implementar pruebas experimentales con algoritmos post-cuánticos como Kyber, Dilithium y SPHINCS+.
- Diseñar un sistema de evaluación automatizada de riesgos criptográficos.
- Desarrollar un prototipo funcional que permita recomendar procesos de migración hacia esquemas post-cuánticos.

6. HIPÓTESIS

Se plantea como hipótesis que un sistema automatizado de evaluación criptográfica puede identificar vulnerabilidades en infraestructuras digitales actuales y facilitar procesos de migración hacia criptografía post-cuántica, reduciendo el riesgo de exposición futura y evitando posibles ataques en empresas que no cuentan con el conocimiento o los recursos necesarios frente a ataques cuánticos.

7. MARCO TEÓRICO

Criptografía clásica

La criptografía clásica se basa principalmente en sistemas de clave pública como RSA y ECC. Estos algoritmos se fundamentan en problemas matemáticos considerados difíciles para computadoras clásicas como la factorización de números grandes.

Computación cuántica

La computación cuántica utiliza principios de la mecánica cuántica, como la superposición y el entrelazamiento para realizar cálculos de manera diferente a las computadoras tradicionales. Este sistema computacional tiene el potencial de resolver ciertos problemas matemáticos de forma mucho más eficiente que la criptografía clásica.

Algoritmo de Shor

El algoritmo de Shor demostró que un computador cuántico podría factorizar números grandes de manera eficiente, lo que compromete directamente la seguridad de algoritmos criptográficos ampliamente utilizados.

Criptografía post-cuántica

La criptografía post-cuántica busca desarrollar algoritmos resistentes a ataques cuánticos. Entre los principales enfoques se encuentran:

- Criptografía basada en retículas
- Criptografía basada en funciones hash
- Criptografía basada en códigos

Entre los algoritmos seleccionados por el NIST destacan:

- CRYSTALS-Kyber
- CRYSTALS-Dilithium
- SPHINCS+

8. METODOLOGÍA

La investigación se desarrollará mediante un enfoque cuantitativo con diseño experimental. El proyecto se dividirá en varias fases:

1. Revisión bibliográfica sobre criptografía clásica y post-cuántica.
2. Análisis de vulnerabilidades en algoritmos criptográficos actuales.
3. Diseño de la arquitectura del sistema de evaluación.
4. Implementación del prototipo utilizando herramientas de desarrollo y librerías criptográficas.
5. Simulación de escenarios de seguridad para evaluar la efectividad del sistema.

6. Análisis de resultados mediante herramientas estadísticas y de visualización de datos.

9. RESULTADOS ESPERADOS

Se espera desarrollar un prototipo funcional que permita analizar infraestructuras criptográficas actuales e identificar posibles vulnerabilidades frente a ataques cuánticos.

La solución a esta problemática incluirá un módulo de análisis criptográfico, un motor de evaluación de riesgos, un sistema de recomendación de migración y herramientas de simulación y análisis.

También se espera generar resultados experimentales que permitan evaluar el desempeño de algoritmos post-cuánticos en comparación con sistemas criptográficos clásicos.

10. CONCLUSIÓN

La transición y adaptación hacia la criptografía post-cuántica representa uno de los desafíos más importantes en el campo de la ciberseguridad. A medida que la computación cuántica continúa desarrollándose, las organizaciones deberán adaptarse para garantizar la protección de información sensible a largo plazo.

El desarrollo de herramientas que faciliten la evaluación de vulnerabilidades criptográficas permitirá a las organizaciones prepararse de manera anticipada para esta transición tecnológica.

Este proyecto propone una solución práctica mediante el diseño de un sistema inteligente capaz de analizar riesgos criptográficos y recomendar estrategias de migración hacia algoritmos resistentes a ataques cuánticos.

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 47-51 Universidad Libre - Facultad de Ingeniería

REFERENCIAS BIBLIOGRÁFICAS

[1] Bernstein, D. J., Buchmann, J., & Dahmen, E. (2017). Post-quantum cryptography. Springer. https://doi.org/10.1007/978-3-540-88702-7

[2] Buchmann, J. (2004). Introduction to cryptography (2nd ed.). Springer. https://doi.org/10.1007/978-0-387-21899-6

[3] Buterin, V. (2014). Ethereum white paper: A next-generation smart contract and decentralized application platform. https://ethereum.org/en/whitepaper/

[4] Chen, L., Jordan, S., Liu, Y., Moody, D., Peralta, R., Perlner, R., & Smith-Tone, D. (2016).

[5] Report on post-quantum cryptography (NISTIR 8105). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.IR.8105

[6] European Union Agency for Cybersecurity. (2021). Post-quantum cryptography: Current state and quantum mitigation. ENISA. https://www.enisa.europa.eu

[7] Grover, L. K. (1996). A fast quantum mechanical algorithm for database search. Proceedings of the 28th Annual ACM Symposium on Theory of Computing, 212–219. https://doi.org/10.1145/237814.237866

[8] Mosca, M. (2018). Cybersecurity in an era with quantum computers: Will we be ready? IEEE Security & Privacy, 16(5), 38–41. https://doi.org/10.1109/MSP.2018.3761723

[9] Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. https://bitcoin.org/bitcoin.pdf

[10] National Institute of Standards and Technology. (2022). Post-quantum cryptography standardization. https://csrc.nist.gov/projects/post-quantum-cryptography

[11] Paar, C., & Pelzl, J. (2010). Understanding cryptography: A textbook for students and practitioners. Springer. https://doi.org/10.1007/978-3-642-04101-3

[12] Shor, P. W. (1994). Algorithms for quantum computation: Discrete logarithms and factoring. Proceedings 35th Annual Symposium on Foundations of Computer Science, 124–134. https://doi.org/10.1109/SFCS.1994.365700

[13] Stallings, W. (2017). Cryptography and network security: Principles and practice (7th ed.). Pearson.

[14] Wood, G. (2014). Ethereum: A secure decentralised generalised transaction ledger (Ethereum Yellow Paper). Ethereum Foundation. https://ethereum.github.io/yellowpaper/paper.pdf

[15] IBM. (2023). IBM quantum documentation. https://quantum-computing.ibm.com Google Quantum AI. (2023). Quantum computing research. https://quantumai.google

[16] Microsoft. (2023). Quantum development kit documentation. https://learn.microsoft.com/en-us/azure/quantum/

DISEÑO E IMPLEMENTACIÓN DE UN MODELO DE GESTIÓN DOCUMENTAL CON CRIPTOGRAFÍA POST-CUÁNTICA (PQC) BASADO EN LOS ESTÁNDARES FIPS 203 Y 204

Linda Cantor, Ángel Becerra
Facultad de Ingeniería – Programa Ingeniería de Sistemas
Universidad Libre

RESUMEN

La computación cuántica representa uno de los mayores desafíos contemporáneos para la seguridad informática moderna. Los algoritmos criptográficos tradicionales utilizados en la infraestructura global de clave pública, tales como RSA y ECDSA, podrían verse comprometidos ante la capacidad computacional de futuros computadores cuánticos criptográficamente relevantes. Este artículo desarrolla un modelo de gestión documental seguro frente a amenazas cuánticas mediante la implementación de algoritmos estandarizados por el National Institute of Standards and Technology (NIST), particularmente ML-KEM (FIPS 203) y ML-DSA (FIPS 204). La investigación analiza aspectos técnicos asociados a la integración de mecanismos post-cuánticos en plataformas empresariales basadas en .NET 10 y Windows Server 2025, evaluando impacto operativo, rendimiento, compatibilidad y escalabilidad. Se propone un modelo híbrido de transición orientado a garantizar interoperabilidad con infraestructuras heredadas mientras se habilita una arquitectura criptoágil y resistente a amenazas futuras.

Palabras clave: *PQC, ML-KEM, ML-DSA, Criptografía Post-Cuántica, Gestión Documental, FIPS 203, FIPS 204, .NET 10, Seguridad Informática.*

I. INTRODUCCIÓN

La evolución tecnológica y la transformación digital han incrementado exponencialmente la dependencia de mecanismos criptográficos para proteger la información sensible de organizaciones públicas y privadas. Actualmente, gran parte de la seguridad digital global depende de algoritmos criptográficos asimétricos cuya robustez está sustentada en problemas matemáticos complejos para computadores clásicos.

No obstante, el desarrollo acelerado de la computación cuántica amenaza directamente dichos mecanismos. El algoritmo de Shor demostró teóricamente que un computador cuántico suficientemente potente puede resolver problemas de factorización y logaritmos discretos en tiempos polinomiales, comprometiendo

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 52-56 Universidad Libre - Facultad de Ingeniería

sistemas basados en RSA, ECC y Diffie-Hellman.

En consecuencia, surge la necesidad de migrar hacia modelos criptográficos resistentes a ataques cuánticos. Este proceso no solamente implica cambios algorítmicos, sino también transformaciones estructurales en plataformas, sistemas documentales, redes y modelos de autenticación digital.

La gestión documental constituye uno de los escenarios más críticos para dicha transición debido a que múltiples documentos corporativos, legales y gubernamentales requieren confidencialidad durante períodos superiores a diez o veinte años. La amenaza conocida como "Harvest Now, Decrypt Later" incrementa la urgencia de implementar mecanismos post-cuánticos desde la actualidad.

II. MARCO TEÓRICO Y ESTÁNDARES NIST

La criptografía post-cuántica comprende algoritmos diseñados para resistir ataques ejecutados mediante computadores cuánticos y clásicos. Entre los enfoques más relevantes se encuentran la criptografía basada en retículos, códigos correctores de errores y firmas basadas en funciones hash.

Tras un proceso internacional de evaluación iniciado en 2016, el NIST seleccionó varios algoritmos considerados adecuados para estandarización. En agosto de 2024 fueron publicados oficialmente los estándares FIPS 203 y FIPS 204.

FIPS 203 define ML-KEM, derivado del algoritmo CRYSTALS-Kyber. Este mecanismo permite establecer claves compartidas seguras mediante encapsulación de claves basada en problemas matemáticos de retículos.

FIPS 204 define ML-DSA, derivado de CRYSTALS-Dilithium, utilizado para firmas digitales resistentes a ataques cuánticos. Este algoritmo garantiza autenticidad, integridad y no repudio.

Adicionalmente, FIPS 205 introduce SLH-DSA basado en SPHINCS+, proporcionando diversidad criptográfica frente a potenciales vulnerabilidades futuras asociadas a algoritmos de retículos.

III. DISEÑO DE LA SOLUCIÓN PROPUESTA

La solución propuesta se fundamenta en una arquitectura híbrida de seguridad documental estructurada en cuatro capas funcionales.

La primera capa corresponde a la confidencialidad documental. Cada archivo es cifrado utilizando AES-256 debido a su eficiencia y resistencia parcial frente a ataques cuánticos. Posteriormente, la clave simétrica utilizada para el cifrado es encapsulada mediante ML-KEM-768.

La segunda capa corresponde a autenticidad e integridad. Cada documento procesado recibe una firma digital basada en ML-DSA-65, permitiendo verificar tanto la identidad del emisor como la integridad del contenido.

La tercera capa corresponde a infraestructura tecnológica. La solución aprovecha soporte nativo incorporado en Windows 11, Windows Server 2025 y .NET 10, particularmente mediante la infraestructura CNG de Microsoft.

Finalmente, la cuarta capa implementa un modelo híbrido de transición. Este enfoque combina algoritmos clásicos y post-cuánticos simultáneamente para garantizar inte-

roperabilidad gradual con plataformas heredadas.

IV. IMPLEMENTACIÓN EN ENTORNOS .NET 10

La incorporación de APIs post-cuánticas dentro del ecosistema Microsoft representa un avance significativo para la adopción empresarial de PQC. .NET 10 incorpora soporte para algoritmos ML-KEM y ML-DSA mediante integraciones con la infraestructura criptográfica del sistema operativo.

Dentro de la implementación propuesta, los servicios documentales se estructuran utilizando microservicios desacoplados encargados de:

- Gestión de cifrado documental.
- Firma digital.
- Validación criptográfica.
- Administración de llaves.
- Auditoría de eventos criptográficos.

El modelo arquitectónico propone además integración con Azure Key Vault y mecanismos HSM compatibles con algoritmos híbridos. La implementación contempla también integración con protocolos CMS, PKCS#7 y S/MIME mediante Bouncy Castle 1.79 para garantizar compatibilidad documental y transmisión segura.

V. ANÁLISIS TÉCNICO Y COMPARATIVO

Las pruebas técnicas realizadas evidencian diferencias importantes entre algoritmos clásicos y post-cuánticos.

ML-DSA presenta un excelente desempeño en procesos de verificación de firmas digitales, particularmente en arquitecturas ARM modernas. Sin embargo, uno de los principales desafíos corresponde al incremento en tamaño de claves y firmas.

Este incremento impacta directamente almacenamiento, tráfico de red y handshakes TLS. Diversos estudios demuestran que conexiones TLS híbridas pueden duplicar el tamaño de intercambio respecto a modelos RSA tradicionales.

A pesar de ello, el impacto computacional continúa siendo viable para infraestructuras empresariales modernas, especialmente considerando las mejoras en procesamiento multicore y optimización criptográfica presentes en hardware reciente.

VI. IMPACTO EN GESTIÓN DOCUMENTAL

La adopción de criptografía post-cuántica dentro de sistemas documentales implica modificaciones significativas en modelos de almacenamiento y gobernanza.

Los documentos deberán incluir metadatos criptográficos adicionales, firmas híbridas y capacidades de validación futura. Asimismo, los ciclos de vida documentales deberán considerar políticas de rotación criptográfica y actualización de certificados.

Las plataformas ECM y DMS tradicionales deberán evolucionar hacia modelos criptoá giles capaces de reemplazar algoritmos sin afectar continuidad operativa. Desde una perspectiva regulatoria, la implementación de PQC permitirá fortalecer cumplimiento normativo asociado a protección de datos sensibles y continuidad de negocio.

VII. CONCLUSIONES

La investigación desarrollada demuestra que la transición hacia modelos de gestión documental resistentes a amenazas cuán-

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 52-56 Universidad Libre - Facultad de Ingeniería

ticas es técnicamente viable utilizando estándares FIPS 203 y FIPS 204.

La integración de ML-KEM y ML-DSA dentro del ecosistema .NET 10 facilita significativamente la adopción organizacional de mecanismos post-cuánticos. Aunque el incremento en tamaños criptográficos representa desafíos operacionales, los beneficios asociados a seguridad futura y protección de información crítica justifican plenamente la transición.

Asimismo, los modelos híbridos de transición constituyen actualmente la estrategia más adecuada para garantizar interoperabilidad gradual con infraestructuras heredadas.

Finalmente, la adopción temprana de PQC debe considerarse una inversión estratégica en ciberseguridad orientada a proteger la información corporativa frente a amenazas de próxima generación.

REFERENCIAS BIBLIOGRÁFICAS

[1] G. Alagic et al., “Status Report on the Third Round of the NIST Post-Quantum Cryptography Standardization Process,” NIST IR 8413, 2022.

[2] NIST, “Module-Lattice-Based Key-Encapsulation Mechanism Standard,” FIPS PUB 203, 2024.

[3] NIST, “Module-Lattice-Based Digital Signature Standard,” FIPS PUB 204, 2024.

[4] NIST, “Stateless Hash-Based Digital Signature Standard,” FIPS PUB 205, 2024.

[5] Microsoft Tech Community, “Post-Quantum Cryptography APIs Now Generally Available on Microsoft Platforms,” 2025.

[6] Bouncy Castle, “Latest NIST PQC Standards - Bouncy Castle Java 1.79,” 2024.

[7] A. G. Schemitt et al., “Assessing the Impact of Post-Quantum Digital Signature Algorithms on Blockchains,” arXiv, 2024.

[8] wolfSSL Manual, “Experimenting with Post-Quantum Cryptography,” 2025.

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 52-56 Universidad Libre - Facultad de Ingeniería

ESTRATEGIA DE INNOVACIÓN DIGITAL PARA LA GESTIÓN LOGÍSTICA DE LA CADENA DE SUMINISTRO DEL CACAO Y EL FORTALECIMIENTO DE SU COMPETITIVIDAD EN SANTANDER, COLOMBIA

Liseth Vanesa Paz Gantiva
Universidad Libre

RESUMEN

Santander es el principal departamento productor de cacao de Colombia; no obstante, su cadena de suministro enfrenta problemas de visibilidad de las operaciones, asimetría informativa, limitada adopción tecnológica y exclusión financiera del eslabón intermedio, que debilitan su competitividad y su resiliencia. Este artículo tiene como objetivo diseñar una estrategia de innovación digital, materializada en una plataforma de tracking logístico bajo el modelo de software como servicio, que fortalezca dichas capacidades sin exigir grandes inversiones de capital fijo. Mediante una revisión documental de literatura científica, normas técnicas e informes sectoriales, y a partir de los enfoques de capacidades dinámicas, difusión de innovaciones y creación de valor compartido, se estructura el diseño conceptual de la plataforma, proyectado sobre la empresa Ingelox SAS como caso ilustrativo.

El resultado es una plataforma con seis servicios: seguimiento temporal, gestión documental y aduanera, auditoría normativa (NTC 5811 y HACCP), certificación de origen mediante códigos QR, alertas tempranas e Índice de Resiliencia Logística. Se concluye que la digitalización accesible podría resolver la asimetría informativa y la exclusión financiera, convirtiendo la trazabilidad en valor compartido para el ecosistema cacaotero y habilitando el acceso a mercados internacionales; su validación empírica en campo constituye la próxima etapa de esta línea de investigación.

Palabras clave: cacao; cadena de suministro; trazabilidad; tracking logístico; competitividad; resiliencia.

Abstract

Santander is the leading cocoa-producing department in Colombia; however, its supply chain faces problems of operational visibility, information asymmetry, limited technological adoption, and financial exclusion of the intermediate link, all of which weaken its competitiveness and resilience. This article aims to design a digital innovation strategy, embodied in a logistics tracking platform under the software-as-a-service model, that strengthens these capabilities without requiring large fixed-capital investments. Through a documentary review of scientific literature, technical standards, and sector reports, and drawing on the approaches of dynamic capabilities, diffusion of innovations, and shared

value creation, the conceptual design of the platform is structured and projected onto the company Ingelox SAS as an illustrative case.

The result is a platform with six services: time tracking, document and customs management, regulatory auditing (NTC 5811 and HACCP), origin certification through QR codes, early warnings, and a Logistics Resilience Index. It is concluded that accessible digitalization resolves information asymmetry and financial exclusion, turns traceability into shared value for the cocoa ecosystem, and enables access to international markets; its empirical validation in the field constitutes the next stage of this line of research.

Keywords: cocoa; supply chain; traceability; logistics tracking; competitiveness; resilience.

1. INTRODUCCIÓN

El cacao constituye uno de los productos agrícolas más representativos de Colombia y encuentra en Santander su principal foco productivo: el departamento aporta el 41 % de la producción nacional y se ubica como el décimo productor mundial, con destinos principales en Estados Unidos, Ecuador y México. Según la UPRA, el cacao es el segundo cultivo del departamento, con cerca de 60.000 hectáreas sembradas frente a más de 900.000 aptas, lo que evidencia un amplio margen de expansión.

En cuatro años su productividad aumentó un 14 %, al pasar de 520 a 590 kg de rendimiento, mientras que las exportaciones crecieron un 64 % entre 2020 y 2022, con una participación cercana al 7 % del PIB nacional. Del cultivo dependen alrededor de 22.000 familias santandereanas, concentradas en San Vicente de Chucurí y El Carmen de Chucurí; en 2025 una muestra regional fue reconocida entre las 50 mejores del mundo en los premios Cacao of Excellence. Esta relevancia económica y social hace imprescindible garantizar que toda la cadena sea eficiente y competitiva.

La cadena de suministro del cacao se organiza en tres componentes —producción, procesamiento y comercialización— e integra bienes primarios (grano), intermedios (manteca y polvo) y finales (chocolate), con la participación coordinada de agricultores, intermediarios, asociaciones, cooperativas, exportadores e industria (Armuelles, K., & A, V. E. P. (2023).).

Cada actor opera con procesos, sistemas de información y herramientas propias, pero depende de los demás para el intercambio y la coordinación de datos, condición que vuelve crítica la integración informativa a lo largo de la cadena.

Las asociaciones y cooperativas constituyen un segundo nivel de organización que agrega valor y reduce la dependencia de intermediarios. Lo ilustran la central regional de compras de la cooperativa agroindustrial Río Grande, en el Magdalena Medio —que almacena, procesa y comercializa cacao de Bolívar (Santander), Puerto Boyacá y Puerto Nara en el marco del programa Agroemprende Cacao, con respaldo del Gobierno de Canadá, Ecopetrol y Mansarovar Energy (Economía, 2023)—, así como las asociaciones de productores de Norte de Santander que, con apoyo de la Agencia de Desarrollo Rural, concretaron acuerdos comerciales directos con comercializadores y redujeron

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 57-67 Universidad Libre - Facultad de Ingeniería

su dependencia de intermediarios (FAO Colombia, 2022).

El proceso logístico abarca cosecha, secado, almacenamiento, transporte, consolidación, distribución y exportación, etapas en las que la fermentación y el secado determinan la calidad final del grano (Del Pilar López-Hernández, M., Melo-Martinez, S. E., & Criollo-Núñez, J. (2023). y en las que resulta decisivo el control de la humedad.

Dado que el grano es higroscópico, debe transportarse lo antes posible tras la cosecha, pues un almacenamiento prolongado ocasiona pérdidas por la alta humedad relativa de las regiones tropicales; asimismo, ha de evitarse la absorción accidental de humedad para prevenir el moho y el deterioro. Con temperaturas de procesamiento recomendadas de 15 a 18 °C en el secado Patricia, F. R. C., (2017), esta sensibilidad fisicoquímica convierte al transporte en uno de los eslabones más vulnerables de la cadena.

Pese a su importancia, la cadena enfrenta obstáculos que se extienden del productor al comprador. La comunicación fragmentada entre actores afecta la distribución equitativa de los beneficios y la eficiencia de la cadena (García-Cáceres et al., 2014).

A ello se suman una infraestructura vial deficiente en las zonas rurales de cultivo, que genera retrasos no contemplados, y riesgos en el comercio internacional derivados de la congestión portuaria y de las interrupciones en corredores estratégicos como los canales de Panamá y Suez, cuyo bloqueo en 2021 evidenció la vulnerabilidad global de las cadenas de suministro (UNCTAD, 2024).

La limitada adopción de tecnologías para la gestión de la información agrava el panorama: en muchas zonas el seguimiento sigue realizándose de forma manual, lo que impide consolidar datos en tiempo real y responder oportunamente ante contingencias (FAO, 2023).

Los efectos del cambio climático sobre los rendimientos y la planeación logística añaden incertidumbre (ICCO, 2024). Paradójicamente, mientras las exportaciones crecen y las importaciones de grano descienden de 6.688 toneladas en 2014 a 50 en 2025 (Fedecacao, 2026), dicho crecimiento no ha estado acompañado de un fortalecimiento proporcional de los factores logísticos.

En este escenario, la competitividad y la resiliencia se consolidan como factores decisivos. La gestión logística deja de ser una función centrada en el transporte y el almacenamiento para convertirse en un elemento estratégico capaz de generar valor (Christopher, 2022); desde la perspectiva de Porter, la competitividad depende de la capacidad de innovar e incorporar tecnología más que de las ventajas comparativas tradicionales (Esguerra, 1990).

A su vez, la resiliencia —entendida como la capacidad de anticipar, responder y recuperarse de eventos disruptivos (Christopher y Peck, 2004)— resulta indispensable en una cadena expuesta a factores climáticos, viales y portuarios. Ambas capacidades se apoyan cada vez más en herramientas tecnológicas que incrementan la visibilidad de las operaciones y facilitan la gestión del riesgo.

No obstante, la cadena de suministro del cacao santandereano no ha logrado maximizar su diseño e impacto, pues persisten sin una solución integral los obstáculos de comunicación, respuesta oportuna ante contingencias logísticas y trazabilidad del producto.

Frente a este vacío, el presente artículo tiene como objetivo diseñar una estrategia

de innovación digital, materializada en una plataforma de tracking logístico bajo el modelo de software como servicio, que fortalezca la competitividad y la resiliencia de la cadena de suministro del cacao en Santander, Colombia.

Se plantea como hipótesis que la adopción de una plataforma de tracking accesible y basada en la nube permite a las PYMES logísticas podría fortalecer dicha competitividad y resiliencia sin requerir grandes inversiones de capital fijo, al resolver la asimetría informativa y la exclusión financiera que afectan al eslabón intermedio de la cadena.

2. METODOLOGÍA

La investigación adoptó un enfoque cualitativo de carácter documental y propositivo. A partir de la revisión de literatura científica reciente, de normas técnicas y de informes sectoriales, se diseñó de forma conceptual una plataforma tecnológica de tracking logístico bajo el modelo de software como servicio.

El corpus documental revisado comprendió, además de la literatura académica, el marco institucional de la cadena —el Acuerdo Nacional de Competitividad (2001), el Consejo Nacional del Cacao (2002) y los pilares del documento de competitividad cacao-chocolate al horizonte 2030: origen en Colombia, productividad sostenible y renovable, investigación e innovación conjunta, extensión rural de impacto y mayor capacidad de transformación (SIOC-MinAgricultura, s.f.)— así como las normas técnicas que parametrizaron el diseño.

A modo de caso ilustrativo, la propuesta se proyectó sobre la empresa Ingelox SAS, que administra fletes terrestres de cacao en Santander, con el fin de mostrar su aplicación sobre los corredores logísticos de exportación del departamento.

El diseño se fundamentó en cuatro perspectivas teóricas complementarias. La Teoría de las Capacidades Dinámicas (Teece, 2007) explica la transformación digital de la PYME logística mediante las capacidades de percibir (sensing), capturar (seizing) y transformar (transforming) su operación, y muestra que dicha transformación depende más de la agilidad gerencial en la administración de los flujos de información que de la magnitud del gasto de capital fijo (Teece et al., 2016).

La Teoría de la Difusión de Innovaciones (Rogers, 2002), junto con el concepto de asimetría de la información (Herrera y Orjuela-Castro, 2021), permite comprender las barreras de apropiación tecnológica que se derivan de sistemas de trazabilidad aislados y operados de forma manual. Finalmente, el enfoque de Creación de Valor Compartido (Borda et al., 2021) y de negocios inclusivos en la base de la pirámide (Maestre Matos et al., 2023) orienta el diseño hacia beneficios distribuidos en todo el ecosistema cacaotero, y no exclusivamente en el operador logístico.

La viabilidad del diseño se sustentó en habilitadores accesibles para empresas de baja madurez digital. Frente a las brechas de conectividad y los altos costos de la Logística 4.0 en economías emergentes (Mantilla et al., 2025), y a la exclusión crediticia estructural que afecta al eslabón intermedio (Ambler et al., 2023), las arquitecturas basadas en la nube y el modelo de software como servicio eliminan los costos fijos de servidores y licencias y extienden el acceso a datos analíticos (Orjuela-Garzón et al., 2024); en efecto, las tecnologías digitales logísticas permiten reducir entre un 5 % y un 20 % los costos operativos (Procolombia). La disponi-

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 57-67 Universidad Libre - Facultad de Ingeniería

bilidad de datos en tiempo real mitiga, además, la propagación de perturbaciones a lo largo de la cadena, o efecto dominó (Ivanov, Dolgui y Sokolov, 2019).

Sobre esta base, el tracking se concibió no como la mera localización satelital de un vehículo, sino como un sistema integrado de información gerencial que enlaza el movimiento físico de la carga con su documentación de soporte y sus indicadores de calidad (Herrera y Orjuela-Castro, 2021; Guevara et al., 2022). Los parámetros del diseño se derivaron de la Norma Técnica Colombiana NTC 5811, que regula la gestión de información del Theobroma cacao, y de la estructura de Análisis de Peligros y Puntos Críticos de Control (HACCP) junto con el Índice de Cumplimiento Global (GCI) validados para centros de beneficio de cacao (Quintero-García et al., 2025).

3. RESULTADOS

El diseño resultante es una plataforma tecnológica independiente y escalable, concebida como un servicio integral de Logística 4.0 bajo el modelo de software como servicio, que ofrece tracking, trazabilidad y gestión aduanera a los distintos actores del ecosistema agroexportador. Su arquitectura articula seis servicios complementarios que transforman la visibilidad física de la carga en un instrumento de gestión, auditoría y creación de valor.

Seguimiento temporal. Un motor de seguimiento paramétrico segmenta la ruta terrestre y portuaria mediante hitos de control estandarizados, en lugar de depender de la geolocalización satelital pasiva, que exige inversiones elevadas en hardware (Mantilla et al., 2025). A partir de un código único por embarque, el sistema registra las variables de tiempo críticas —Tiempo Estimado de Arribo (ETA) frente a Tiempo Real de Arribo (ATA)— en cada punto de la cadena: finca, centro de acopio, tránsito por carretera e inspección aduanera.

Ello permite identificar analíticamente los cuellos de botella, como la congestión en el ingreso a terminales marítimos o las demoras de consolidación, y sustituye la verificación por correos y llamadas telefónicas por un panel de control en tiempo real que reduce la asimetría informativa (Herrera y Orjuela-Castro, 2021).

Gestión documental y aduanas. La plataforma centraliza la gestión documental: asocia a cada código de embarque una carpeta digital inmutable en la nube donde se cargan y validan el Bill of Lading, la factura comercial, el Packing List y los certificados fitosanitarios.

Su innovación de proceso es una compuerta lógica de validación aduanera que impide que el contenedor avance a la fase de tránsito a puerto sin la aprobación previa de la factura comercial, verificando desde el origen que las subpartidas arancelarias, los valores y los volúmenes coincidan con el manifiesto físico. De este modo se habilita una auditoría aduanera a distancia que acelera los tiempos de levante y mitiga los sobrecostos por almacenamiento no planificado o pérdida de itinerarios navieros.

Auditoría normativa. El sistema audita de forma automatizada el cumplimiento normativo sobre los flujos de datos operativos. Por una parte, parametriza las especificaciones de la NTC 5811 para el grano de cacao, que al finalizar el beneficio exige tres parámetros: un nivel de humedad técnica controlado, el control de las curvas térmicas de fermentación y el registro de las cinco fases de buenas prácticas —presiembra, siembra, mantenimiento, producción y beneficio—. En lugar de operar como barrera,

la plataforma guía al operador del centro de acopio, mediante interfaces móviles, para registrar y validar esas curvas y los porcentajes de humedad.

Por otra parte, digitaliza el Índice de Cumplimiento Global (GCI) y la estructura HACCP —en sus cuatro ejes: requisitos sanitarios estructurales y locativos, aseguramiento de la calidad, operaciones en proceso y el sistema HACCP propiamente dicho— validados para centros de beneficio de cacao (Quintero-García et al., 2025).

Cuando el tránsito por las carreteras nacionales produce una desviación térmica que amenaza la inocuidad, el sistema bloquea el indicador de conformidad sanitaria y exige el registro de medidas correctivas.

Certificación de origen. Este servicio convierte un flete genérico en una experiencia de valor. Mediante la lectura de códigos QR en las zonas de recolección (Paz-Orozco et al., 2025), captura de forma descentralizada la identidad del productor, la finca, el municipio de cultivo y la fecha de cosecha, y garantiza la segregación del lote, de modo que el grano llega al cliente con su perfil botánico e historia comunitaria intactos.

Esta información respalda la autenticidad del producto: aunque el sabor sigue siendo el principal determinante de compra, los atributos extrínsecos verificables sobre la procedencia incrementan la confianza y la disposición a pagar un precio superior (Ballesteros et al., 2023), que en el caso del cacao especial comercializado con chocolateros artesanales puede alcanzar cerca de un 20 % por encima del precio. La narrativa de origen construida con estos datos fortalece, además, la percepción de exclusividad y calidad (Escobar et al., 2021).

Alerta temprana. Un motor de alertas tempranas monitorea la operación y notifica cualquier desviación que amenace la integridad de la carga: la ausencia de documentos obligatorios antes del ingreso portuario; los retrasos con riesgo térmico, pues la inercia térmica y las fluctuaciones ambientales derivadas de detenciones prolongadas alteran el pH y provocan fermentaciones secundarias que degradan el perfil aromático del grano (Santander et al., 2025); los cambios de itinerario naviero, o roll-overs; y los puntos ciegos que se producen cuando los operadores exceden los tiempos máximos sin registrar novedades.

Ante estas alertas es posible desviar la ruta o priorizar el descargue para evitar la destrucción de los atributos del cacao.

Índice de Resiliencia Logística. La plataforma incorpora un Índice de Resiliencia Logística (IRL), un indicador que sintetiza el comportamiento de cada embarque a partir de cinco dimensiones: el nivel de trazabilidad alcanzado, la integridad y disponibilidad de la documentación, la frecuencia de actualización de eventos, el número de incidencias y el tiempo de respuesta ante ellas.

Estas variables se consolidan en una calificación de 0 a 100 puntos que permite comparar operaciones, identificar tendencias y clasificar su desempeño por niveles de resiliencia. Más allá de la calificación, el sistema genera recomendaciones automáticas asociadas a las variables de menor desempeño, con lo que se convierte en una herramienta de gestión y no solo de control.

Así democratiza el acceso a la analítica operacional propia de la Logística 4.0, habitualmente reservada a las grandes organizaciones, y permite a las PYMES evaluar y fortalecer sus procesos.

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 57-67 Universidad Libre - Facultad de Ingeniería

Tabla 1

Servicios de la plataforma propuesta y su fundamento

Servicio	Función principal	Base normativa / fuente
Seguimiento temporal	Registro de hitos ETA vs. ATA (finca, acopio, carretera, aduana) e identificación de cuellos de botella	Mantilla et al. (2025); Herrera y Orjuela-Castro (2021)
Gestión documental y aduanas	Carpeta digital inmutable (BL, factura, Packing List, fitosanitarios) con compuerta lógica de validación aduanera	—
Auditoría normativa	Verificación de NTC 5811 (3 parámetros, 5 fases) y de GCI/HACCP (4 ejes)	NTC 5811; Quintero-García et al. (2025)
Certificación de origen	Captura descentralizada por QR de la identidad del lote y segregación del grano	Paz-Orozco et al. (2025); Ballesteros et al. (2023); Escobar et al. (2021)
Alerta temprana	Notificación de documento faltante, riesgo térmico, roll-overs y puntos ciegos	Santander et al. (2025)
Índice de Resiliencia Logística	Calificación de 0 a 100 sobre cinco dimensiones y recomendaciones automáticas	Ivanov y Dolgui (2020)

Nota. Elaboración propia a partir del diseño de la plataforma.

4. DISCUSIÓN

Los resultados permiten sostener la hipótesis planteada: el diseño muestra que una plataforma de tracking accesible y basada en la nube puede fortalecer la competitividad y la resiliencia de la cadena exportadora del cacao sin exigir grandes inversiones de capital fijo. La contribución no reside en la movilización física de la carga, sino en la capacidad gerencial de resolver la asimetría informativa y de superar la exclusión financiera que afecta al eslabón intermedio (Ambler et al., 2023).

En lugar de competir por volumen o por minimización de fletes, la plataforma habilita una estrategia de diferenciación basada en la calidad y el origen (Yana et al., 2026), coherente con la exigencia de rigor científico en la cadena de custodia: la identidad botánica del grano —cuya divergencia entre variedades se ha documentado mediante 21 metabolitos clave por cromatografía de gases (Murillo et al., 2026)— solo puede certificarse si la trazabilidad se mantiene íntegra a lo largo del transporte.

Esta visibilidad en tiempo real trasciende la mera localización de la carga. Al registrar de forma abierta y verificable las prácticas agrícolas y las etapas logísticas, genera una transparencia antes inalcanzable que permite al comprador evaluar la calidad del grano a distancia (Steinke et al., 2024). En un contexto de baja madurez digital, la captura descentralizada mediante códigos QR demuestra que es viable estructurar trazabilidad eficiente para pequeños productores (Paz-Orozco et al., 2025), de modo que la transparencia deja de ser un valor agregado opcional para convertirse en condición de acceso a los mercados internacionales.

La plataforma desplaza además el control de calidad desde un enfoque reactivo —inspecciones tardías o muestreos en puerto— hacia una gestión preventiva del riesgo. La digitalización del GCI y del HACCP permite detectar las desviaciones mientras la operación está en curso (Quintero-García et al., 2025), y el monitoreo de los tiempos de tránsito frente a la sensibilidad fisicoquímica del grano (Santander et al., 2025) habilita respuestas anticipadas que preservan el perfil aromático. Esta capacidad de segregación y monitoreo continuo construye resiliencia frente al efecto dominó de las disrupciones viales (Ivanov, Dolgui y Sokolov, 2019), al evitar que un contratiempo se propague a lo largo de la cadena.

Más allá de lo técnico, la trazabilidad digital opera como catalizador de inclusión social. La estandarización y digitalización de los eslabones de la cadena de valor constituyen un mecanismo para dotar de transparencia los procesos poscosecha (Escobar et al., 2020), y su articulación con redes asociativas permite registrar en el origen el capital social y la equidad de género —como el liderazgo de las mujeres rurales en los nodos de producción cacaotera de Sardinata y Tibú (Albornoz-Arias et al., 2025a, 2025b)—, en línea con los factores institucionales que impulsan los negocios inclusivos en la base de la pirámide (Maestre Matos et al., 2023). De este modo, la innovación logística deja de ser un beneficio exclusivo del operador de transporte para convertirse en valor compartido por todo el ecosistema cacaotero (Borda et al., 2021).

El presente trabajo corresponde a una fase de diseño conceptual cuyos componentes se apoyan en evidencia empírica previa: cada servicio de la plataforma se sustenta en resultados ya validados por la literatura, desde la digitalización del HACCP en centros de beneficio (Quintero-García et al., 2025) hasta la trazabilidad por códigos QR en pequeños productores (Paz-Orozco et al., 2025). Su alcance se circunscribe, por ahora, a la proyección sobre un caso ilustrativo, de modo que la cuantificación de sus efectos sobre costos, tiempos y niveles de servicio constituye el paso natural siguiente. En esa dirección, la implementación de un piloto permitiría validar el Índice de Resiliencia Logística con información operativa real y evaluar la transferibilidad de la plataforma a otras cadenas agroexportadoras y regiones del país, consolidando una propuesta que ya cuenta con un fundamento teórico y normativo sólido.

5. CONCLUSIÓN

La gestión de la innovación aplicada al tracking digital convierte la logística de transporte en un motor de creación de valor compartido para el ecosistema cacaotero santandereano. Al reducir la opacidad informativa del eslabón intermedio y armonizar los requisitos de la Norma Técnica Colombiana NTC 5811 (Guevara et al., 2022), la PYME logística se afirma como un aliado estratégico de la agroexportación, y la visibilidad tecnológica deja de ser un accesorio operativo para convertirse en el pilar que faculta a la agroindustria colombiana a acceder a cadenas de valor globales más exigentes y a sostener su participación con rentabilidad, gobernanza ética y seguridad científica. En síntesis, la competitividad agroexportadora contemporánea no depende únicamente de movilizar la carga, sino de la capacidad de resolver la asimetría informativa y la exclusión financiera mediante una digitalización accesible; la plataforma propuesta ofrece un vehículo operativo para lograrlo, cuya validación empírica en campo constituye la próxima etapa de esta línea de investigación.

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 57-67 Universidad Libre - Facultad de Ingeniería

REFERENCIAS BIBLIOGRÁFICAS

[1] Albornoz-Arias, N., Rojas-Sanguino, C., & Santafe-Rojas, A.-K. (2025a). Empowering rural women in the cocoa production chain in Sardinata, Norte de Santander, Colombia. Social Sciences, 14(2), 94. https://doi.org/10.3390/socsci14020094

[2] Albornoz-Arias, N., Rojas-Sanguino, C., & Santafe-Rojas, A.-K. (2025b). Rural women's leadership within the cocoa production chain in Tibú, Norte de Santander, Colombia: A gender perspective. Agriculture, 15(2), 370. https://doi.org/10.3390/agriculture15020370 (DOI a verificar: en el original figura malformado)

[3] Ambler, K., de Brauw, A., Herskowitz, S., & Pulido, C. (2023). Viewpoint: Finance needs of the agricultural midstream. Food Policy. https://doi.org/10.1016/j.foodpol.2023.102530

[4] Armuelles, K., & A, V. E. P. (2023). Gestión de la cadena de suministro del cacao: Estado del Arte. Investigación y Pensamiento Crítico, 11(3), 12-20. https://doi.org/10.37387/ipc.v11i3.364

[5] Borda, A., Morales, O., Teegen, H., Rees, G. H., & Gonzalez-Perez, M. A. (2021). Addressing sustainable rural development with shared value: A Peruvian model from the cacao industry. Sustainability, 13(14), 8028. https://doi.org/10.3390/su13148028

[6] Christopher, M. (2022). Logistics and supply chain management (4.ª ed.). https://rudyct.com/supchn/Christopher%20Logistics%20and%20Supply%20Chain%20Management%204th%20txtbk.pdf

[7] Christopher, M., & Peck, H. (2004). Building the resilient supply chain. The International Journal of Logistics Management. (volumen y páginas incompletos en el original; verificar)

[8] Del Pilar López-Hernández, M., Melo-Martinez, S. E., & Criollo-Núñez, J. (2023). Effect of the maturity stage, genotype, and geographical location on the physicochemical characteristics of the cocoa bean during fermentation. Ingeniería y Competitividad, 25(3). https://doi.org/10.25100/iyc.v25i2.12503

[9] Escobar, S., Santander, M., Useche, P., Contreras, C., & Rodríguez, J. (2020). Aligning strategic objectives with research and development activities in a soft commodity sector: A technological plan for Colombian cocoa producers. Agriculture, 10(5), 141. https://doi.org/10.3390/agriculture10050141 (el original indica «143» en el texto; número de artículo a verificar)

[10] Escobar, S., Santander, M., Zuluaga, M., Chacón, I., Rodríguez, J., & Vaillant, F. (2021). Fine cocoa beans production: Tracking aroma precursors through a comprehensive analysis of flavor attributes formation. Food Chemistry, 365, 130627. https://doi.org/10.1016/j.foodchem.2021.130627

[11] Esguerra, M. del P. (1990). Nuevas hipótesis sobre los patrones del comercio internacional: "Las ventajas competitivas de las naciones". Banco de la República de Colombia.

[12] Ivanov, D., & Dolgui, A. (2020). Viability of intertwined supply networks: extending the supply chain resilience angles towards survivability. A position paper motivated by COVID-19 outbreak. International Journal Of Production Research, 58(10), 2904-2915. https://doi.org/10.1080/00207543.2020.1750727

[13] FAO Colombia. (2022, 22 de agosto). 54 familias de cacaoteros en Norte de Santander implementan un Sistema Agroforestal innovador. https://www.fao.org/colombia/noticias/detail-events/ru/c/1602889/

[14] Federación Nacional de Cacaoteros. (2026, 4 de mayo). Cacao colombiano conquista mercados: exportaciones alcanzan cifras históricas en 2025, con 36.097 toneladas. Fedecacao. https://www.fedecacao.com.co/post/cacao-colombiano-conquista-mercados

[15] Food and Agriculture Organization of the United Nations. (2023). The Status of Digital Agriculture in Latin America and the Caribbean. FAO.

[16] García-Cáceres, R. G., Perdomo, A., Ortiz, O., Beltrán, P., & López, K. (2014). Caracterización de las cadenas de abastecimiento y valor del cacao en Colombia. DYNA, 81(187), 30–40. https://revistas.unal.edu.co/index.php/dyna/article/view/39555

[17] Guevara, E., Rojas, A. E., & Florez, H. (2022). Technology platform for the information management of Theobroma cacao crops based on the Colombian Technical Standard 5811. Engineering Letters, 30(1), 36–50.

[18] Herrera, M. M., & Orjuela-Castro, J. (2021). An appraisal of traceability systems for food supply chains in Colombia. International Journal on Food System Dynamics, 12(1), 37–50. https://doi.org/10.18461/ijfsd.v12i1.74

[19] Ivanov, D., Dolgui, A., & Sokolov, B. (2019). The impact of digital technology and Industry 4.0 on the ripple effect and supply chain risk analytics. International Journal of Production Research, 57(3), 829–846. https://doi.org/10.1080/00207543.2018.1488086

[20] International Cocoa Organization. (2024). ICCO Annual Report 2023/2024. https://www.icco.org/wp-content/uploads/ICCO-Annual-Report-2023-2024-EN.pdf

[21] Maestre-Matos, L. M. M., Lombana-Coy, J. L., Mesías, F. J. M., & Cabas, A. P. P. (2023). The role of institutional factors in the development and consolidation of inclusive businesses with small farmers at the base of the pyramid. REVESCO. Revista de Estudios Cooperativos, 145, 1–23. https://doi.org/10.5209/REVE.91312

[22] Mantilla, F., Mejía, G., & Tascón, D. (2025). The role of Industry 4.0 technologies in the export flower industry: Insights from a systematic literature review and surveys in emerging economies. Results in Engineering, 25, 104507. https://doi.org/10.1016/j.rineng.2025.104507

[23] Murillo, J. S., Tirado, M., Castillo, E., & Sinuco, D. C. (2026). Differentiation of Criollo and Trinitario cocoa beans: Study case Santander farms, Colombia. ACS Food Science & Technology, 6(2), 368–378. https://doi.org/10.1021/acsfoodscitech.5c00854

[24] Orjuela-Garzón, W. A., Sandoval-Aldana, A., & Méndez-Arteaga, J. J. (2024). Systematic literature review of barriers and enablers to implementing food informatics technologies: Unlocking agri-food chain innovation. Foods, 13(21), 3349. https://doi.org/10.3390/foods13213349

[25] Patricia, F. R. C. (2017, 15 marzo). Desarrollo de una nueva tecnología de procesamiento de post-cosecha de cacaos finos y de aroma en centros de acopio para mercados internacionales. https://repositorio.minciencias.gov.co/entities/publication/cfb65b08-d479-4a64-b199-f55f582ed0be?

[26] Paz-Orozco, H., Antunes de Souza Lima, I., Anacona-Mopan, Y., de Brito Junior, I., Luna, A., & Chong, M. (2025). Traceability in the supply chain of specialty coffee small producers through QR technology in Colombia. Brazilian Journal of Biosystems Engineering, 19, 1286. https://doi.org/10.18011/bioeng.2025.v19.1286

[27] Quintero-García, J. C., Castro-Camacho, J. K., Rodríguez-Polanco, L., & Criollo-Núñez, J. (2025). Development and validation of an inspection and compliance tool for sanitary requirements applied to cacao processing centers in the department of Huila. Vitae, 32(2), 358408. https://doi.org/10.17533/udea.vitae.v32n2a358408

[28] Redacción Economía. (2023, 10 de febrero). Este es el nuevo impulso para los cacaoteros del Magdalena Medio, ¿cómo lo hicieron? Vanguardia. https://www.vanguardia.com/economia/local/2023/02/10/este-es-el-nuevo-impulso-para-los-cacaoteros-del-magdalena-medio-como-lo-hicieron/

[29] Rogers, E. M. (2002). Diffusion of preventive innovations. Addictive Behaviors, 27(6), 989–993. https://doi.org/10.1016/S0306-4603(02)00300-3

[30] Santander, M., Leguizamón, L., Vaillant, F., Boulanger, R., Zuluaga, M., Maraval, I., Rodríguez, J., Llano, S., Sommerer, N., Meudec, E., Meter, A., & Escobar, S. (2025). Influence of driven fermentation of cacao in bioreactors on quality: Decoding the effect of temperature, mixing, and pH on metabolomic, sensory, and volatile profiles. LWT – Food Science and Technology, 231, 118313. https://doi.org/10.1016/j.lwt.2025.118313

[31] SIOC-MinAgricultura. (s.f.). Sistema de Información de Gestión y Desempeño de Organizaciones de Cadena. https://sioc.minagricultura.gov.co/Cacao/Pages/default.aspx

[32] Steinke, J., Ivanova, Y., Jones, S. K., Minh, T., Sánchez, A., Sánchez-Choy, J., & Mockshell, J. (2024). Digital sustainability tracing in smallholder context: Ex-ante insights from the Peruvian cocoa supply chain. World Development Sustainability, 5, 100185. https://doi.org/10.1016/j.wds.2024.100185

[33] Teece, D. J. (2007). Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. Strategic Management Journal, 28(13), 1319–1350. https://doi.org/10.1002/smj.640

[34] Teece, D., Peteraf, M., & Leih, S. (2016). Dynamic capabilities and organizational agility: Risk, uncertainty, and strategy in the innovation economy. California Management Review, 58(4), 13–35. https://doi.org/10.1525/cmr.2016.58.4.13

[35] UNCTAD. (2024). Review of maritime transport 2024: Navigating maritime chokepoints. United Nations Conference on Trade and Development. https://unctad.org/publication/review-maritime-transport-2024

[36] Yana, S., Maryam, Hafasnuddin, Rusmina, C., & Mukhdasir. (2026). A review of local commodity development and export strategies for coffee, cocoa, and major agricultural products. World Development Sustainability, 8, 100295. https://doi.org/10.1016/j.wds.2026.100295

SISTEMA DE IDENTIDAD DIGITAL POST-CUÁNTICA PARA EL ACCESO SEGURO A MODELOS DE INTELIGENCIA ARTIFICIAL

Carlos Jose Petro Almario
Universidad Libre, Bogotá, Colombia
Correo: Carlosj-petroa@unilibre.edu.co

RESUMEN

El crecimiento acelerado de la Inteligencia Artificial (IA) ha generado nuevos desafíos en ciberseguridad, especialmente en la autenticación de usuarios y el control de acceso a sistemas inteligentes. Los mecanismos tradicionales de autenticación dependen de algoritmos criptográficos clásicos, como RSA y ECC, que podrían volverse vulnerables frente al avance de la computación cuántica. Esta situación representa un riesgo crítico para la protección de identidades digitales y la seguridad de modelos de Inteligencia Artificial.

El presente artículo propone un sistema de identidad digital post-cuántica orientado al acceso seguro a modelos de IA mediante la integración de criptografía post-cuántica, identidad descentralizada (DID) y tecnología blockchain. La propuesta utiliza el algoritmo CRYSTALS-Dilithium, seleccionado por el National Institute of Standards and Technology (NIST) como estándar de firmas digitales resistentes a ataques cuánticos.

La arquitectura diseñada busca garantizar autenticación segura, integridad de la información, trazabilidad de accesos y control soberano de la identidad digital por parte del usuario. Asimismo, la implementación de blockchain permite almacenar registros de acceso de forma distribuida e inmutable, reduciendo riesgos asociados a manipulación y centralización.

La investigación se desarrolló bajo un enfoque descriptivo y aplicado, mediante revisión bibliográfica, análisis comparativo de tecnologías actuales y diseño conceptual del sistema. Aunque el proyecto se encuentra en fase de desarrollo conceptual, se definieron escenarios de validación orientados a evaluar resistencia frente a ataques clásicos y cuánticos, tiempos de autenticación y detección de accesos no autorizados. Los resultados obtenidos evidencian que la integración de tecnologías emergentes como DID, blockchain y criptografía post-cuántica representa una alternativa viable para fortalecer la seguridad digital en entornos basados en Inteligencia Artificial y preparar infraestructuras tecnológicas frente a futuros escenarios de computación cuántica.

Palabras clave: Criptografía Post-Cuántica, Identidad Digital, Blockchain, Inteligencia Artificial, CRYSTALS-Dilithium, DID.

Abstract

The rapid growth of Artificial Intelligence (AI) has created new cybersecurity challenges, particularly in user authentication and access control mechanisms. Traditional authentication systems rely on classical cryptographic algorithms such as RSA and ECC, which may become vulnerable due to advances in quantum computing. This situation represents a critical risk for digital identity protection and AI system security.

This article proposes a post-quantum digital identity system designed to ensure secure access to Artificial Intelligence models through the integration of post-quantum cryptography, decentralized identity (DID), and blockchain technology. The proposal uses the CRYSTALS-Dilithium algorithm, selected by the National Institute of Standards and Technology (NIST) as a standard for quantum-resistant digital signatures.

The proposed architecture aims to guarantee secure authentication, data integrity, access traceability, and self-sovereign identity management. Additionally, blockchain technology enables immutable and distributed storage of access records, reducing risks associated with manipulation and centralization. The research follows a descriptive and applied methodology based on literature review, comparative analysis, and conceptual system design.

Although the project remains at a conceptual stage, several validation scenarios were defined to evaluate resistance against classical and quantum attacks, authentication performance, and unauthorized access detection. The results indicate that integrating emerging technologies such as DID, blockchain, and post-quantum cryptography provides a viable alternative for strengthening cybersecurity in Artificial Intelligence environments and preparing digital infrastructures for future quantum computing scenarios.

Keywords: Post-Quantum Cryptography, Digital Identity, Blockchain, Artificial Intelligence, CRYSTALS-Dilithium, DID.

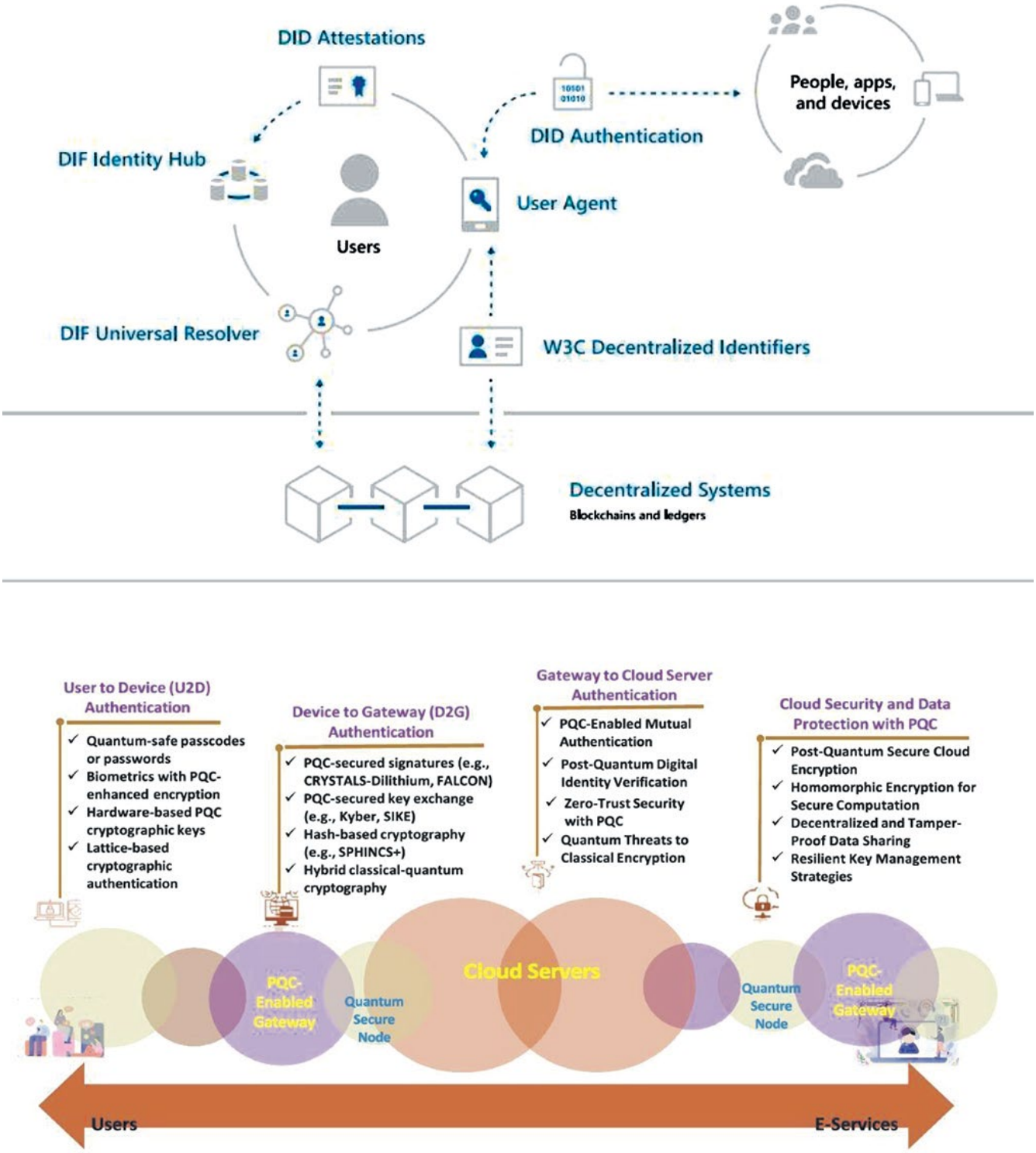
1. INTRODUCCIÓN

La transformación digital y el crecimiento exponencial de la Inteligencia Artificial han permitido el desarrollo de sistemas capaces de automatizar procesos, analizar grandes volúmenes de información y optimizar la toma de decisiones en múltiples sectores. Sin embargo, este avance tecnológico también ha incrementado los riesgos relacionados con la seguridad informática y la protección de identidades digitales.

Actualmente, gran parte de los sistemas de autenticación dependen de mecanismos centralizados y algoritmos criptográficos clásicos como RSA y ECC. Aunque estos algoritmos han sido considerados seguros durante décadas, investigaciones relacionadas con computación cuántica demuestran que podrían ser vulnerables frente a algoritmos cuánticos como el algoritmo de Shor.

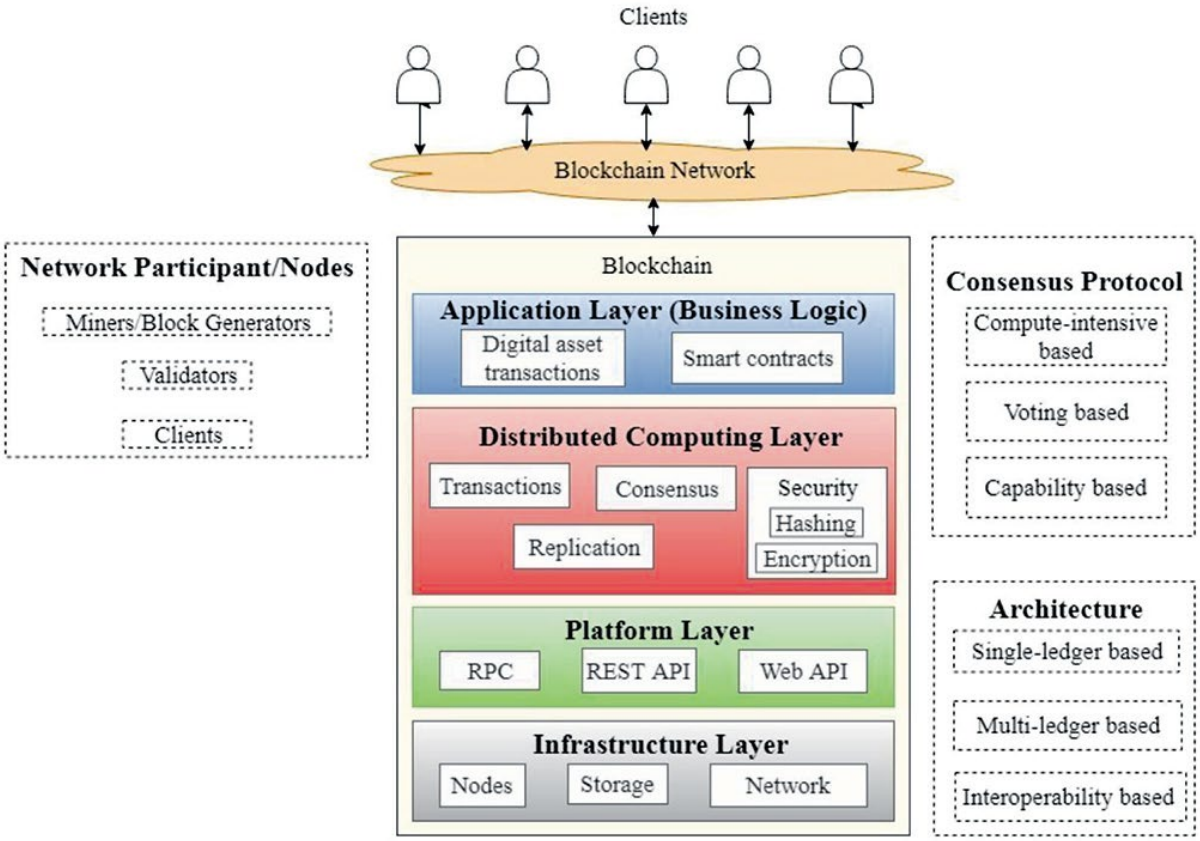
Figura 1.

Arquitectura General del Sistema de Identidad Digital PostCuántica



Fuente: Adaptado de Microsoft Entra – Decentralized Identity Architecture

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 68-78 Universidad Libre - Facultad de Ingeniería



Además, la dependencia de entidades centralizadas genera problemas asociados a privacidad, manipulación de información, fuga de datos y control limitado de la identidad digital por parte del usuario.

En este contexto, la criptografía post-cuántica surge como una alternativa orientada a desarrollar mecanismos criptográficos resistentes frente a ataques cuánticos. Paralelamente, tecnologías como blockchain e identidad descentralizada (DID) permiten construir sistemas más transparentes, distribuidos y seguros.

El presente trabajo propone un sistema de identidad digital post-cuántica para el acceso seguro a modelos de Inteligencia Artificial, integrando criptografía post-cuántica mediante CRYSTALS-Dilithium, blockchain y DID con el objetivo de fortalecer la autenticación digital y garantizar trazabilidad e integridad de los accesos.

2. MATERIALES Y MÉTODOS

Diseño Metodológico

La investigación se desarrolló bajo un enfoque descriptivo y aplicado orientado al análisis de tecnologías emergentes relacionadas con seguridad digital e identidad descentralizada.

Para el desarrollo del estudio se realizó revisión bibliográfica de estándares internacionales, artículos científicos y documentación técnica sobre blockchain, identidad digital descentralizada y criptografía post-cuántica.

Asimismo, se diseñó una arquitectura conceptual basada en integración de DID,

blockchain y firmas digitales resistentes a ataques cuánticos utilizando el algoritmo CRYSTALSDilithium.

El crecimiento acelerado de plataformas basadas en Inteligencia Artificial ha generado nuevos desafíos en seguridad digital, especialmente relacionados con autenticación de usuarios y control de acceso.

Diversos estudios internacionales indican que un alto porcentaje de vulnerabilidades en sistemas digitales están relacionadas con fallas de autenticación, credenciales comprometidas y suplantación de identidad. Asimismo, los sistemas tradicionales continúan dependiendo de algoritmos criptográficos vulnerables frente al avance de la computación cuántica.

Entre las principales problemáticas identificadas se encuentran:

- Vulnerabilidad frente a ataques de suplantación de identidad.
- Dependencia de entidades centralizadas.
- Falta de trazabilidad confiable en accesos digitales.
- Riesgo futuro asociado a computadores cuánticos.
- Limitado control del usuario sobre su identidad digital.

Estas limitaciones evidencian la necesidad de implementar nuevas arquitecturas de autenticación resistentes a amenazas futuras y alineadas con modelos de seguridad descentralizada.

3. OBJETIVOS DEL ESTUDIO

3.1 Objetivo General

Diseñar una arquitectura de identidad digital post-cuántica que garantice autenticación segura, descentralizada y resistente a ataques clásicos y cuánticos en el acceso a modelos de Inteligencia Artificial.

3.2 Objetivos Específicos

- Implementar identidad digital descentralizada mediante DID.
- Integrar firmas digitales utilizando CRYSTALS-Dilithium.
- Registrar accesos mediante blockchain.
- Diseñar un modelo de autenticación resistente a ataques cuánticos.
- Definir escenarios de validación y pruebas de seguridad.

4. RESULTADOS

Estado del Arte y Hallazgos

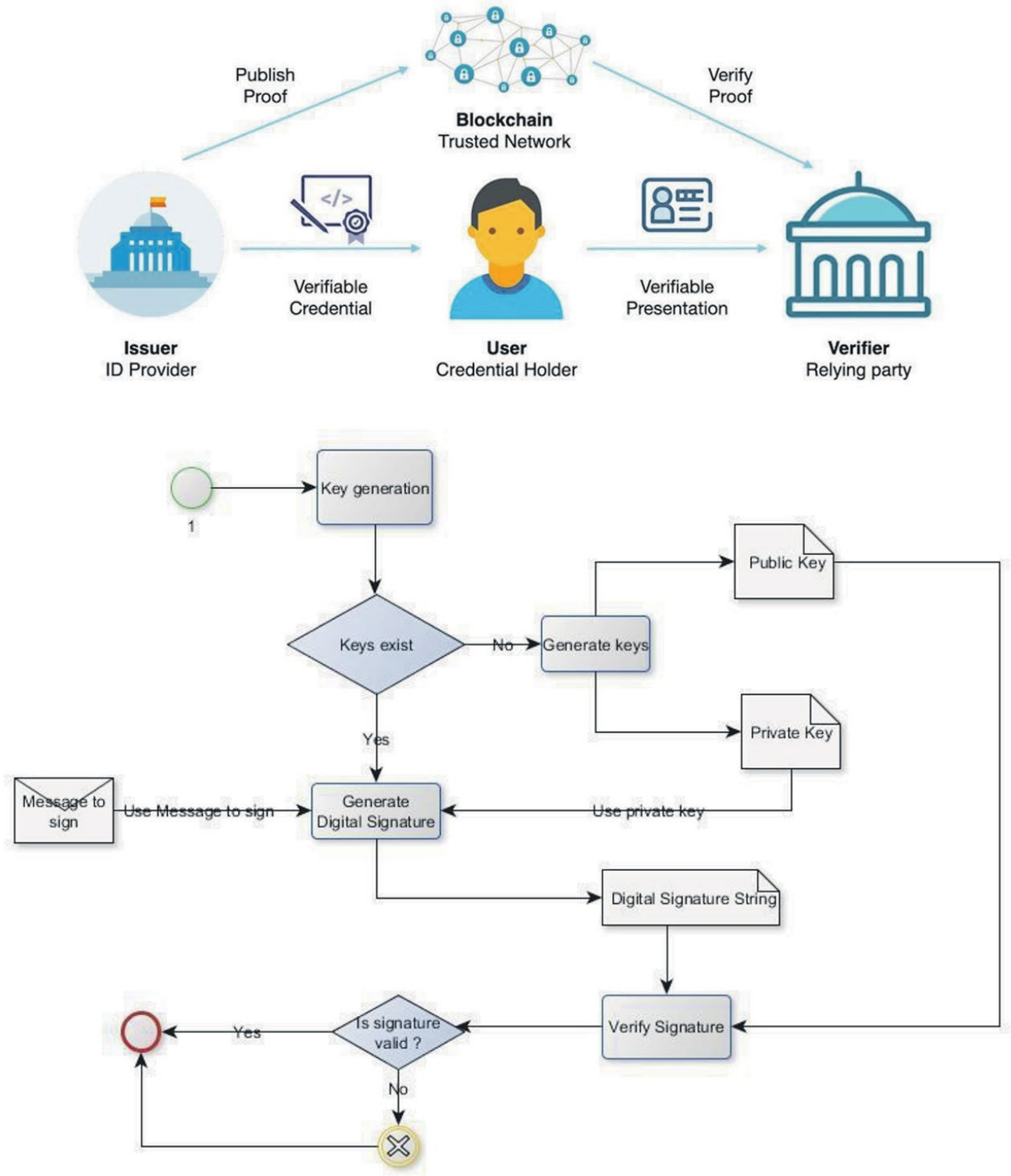
La identidad digital descentralizada ha sido impulsada por el World Wide Web Consortium (W3C) mediante el estándar DID, el cual propone sistemas de identidad autosoberana donde el usuario mantiene control total sobre sus credenciales.

Por otra parte, la criptografía post-cuántica ha cobrado relevancia debido al avance de la computación cuántica. En 2022, el National Institute of Standards and Technology (NIST) seleccionó CRYSTALS-Dilithium como estándar para firmas digitales post-cuánticas debido a su eficiencia y resistencia criptográfica.

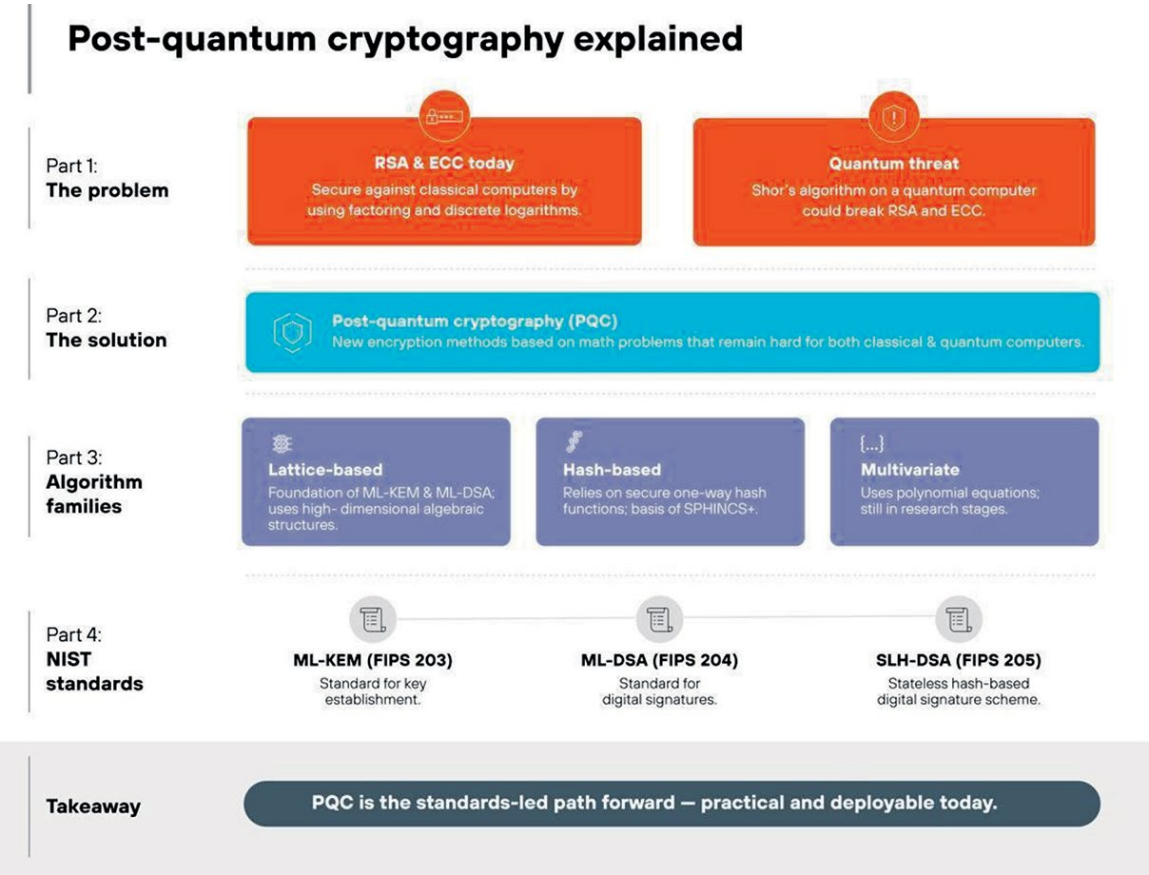
ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 68-78 Universidad Libre - Facultad de Ingeniería

Figura 2.

Flujo de Autenticación con DID y CRYSTALS-Dilithium



Fuente: Adaptado de NIST – CRYSTALS-Dilithium Digital Signature Scheme



Asimismo, blockchain se ha consolidado como una tecnología capaz de garantizar almacenamiento distribuido, inmutabilidad y trazabilidad de registros digitales. A pesar de los avances en seguridad informática, la mayoría de investigaciones en Inteligencia Artificial se enfocan principalmente en protección de modelos y ataques adversariales, dejando en segundo plano la autenticación robusta del usuario.

La propuesta desarrollada en esta investigación busca integrar estas tecnologías emergentes en una única arquitectura de seguridad orientada a entornos de IA.

5. ARQUITECTURA Y COMPONENTES DEL SISTEMA

La arquitectura propuesta fue diseñada bajo un enfoque distribuido y seguro, integrando tecnologías de identidad descentralizada, criptografía post-cuántica y blockchain para proteger el acceso a modelos de Inteligencia Artificial. El objetivo principal de esta arquitectura es garantizar autenticación robusta, trazabilidad de accesos, integridad de la información y resistencia frente a ataques clásicos y cuánticos.

El sistema se compone de varios módulos interconectados que trabajan de manera coordinada para validar la identidad del usuario antes de permitir el acceso a modelos de IA.

Módulo de Identidad Descentralizada (DID)

Este componente permite la creación y gestión de identidades digitales auto-soberanas. Cada usuario posee un identificador único descentralizado que elimina la de-

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 68-78 Universidad Libre - Facultad de Ingeniería

pendencia de entidades centrales de autenticación. Las credenciales asociadas al usuario son administradas directamente por el propietario de la identidad, aumentando privacidad, seguridad y control de datos personales.

Principales funciones:
- Generación de identidades digitales descentralizadas.
- Asociación de claves criptográficas.
- Gestión de credenciales verificables.
- Eliminación de intermediarios de autenticación.

Módulo Criptográfico Post-Cuántico

Este módulo implementa el algoritmo CRYSTALS-Dilithium para generación y validación de firmas digitales resistentes a ataques cuánticos. Cada usuario genera un par de claves criptográficas:

- Clave privada: utilizada para firmar solicitudes de acceso.
- Clave pública: utilizada para verificar autenticidad.

La utilización de criptografía post-cuántica fortalece significativamente la seguridad del sistema frente a futuros escenarios donde computadores cuánticos puedan comprometer algoritmos tradicionales como RSA y ECC.

Asimismo, las firmas digitales garantizan:
- Integridad de la información.
- Autenticidad del usuario.
- No repudio.
- Protección frente a suplantación de identidad.

Blockchain y Registro de Accesos

La blockchain funciona como una capa de almacenamiento distribuido encargada de registrar de manera inmutable cada intento de autenticación realizado dentro del sistema. Cada acceso validado genera un registro criptográfico que incluye:
- Identificador del usuario.
- Marca temporal.
- Estado de autenticación.
- Huella criptográfica de la solicitud.

La naturaleza distribuida de blockchain permite evitar manipulación de registros y proporciona transparencia en el proceso de autenticación. Además, el sistema puede implementar contratos inteligentes para automatizar validaciones de identidad y permisos de acceso.

Sistema de Autenticación

El sistema de autenticación coordina todos los módulos de seguridad. Cuando un usuario intenta acceder a un modelo de Inteligencia Artificial, se ejecuta el siguiente flujo:
1. El usuario genera una solicitud firmada digitalmente.
2. El sistema obtiene la clave pública asociada al DID.
3. Se verifica la firma digital mediante CRYSTALS-Dilithium.
4. Se valida la autenticidad del usuario.
5. El acceso es registrado en blockchain.
6. El sistema autoriza o rechaza el acceso.

Este proceso permite garantizar autenticación descentralizada y resistente a ataques clásicos y cuánticos.

Integración con Modelos de Inteligencia Artificial

La arquitectura propuesta funciona como una capa de seguridad previa al acceso a modelos de Inteligencia Artificial. Esto significa que ningún usuario puede interactuar con el modelo sin superar previamente el proceso de validación criptográfica.

Este enfoque permite:
- Controlar accesos de manera segura.
- Registrar interacciones realizadas.

- Reducir riesgos de uso malicioso.
- Garantizar trazabilidad de operaciones.

Además, la propuesta puede adaptarse a diferentes entornos tecnológicos como plataformas web, sistemas empresariales, aplicaciones gubernamentales y servicios basados en IA.

6. DESARROLLO DE LA PROPUESTA

La investigación se desarrolló bajo un enfoque descriptivo-aplicado dividido en cuatro fases principales:

Fase 1. Investigación
Se realizó revisión bibliográfica sobre:
- Identidad descentralizada.
- Blockchain.
- Criptografía post-cuántica.
- Seguridad en Inteligencia Artificial.

Fase 2. Diseño
Se definió la arquitectura del sistema, el flujo de autenticación y la interacción entre blockchain, DID y firmas digitales.

Fase 3. Desarrollo Conceptual
Se planteó el diseño de módulos criptográficos, generación de claves y validación de accesos.

Fase 4. Validación
Se definieron escenarios de prueba orientados a:
- Simulación de suplantación de identidad.
- Verificación de firmas digitales.
- Registro blockchain.
- Evaluación de rendimiento.

7. INTEGRACIÓN TECNOLÓGICA

La arquitectura diseñada integra múltiples componentes tecnológicos:
- Módulo DID.
- Módulo de generación de claves post-cuánticas.
- Sistema de autenticación.
- Blockchain.
- Contratos inteligentes.
- Modelos de Inteligencia Artificial.

El sistema funciona mediante generación de pares de claves criptográficas. La clave privada permite firmar solicitudes de autenticación, mientras que la clave pública verifica la identidad del usuario.

Los registros de acceso son almacenados en blockchain para garantizar integridad, trazabilidad y transparencia.

8. RESULTADOS ESPERADOS Y VALIDACIÓN

Aunque el sistema se encuentra actualmente en fase conceptual, se definieron métricas de evaluación para futuras implementaciones:

Métrica	Objetivo
Tiempo de autenticación	Evaluar eficiencia
Detección de accesos no autorizados	Validar seguridad
Intentos bloqueados	Medir resistencia
Integridad blockchain	Garantizar trazabilidad
Consumo de recursos	Analizar rendimiento

Se espera que la integración de criptografía post-cuántica y DID reduzca significativamente riesgos asociados a autenticación tradicional.

9. DISCUSIÓN Y CONCLUSIONES

Discusión
La integración de blockchain, identidad descentralizada y criptografía post-cuántica representa una alternativa innovadora frente a modelos tradicionales de autenticación.

ISSN 2322-8415 / Año 2026 / No. 26 / Páginas 68-78 Universidad Libre - Facultad de Ingeniería

Entre las principales ventajas identificadas destacan:

- Resistencia frente a ataques cuánticos.
- Eliminación de intermediarios centrales.
- Mayor control del usuario sobre su identidad.
- Registro inmutable de accesos.
- Transparencia y trazabilidad.

Sin embargo, también existen desafíos relacionados con costos computacionales, escalabilidad blockchain y complejidad de implementación. A pesar de ello, el sistema propuesto presenta una arquitectura viable y preparada para futuros escenarios digitales basados en Inteligencia Artificial.

10. CONCLUSIONES

La investigación permitió diseñar una propuesta de sistema de identidad digital postcuántica orientada a fortalecer la seguridad en el acceso a modelos de Inteligencia Artificial.

La integración de tecnologías como DID, blockchain y CRYSTALS-Dilithium permitió establecer una arquitectura moderna, descentralizada y resistente a amenazas futuras relacionadas con la computación cuántica.

Asimismo, se identificaron vulnerabilidades críticas en los sistemas tradicionales de autenticación, especialmente relacionadas con suplantación de identidad y dependencia de algoritmos criptográficos clásicos.

Como trabajo futuro se plantea desarrollar un prototipo funcional, ejecutar pruebas experimentales y validar el sistema en escenarios reales.

REFERENCIAS BIBLIOGRÁFICAS

[1] National Institute of Standards and Technology (NIST). (2022). Post-Quantum Cryptography Standardization.

[2] NIST. (2023). CRYSTALS-Dilithium Specification.

[3] World Wide Web Consortium (W3C). (2022). Decentralized Identifiers (DID) v1.0.

[4] Nakamoto, S. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System.

[5] Wood, G. (2014). Ethereum Whitepaper.

[6] Boneh, D., & Shoup, V. (2020). A Graduate Course in Applied Cryptography.

[7] Bernstein, D. J., et al. (2017). Post-Quantum Cryptography.

[8] Shor, P. (1994). Algorithms for Quantum Computation.

[9] OWASP. (2023). Identity and Authentication Failures.

[10] IEEE. (2022). Blockchain Security Advances.

[11] Microsoft. (2023). Decentralized Identity Overview.

[12] IBM Quantum. (2023). Quantum Computing Research.

[13] ENISA. (2021). Post-Quantum Cryptography: Current State.

[14] Hyperledger Foundation. (2022). Blockchain Frameworks.

[15] ISO/IEC. (2020). Information Security Standards.