UNIVERSIDAD TÉCNICA DE MANABÍ
Fundada en 1952

INFORMÁTICA Y SISTEMAS
Revista de Tecnologías de la Informática y las Comunicaciones

Buenas Prácticas de Seguridad para la Implementación de Inteligencia Artificial en Entornos de Computación en la Nube

Best Practices for Securing the Implementation of Artificial Intelligence in Cloud-Based Environments

Autores

**** Yasbeck Jemima Mora Chávez***
ymora6@utmachala.edu.ec

Reinaldo Benedicto Fernández Gonzalez
rfernande4@utmachala.edu.ec

Joofre Antonio Honores Tapia
jhonores@utmachala.edu.ec

Milton Rafael Valarezo Pardo
mvalarezo@utmachala.edu.ec

Universidad Técnica de Machala, Facultad de Ingeniería Civil, Carrera de Tecnologías de la Información, Machala, El Oro, Ecuador.
*Autor para correspondencia

Comó citar el artículo:
Mora Chavez, Y.J., Fernández Gonzalez, R.B., Honores Tapia, J.A. & Valarezo Pardo, M.R. 2025. Buenas Prácticas de Seguridad para la Implementación de Inteligencia Artificial en Entornos de Computación en la Nube. *Informática y Sistemas. 9 (2)*, pp. 152-163. https://doi.org/10.33936/isrtic.v9i2.7687

Enviado: 03/07/2025
Aceptado: 19/08/2025
Publicado: 26/08/2025

Resumen

La incorporación de la Inteligencia Artificial (IA) en los entornos de Cloud Computing (Computación en la Nube) ha permitido alcanzar mayores niveles de escalabilidad, eficiencia y automatización en la creación de sistemas inteligentes. Sin embargo, esta confluencia también ha resaltado nuevas amenazas y desafíos críticos en aspectos de seguridad y privacidad. El presente artículo desarrolla una revisión sistemática de literatura científica publicada entre 2021 y 2025, siguiendo la metodología PRISMA 2020, para identificar buenas prácticas de seguridad aplicadas al despliegue de proyectos de IA en entornos de Computación en la Nube. Se analizaron veinticinco estudios relevantes, lo que permitió sintetizar un conjunto de recomendaciones centradas en la gestión de accesos, protección de APIs, cifrado de datos y auditoría de eventos. Se estableció un prototipo de inteligencia artificial que incorporó controles de seguridad y privacidad de datos para ser finalmente desplegado en un entorno de Computación en la Nube, buscando validar la aplicabilidad de las buenas prácticas identificadas en la revisión sistemática. Los resultados obtenidos evidencian la efectividad de los controles aplicados, así como la necesidad de adoptar estrategias proactivas de seguridad desde las fases iniciales de los proyectos de despliegue de la IA en la nube.

Palabras clave: inteligencia artificial; computación en la nube; seguridad informática; buenas prácticas; revisión sistemática.

Abstract

The incorporation of Artificial Intelligence (AI) into Cloud Computing environments has enabled greater levels of scalability, efficiency, and automation in the creation of intelligent systems. However, this convergence has also highlighted new threats and critical challenges in terms of security and privacy. The present article develops a systematic review of scientific literature published between 2021 and 2025, following the PRISMA 2020 methodology, to identify good security practices applied to the deployment of AI projects in cloud computing environments. Twenty-five relevant studies were analyzed, which allowed for the synthesis of a set of recommendations focused on access management, API protection, data encryption, and event auditing. An artificial intelligence prototype incorporating security and data privacy controls was established and finally deployed in a cloud computing environment to validate the applicability of the best practices identified in the systematic review. The results obtained demonstrate the effectiveness of the controls applied, as well as the need to adopt proactive security strategies from the initial phases of AI deployment projects in the cloud.

Keywords: artificial intelligence; cloud computing; information security; best practices; systematic review.

INFORMÁTICA Y SISTEMAS
Revista de Tecnologías de la Informática y las Comunicaciones

Facultad de Ciencias Informáticas
Universidad Técnica de Manabí
Av. Urbina y Che Guevara, Portoviejo, Ecuador
revista.iys@utm.edu.ec

1. Introducción

El desarrollo acelerado de tecnologías como la IA junto a la Computación en la Nube ha cambiado de forma radical el diseño y funcionamiento de los sistemas digitales en diferentes áreas. Por un lado, la IA automatiza procesos, permite analizar grandes volúmenes de datos y facilita la toma de decisiones; por otro, la computación en la nube brinda una infraestructura flexible, escalable y accesible para desplegar estas soluciones sin requerir recursos locales (Chen & Ying, 2022; Sunyaev, 2020; Bello et al., 2021). Esta combinación ha contribuido al aumento de los modelos de IA como servicio (AIaaS) para aplicaciones inteligentes en sectores como: salud, comercio electrónico, educación y seguridad pública (Abioye et al., 2021; Zhang et al., 2024; Kyivska & Tsiutsiura, 2021). A su vez, esta convergencia ha dado lugar a arquitecturas distribuidas que dan soporte a operaciones en tiempo real, integran datos de diversas fuentes y se adaptan rápidamente a las variaciones del entorno. Por el contrario, este avance conlleva nuevos retos en términos de la seguridad de estos sistemas. Varios estudios han identificado amenazas recurrentes como el envenenamiento de datos, ataques adversarios, filtración de información sensible, manipulación de APIs y una gestión ineficiente del control de accesos (Calle & Barriga, 2025; Bai et al., 2025). Estas vulnerabilidades aumentan su gravedad en entornos multiusuario, donde los datos se almacenan de forma compartida y las interfaces abiertas, como las APIs, son puntos críticos de exposición. Si bien existen marcos normativos reconocidos, como la norma ISO/IEC 27001, el OWASP Top 10 for Large Language Models (OWASP, 2023) y el NIST AI Risk Management Framework (NIST, 2023), estos proporcionan guías generales que requieren ser adaptadas a un contexto determinado de IA en la nube. Por ejemplo, la mayoría de estos no contemplan de forma específica temas como el control del aprendizaje continuo, la integridad de los conjuntos de entrenamiento, la auditabilidad de decisiones automatizadas y la supervisión post-despliegue de modelos. Esto genera una diferencia significativa entre las recomendaciones normativas y la realidad operativa en entornos distribuidos basados en IA (Grechi et al., 2025; Flores-Cedeño et al., 2024; Gutiérrez Rodríguez & Castellanos-Sánchez, 2023; Acosta Cortez, 2024). Pese a que muchos estudios abordan los temas de seguridad de la IA o de la nube, suelen hacerlo por separado, sin proponer soluciones conjuntas o validar su aplicabilidad mediante casos experimentales. Por ejemplo, trabajos como el de Alnami et al. (2025) se enfocan puntualmente en la protección de modelos de IA frente a ataques adversarios, sin contemplar los riesgos asociados al entorno de despliegue en la nube. A diferencia del estudio mencionado, Bai et al. (2025) analizan vulnerabilidades específicas en plataformas de computación en la nube, como la gestión de accesos y el aislamiento de recursos, pero sin considerar la integración con sistemas de IA. Por su parte, Ahmed et al. (2025) proponen un marco de cifrado para proteger datos en entornos cloud, pero con un enfoque centrado en el almacenamiento seguro, sin integrar elementos de seguridad específicos de modelos de IA. Estos estudios, aunque valiosos, no consideran de forma integral los desafíos de seguridad que surgen al combinar IA y Cloud en un mismo entorno. En contraste, el presente estudio propone un enfoque original al integrar la revisión teórica y la validación práctica ante esta convergencia, lo que representa una contribución significativa frente a trabajos anteriores que solo abordan aquellos temas de manera aislada o sin evidencia experimental.

Según el contexto, se plantea la siguiente pregunta de investigación: ¿Cómo pueden implementarse buenas prácticas de seguridad en un proyecto de IA desplegado en la nube, y qué tan favorables resultan frente a amenazas comunes? A su vez, surge la siguiente hipótesis: La implementación de buenas prácticas de seguridad permite fortalecer los sistemas de IA desplegados en la nube. El propósito de la investigación es detectar, agrupar y validar buenas prácticas de seguridad aplicables en proyectos de inteligencia artificial en la nube. Para ello, se llevó a cabo una revisión sistemática de literatura apoyada en artículos científicos del periodo 2021-2025, siguiendo los lineamientos de la metodología PRISMA 2020. Las buenas prácticas extraídas buscan mitigar vulnerabilidades críticas como por ejemplo el mencionado acceso no autorizado, la fuga de información a través de APIs, la falta de cifrado o la ausencia de registro sobre los procesos de toma de decisiones automatizadas. Como aporte, se desarrolló un prototipo basado en inteligencia artificial al que se le incorporaron algunas de las buenas prácticas de seguridad identificadas en la revisión, tales como el uso de tokens para proteger APIs, el cifrado de datos sensibles en tránsito y en reposo, entre otras. Este prototipo fue desplegado en un entorno cloud y se sometió a pruebas para evaluar la aplicabilidad de dichas prácticas.

2. Materiales y Métodos

Esta investigación se realizó siguiendo la metodología Preferred Reporting Items for Systematic Reviews and Meta-Analyses

(PRISMA 2020). Esta fue seleccionada por su reconocimiento internacional y por ofrecer un marco bien estructurado que permite organizar, seleccionar y evaluar críticamente la literatura científica relevante. La metodología PRISMA está estructurada en tres fases principales: Identificación, Cribado e Inclusión, ilustradas y brevemente descritas en la Figura 1:

e-ISNN 2550-6730 | Vol. 9, Núm. 2 (152-163): Julio-Diciembre, 2025 | DOI: 10.33936/isrtic.v9i2.7687

Figura 1. Fases de la Metodología PRISMA 2020.
Fuente: Los Autores

2.1. Fase 1: Identificación

2.1.1. Preguntas de Investigación

Las siguientes preguntas guiaron la revisión sistemática y el desarrollo del prototipo práctico: ¿Qué buenas prácticas de seguridad propone la literatura científica reciente para proteger sistemas de inteligencia artificial desplegados en entornos de computación en la nube?, ¿Qué mecanismos, marcos o controles se han documentado para mitigar amenazas comunes en soluciones de IA en la nube? y ¿Cómo pueden implementarse estas buenas prácticas y qué tan aplicables son?

2.1.2. Palabras Clave y Cadenas de Búsqueda

Las palabras clave se seleccionaron considerando su frecuencia de uso en la literatura científica, así como su capacidad para recuperar estudios pertinentes que aporten evidencia sobre la mitigación de amenazas comunes en sistemas distribuidos basados en IA. Estas se complementaron con términos asociados a estándares internacionales, marcos de referencia y estrategias aplicadas (como OWASP, ISO/IEC 27001, cifrado, autenticación, etc.), con el fin de captar investigaciones que presentarán controles técnicos validados o propuestas concretas de implementación. A continuación, en la Tabla 1 se detallan las palabras clave seleccionadas, junto con su justificación. De la misma manera en la Tabla 2 se evidencia la construcción de las cadenas de búsqueda específicas en inglés y español, adaptadas a los motores de búsqueda académicos utilizados:

2.1.3. Selección de Bases de Datos Científicas

Para la ejecución de la búsqueda bibliográfica se seleccionaron tres bases de datos ampliamente reconocidas por su cobertura académica y rigor científico: Scopus, Google Scholar y Redalyc. Estas fuentes fueron elegidas por su diversidad de publicaciones, su acceso a literatura multidisciplinaria y la posibilidad de filtrar resultados relevantes mediante el uso de operadores booleanos, lo cual facilita una recuperación más precisa de estudios relacionados con la aplicación de la inteligencia artificial en

Tabla 1: Palabras clave utilizadas en la búsqueda bibliográfica.
Fuente: Los Autores

Palabra Clave	Justificación y contexto de uso
Inteligencia Artificial	Concepto central del estudio. Se refiere al desarrollo de sistemas capaces de realizar tareas que normalmente requieren inteligencia humana, como el análisis de datos, toma de decisiones y automatización.
Cloud Computing	Describe el entorno de operación de los sistemas de IA estudiados.
Seguridad informática	Permite enfocar la búsqueda en estudios que aborden amenazas, controles o marcos de protección.
Buenas prácticas	Palabra clave orientada a identificar recomendaciones prácticas o estrategias de implementación.
Protección de datos	Ayuda a filtrar literatura centrada en el resguardo de información personal y sensible.
Marcos de seguridad	Se refiere a estándares, normas y frameworks como OWASP, ISO/IEC 27001, NIST RMF, entre otros.

Tabla 2: Cadenas de búsqueda utilizadas en los motores de bases de datos.
Fuente: Los Autores

Idioma	Cadenas de Búsqueda
Inglés	(Artificial Intelligence OR AI systems) AND (Cloud computing OR distributed systems) AND (security best practices OR cybersecurity frameworks OR data protection OR “security controls)
Español	(Inteligencia Artificial OR sistemas de IA) AND (computación en la nube OR sistemas distribuidos) AND (buenas prácticas de seguridad OR marcos de ciberseguridad OR protección de datos OR controles de seguridad)

entornos de computación en la nube. En la figura 2 se muestra las bases de datos mencionadas:

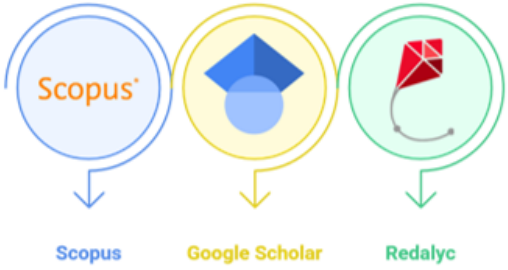


Figura 2. Bases de Datos Científicas utilizadas.
Fuente: Los Autores

e-ISNN 2550-6730 | Vol. 9, Núm. 2 (152-163): Julio-Diciembre, 2025 | DOI: 10.33936/isrtic.v9i2.7687

INFORMÁTICA Y SISTEMAS
Revista de Tecnologías de la Informática y las Comunicaciones

Facultad de Ciencias Informáticas
Universidad Técnica de Manabí
Av. Urbina y Che Guevara, Portoviejo, Ecuador
revista.iys@utm.edu.ec

2.2 Fase 2: Cribado

2.2.1. Criterios de Inclusión y Exclusión

Los resultados obtenidos durante la fase anterior fueron filtrados, de forma que se descartaron los documentos que no cumplían con los objetivos específicos de la investigación. El eje fundamental fue garantizar que los artículos analizados aportaran en la identificación de enunciados que enumeren buenas prácticas de la seguridad en entornos de IA, concretamente cuando se despliegan en una nube. A continuación, en la Tabla 3 se detallan los criterios definidos:

Tabla 3: Aplicación de criterios de inclusión y exclusiónFuente: Las autoras

Tipo de criterio	Descripción
Criterios de Inclusión	- Artículos científicos publicados entre enero de 2021 y marzo de 2025. -Artículos revisados por expertos/as, y publicados en revistas indexadas o en repositorios institucionales. - Artículos que aborden explícitamente la seguridad de sistemas de IA desplegados en entornos de computación en la nube. - Artículos que propongan buenas prácticas, marcos normativos o controles técnicos aplicables a dichos entornos. - Artículos que estén disponibles en texto completo ya sea en idioma inglés o español.
Criterios de Exclusión	- Documentos previos a 2021. - Documentos centrados unicamente en IA o en el cloud computing sin abordar su interrelación. - Informes técnicos, TFGs o resúmenes/abstracts que carecieran de un desarrollo metodológico en su totalidad. - Artículos duplicados en bases de datos, o bien, documentos que no contuvieran suficiente información para su posible verificación y análisis.

2.2.2. Selección de Artículos

Inicialmente se obtuvieron 6.049 registros provenientes de las bases de datos científicas mencionadas. En la fase de cribado, se eliminaron 1.246 registros duplicados. Posteriormente, al aplicar criterios de inclusión y exclusión, se descartaron 4.764 documentos por no abordar directamente temas de seguridad en sistemas de inteligencia artificial desplegados en la nube, por carecer de validez metodológica o por estar incompletos. Luego, se revisaron 39 artículos a texto completo, de los cuales 14 fueron excluidos por no presentar evidencia útil para la formulación de buenas prácticas de seguridad. Finalmente, 25 artículos científicos fueron seleccionados para aplicarlos en la investigación ya que cumplen con los criterios establecidos, y aportan significativamente al objetivo central de la investigación. Este proceso se resume gráficamente en la Figura 3, siguiendo el diagrama de flujo PRISMA 2020:

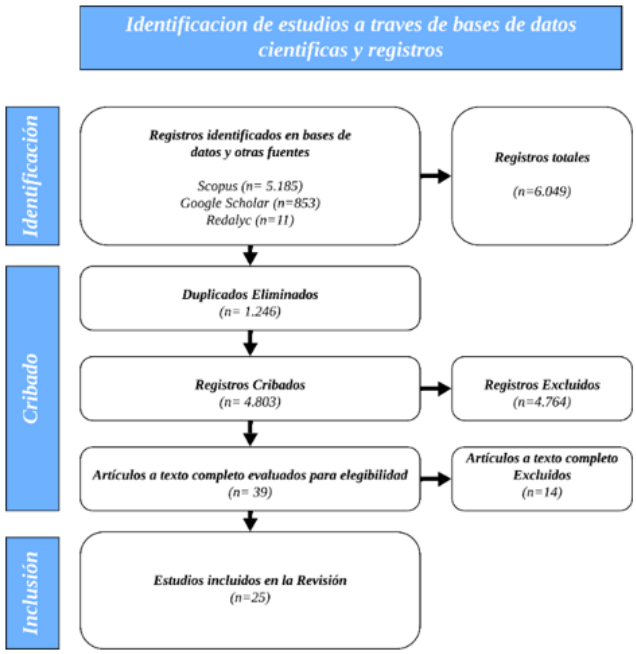


Figura 3. Diagrama PRISMA 2020.

2.3 Fase 3: Inclusión

2.3.1. Evaluación de Relevancia

Se aseguró que los artículos no solo cumplieran con los criterios aplicados, sino que también fueran metodológicamente sólidos y fuera posible identificar en ellos buenas prácticas de seguridad aplicables al despliegue de sistemas de IA en la nube. Los aspectos que se evaluaron están descritos en la Tabla 4:

2.3.2. Extracción de Datos

Se revisó cada uno de los 25 artículos seleccionados con el fin de identificar información que pudiera contribuir a la formulación, análisis y validación de buenas prácticas de seguridad en sistemas de IA desplegados en la nube. La extracción de datos se organizó según una serie de aspectos clave, como se evidencia en la Tabla 5:

e-ISNN 2550-6730 | Vol. 9, Núm. 2 (152-163): Julio-Diciembre, 2025 | DOI: 10.33936/isrtic.v9i2.7687

Tabla 4: Criterios aplicados para la evaluación de relevancia de estudios.
Fuente: Los Autores

Aspecto Evaluado	Detalle
Definición del objetivo del estudio	Se evaluó si el artículo presenta de forma clara y específica los objetivos o preguntas de investigación.
Transparencia metodológica	Se verificó la descripción de la metodología utilizada, incluyendo enfoque, instrumentos o criterios de selección.
Profundidad en el tratamiento de amenazas o riesgos	Se analizó la profundidad con la que el estudio aborda amenazas de seguridad, prácticas de mitigación o análisis de vulnerabilidades específicas.
Integración con marcos normativos o estándares	Se consideró la incorporación de marcos como ISO/IEC 27001, OWASP, NIST RMF u otros enfoques éticos o regulatorios aplicables.
Aplicabilidad de hallazgos en contextos reales o simulados	Se valoró si las propuestas, recomendaciones o resultados del artículo son susceptibles de ser implementados en proyectos reales o prototipos simulados.

Tabla 5: Aspectos analizados a partir de los resultados de la extracción de datos.
Fuente: Los Autores

Aspecto Evaluado	Detalle
Tipo de estudio	Categoriza los artículos según corresponda: una revisión, estudio de caso, propuesta técnica, mapeo sistemático, entre otros.
Contexto de aplicación	Describe el sector o área en el que se desarrolla la investigación.
Percepciones respecto a los riesgos en el tema de la seguridad	Expone las debilidades señaladas en el estudio, como el acceso ilegal, la alteración de API, ataques maliciosos, etc.
Buenas prácticas o controles sugeridos	Detalla acciones o recomendaciones para mitigar los riesgos mencionados.
Normas o lineamientos utilizados como referencia	Indica si el texto examinado menciona o se ajusta a estándares reconocidos como OWASP, ISO/IEC 27001, NIST u otros similares.
Nivel de validación o experimentación	Señala si las buenas prácticas fueron validadas teóricamente, mediante simulaciones, prototipos o casos reales.
Aplicabilidad práctica	Evalúa el potencial de las buenas prácticas detectadas de ser implementadas en entornos reales o simulados.
Limitaciones del estudio	Identifica restricciones metodológicas o temáticas que puedan afectar la generalización de los hallazgos.

2.4 Metodología de Desarrollo del Prototipo.

La presente investigación adopta un enfoque experimental centrado en el desarrollo de un prototipo, con el fin de validar las buenas prácticas establecidas en la revisión sistemática, fue desarrollado en el entorno de programación Visual Studio Code, usando el lenguaje Python en su versión 3.13. Para el backend, se implementó el framework FastAPI para construir una API REST segura, la autenticación fue basada en JSON Web tokens (tokens JWT) y también se empleó PostgreSQL para ser desplegado en la nube a través del servicio Aiven. Toda la información de inicios de sesión y eventos de seguridad fueron almacenados en dicha base de datos. Asimismo, el prototipo se integró con un motor de inteligencia artifical Groq, accesible por API. La fase de pruebas y validación se ejecutó utilizando la herramienta Thunder Client, permitiendo simular peticiones HTTP controladas a la API y verificar el comportamiento del sistema ante diferentes escenarios. En la Figura 4 se observa un diagrama descriptivo de la arquitectura del prototipo:

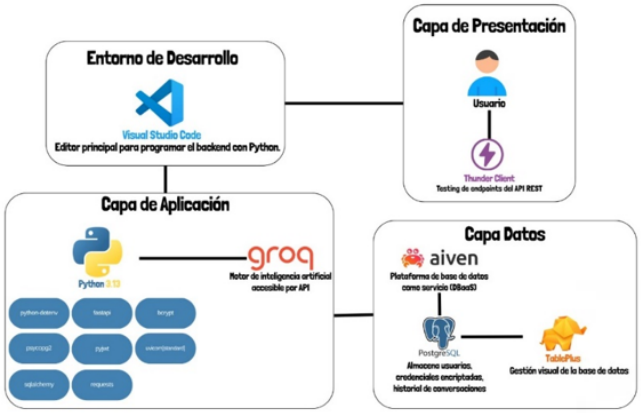


Figura 4. Arquitectura del prototipo.
Fuente: Los Autores

3. Resultados y Discusión

3.1 Resultados de Revisión Literaria

3.1.1. Identificación de Amenazas Recurrentes

Los estudios analizados revelan un conjunto de amenazas comunes que afectan directamente la seguridad de los sistemas de inteligencia artificial cuando son desplegados en entornos de computación en la nube. Estas amenazas se presentan en contextos diversos como salud, redes empresariales, investigación científica, transporte inteligente y sistemas de análisis forense digital, lo que evidencia la transversalidad del problema. En la Figura 5 se ilustran las amenazas más destacadas en los artículos revisados:

3.1.3. Síntesis de Buenas Prácticas extraídas

A partir del análisis de los artículos, se identificaron varias buenas prácticas de seguridad aplicables en el diseño, despliegue y operación de proyectos de inteligencia artificial en la nube:

e-ISNN 2550-6730 | Vol. 9, Núm. 2 (152-163): Julio-Diciembre, 2025 | DOI: 10.33936/isrtic.v9i2.7687

INFORMÁTICA Y SISTEMAS
Revista de Tecnologías de la Informática y las Comunicaciones

Facultad de Ciencias Informáticas
Universidad Técnica de Manabí
Av. Urbina y Che Guevara, Portoviejo, Ecuador
revista.iys@utm.edu.ec

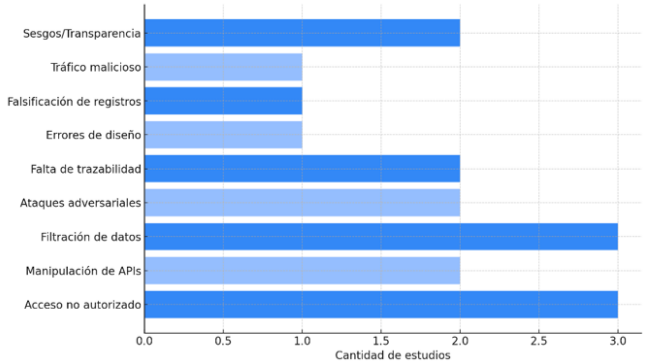


Figura 5. Frecuencia de amenazas abordadas en los artículos.
Fuente: Los Autores

• Autenticación multifactor y segmentación de acceso.

• Cifrado homomórfico.

• Blockchain.

• Análisis en tiempo real de tráfico y verificación de consistencia en datos distribuidos.

• Modelado de riesgos con coanálisis de seguridad y confiabilidad (safety).

• Supervisión continua, monitoreo de eventos y registro de auditoría.

• Políticas de transparencia algorítmica y ética en IA.

Aunque estas prácticas son heterogéneas en su formulación, convergen en la necesidad de incorporar controles proactivos desde las fases tempranas de diseño y no solo en la operación post-despliegue.

3.1.4. Discusión Crítica

Uno de los hallazgos más relevantes es el hecho de que no existe ninguna estandarización entre los marcos de implementación, pues si bien algunos estudios se refieren a normas como ISO/IEC 27001, OWASP o NIST RMF (Calle-Méndez & Barriga-Andrade, 2025; Grechi et al., 2025), la mayoría de ellos proponen soluciones ad-hoc que son muy específicas a un contexto concreto. Además, únicamente cuatro de los estudios propuestos incluyen algún tipo de validación técnica (Alnami et al, 2025; Bai et al., 2025; Ahmed et al., 2025; Cajamarca et al., 2025), aunque centradas en IA o Cloud por separado, sugiriendo que las buenas prácticas no han sido suficientemente validadas. Esta falta de validación técnica será abordada en la fase experimental de la investigación a través del uso de un prototipo funcional. Además, casi la mitad de los estudios analizados cuenta con un enfoque meramente teórico (45,5%), mientras que solo una minoría plantean soluciones mediante simulaciones, prototipos o entornos reales (cada uno con el 18,2%), lo cual es una muestra clara de la creciente necesidad de investigaciones que cuenten con validación empírica. Lo mencionado se observa a continuación en la Figura 6:

Frente a esto, algunos autores han planteado la posibilidad de una convergencia de marcos mediante arquitecturas híbridas que combinen, por ejemplo, el NIST AI RMF como marco de análisis de riesgos, controles técnicos del OWASP Top 10 for LLMs y complementando con los lineamientos de gestión de la ISO/IEC 27001 (Sharma & Dhiman, 2025). Así se combinaría

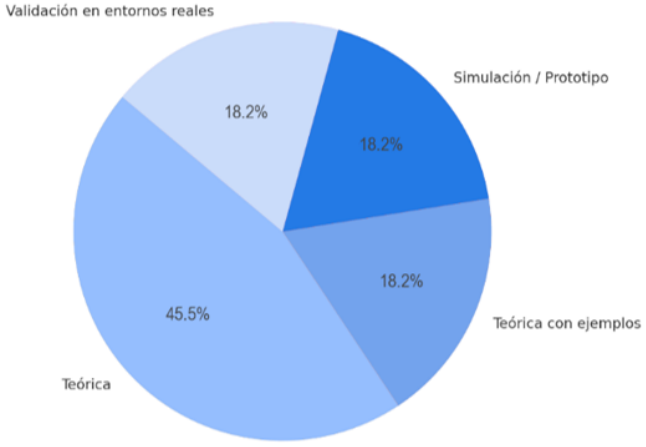


Figura 6. Nivel de validación de los estudios analizados.
Fuente: Los Autores

la gobernanza, el control técnico y la gestión de riesgos en un mismo enfoque, permitiendo una cobertura completa de proyectos de IA en la nube. Por otro lado, se observa la validez de enfoques complementarios como blockchain y cifrado avanzado que permiten cubrir las limitaciones clásicas en la protección de datos. Estudios como los de Lamar Peña et al. (2024), Vaca & Dulce-Villarreal (2024), Rueda-Castañeda et al. (2024), Dewangan y Chandrakar (2025) y Wazid et al. (2025) han demostrado la utilidad del blockchain para garantizar la trazabilidad, la autenticación segura y la emisión confiable de datos digitales, aportando valiosos mecanismos para fortalecer la integridad de modelos de IA desplegados en la nube. De manera similar, se han explorado mecanismos de autenticación biométrica multimodal en entornos cloud (Lin et al., 2023; Cui et al., 2025) y sistemas de detección de intrusiones basados en IA aplicados a infraestructuras críticas (Gujar, 2024), aunque estos enfoques siguen siendo específicos y carecen de validación conjunta en arquitecturas distribuidas de IA en la nube. Por último, se detecta un creciente interés por las preocupaciones éticas, pero aún no

e-ISNN 2550-6730 | Vol. 9, Núm. 2 (152-163): Julio-Diciembre, 2025 | DOI: 10.33936/isrtic.v9i2.7687

están representadas en políticas técnicas claras, lo que representa una futura oportunidad de convergencia entre marcos normativos y principios éticos aplicables.

3.1.5. Propuesta Preliminar de Buenas Prácticas a validar

Como resultado de la revisión, se propone la siguiente lista preliminar de buenas prácticas para ser validadas en el prototipo experimental del estudio:

• Protección de API mediante tokens de acceso temporales.

• Autenticación multifactor y segregación de privilegios.

• Cifrado de datos sensibles en tránsito y en reposo.

• Registro detallado de eventos de seguridad para auditoría.

como el sistema diferencia entre usuario (user) y administrador (admin), permitiendo una gestión de privilegios efectiva. Además, el sistema de bloqueo automático de cuentas tras múltiples intentos fallidos de inicio de sesión y la trazabilidad de cada evento queda registrado en la base de datos, accesible mediante auditorias que podemos observar en la Figura 8. Por último, la Figura 9 detalla que de los 100 intentos con tokens JWT entre inválidos y expirados fueron 100% rechazados por parte del sistema, garantizando que las sesiones no se mantengan abiertas indefinidamente y previene ataques por fuerza bruta.

3.2.2 Registro de Eventos de Seguridad

El sistema realiza auditoría detallada de eventos críticos, incluyendo autenticaciones exitosas o fallidas, intentos de acceso desde IPs bloqueadas y uso de token inválidos o expirados. En las Figuras 10 y 11 se observan las validaciones de intentos

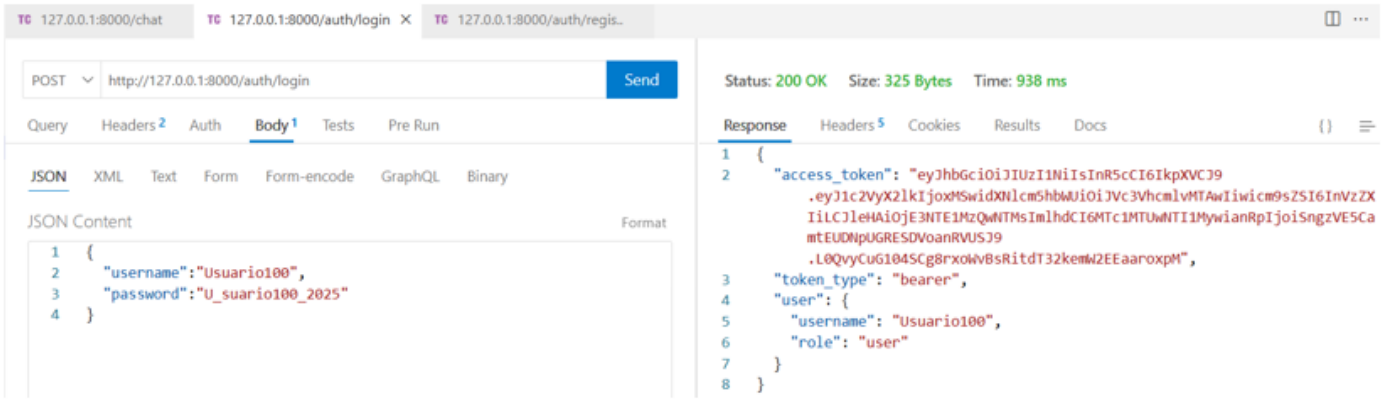
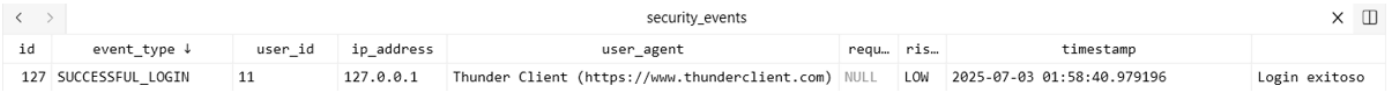


Figura 7. Validación en el inicio de sesión mediante Thunder Client (Credenciales Válidas) y generación del token JWT.
Fuente: Los Autores



id	event_type	user_id	ip_address	user_agent	requ…	ris…	timestamp	
127	SUCCESSFUL_LOGIN	11	127.0.0.1	Thunder Client (https://www.thunderclient.com)	NULL	LOW	2025-07-03 01:58:40.979196	Login exitoso

Figura 8. Registro del evento de inicio exitoso en PostgreSQL.
Fuente: Los Autores

• Verificación de integridad mediante tecnologías distribuidas (ej. IOTA).

• Supervisión automática de tráfico mediante IA.

• Revisión de código y modelos bajo principios de transparencia.

Estas prácticas se incorporarán al prototipo con el fin de verificar su factibilidad técnica y analizar su efecto en la mitigación de riesgos dentro de un entorno de computación en la nube.

3.2 Resultados del Prototipo

3.2.1. Acceso al Sistema

En la Figura 7 se visualiza la validación del mecanismo de autenticación basado en tokens JWT con expiración automática, lo que asegura sesiones temporales y protegidas, también de

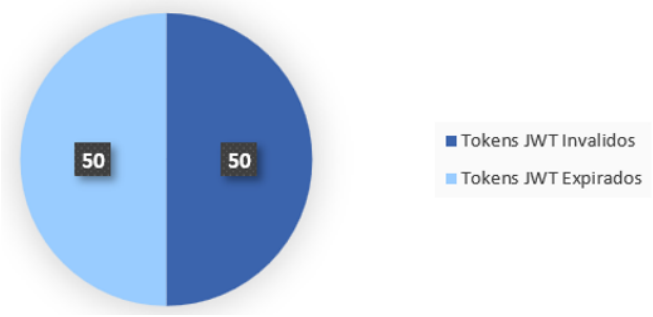


Figura 9. Distribución de 100 intentos fallidos de uso del chat con tokens JWT inválidos y expirados
Fuente: Los Autores

e-ISNN 2550-6730 | Vol. 9, Núm. 2 (152-163): Julio-Diciembre, 2025 | DOI: 10.33936/isrtic.v9i2.7687

maliciosos y bloqueo de patrones de riesgos con respuestas automáticas. Todos los eventos son almacenados con su nivel de riesgo, usuario, IP y hora exacta. Al final, en la Figura 12 se muestra que, de 100 eventos aleatorios legítimos y maliciosos, fueron registrados en su totalidad con detalles proporcionando una trazabilidad completa para análisis de los datos y detención de amenazas.

rate_limit (límite máximo 10 peticiones en un minuto por usuario) para mitigar abusos del sistema. En la Figura 17 se muestra la validación del método aplicado. Además, en la Figura 18 se demuestra que mediante pruebas de estrés y comparación de tiempos de respuesta promedio con respecto al límite de peticiones, se pudo obtener que dicho control contribuye significativamente a mantener un rendimiento óptimo del sistema, garantizando tiempos de respuesta estables.

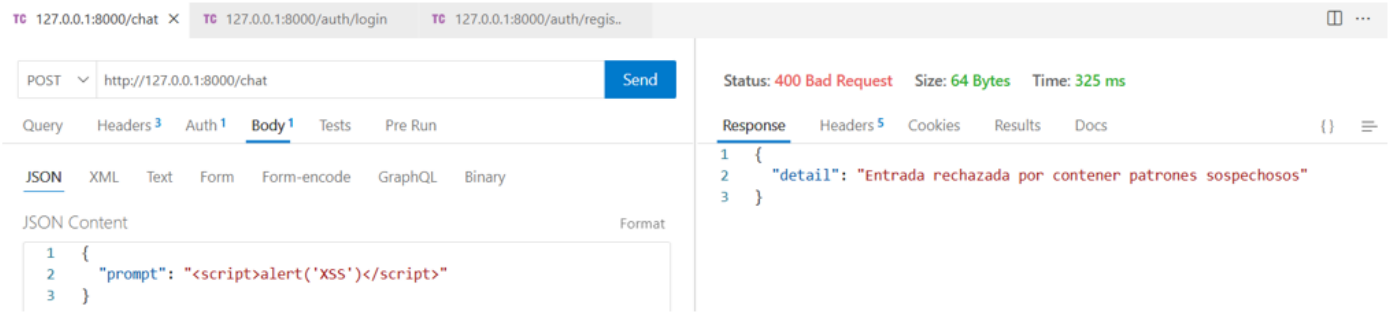


Figura 10. Validación y clasificación de intentos maliciosos en el endpoint (/chat).
Fuente: Los Autores

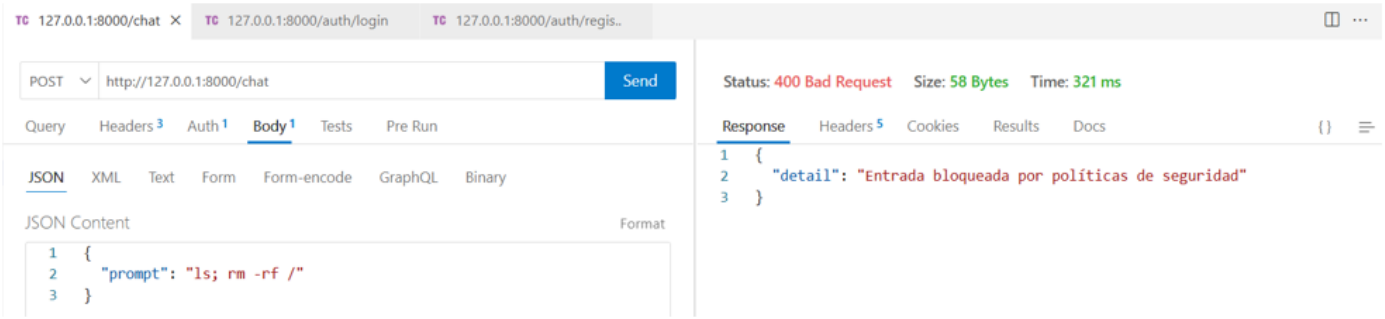


Figura 11. Validación de detección y bloqueo de patrones de riesgo en el endpoint (/chat).
Fuente: Los Autores

3.2.3 Gestión de Vulnerabilidades

El sistema integra un módulo de detección de vulnerabilidades que analiza en tiempo real las solicitudes enviadas a la IA. Este módulo identifica patrones relacionados con inyecciones SQL, XSS, inyecciones de comandos y prompt injection en entradas del usuario como se muestra en la Figura 13. La solicitud puede ser bloqueada o permitida bajo advertencia dependiendo del nivel de riesgo detectado, pero siempre registrada en la base de datos. En la Figura 14 se detalla uno de los registros de Inyección SQL con Patrones maliciosos. Para terminar, la Figura 15 destaca que el sistema detectó y bloqueó el 100% de ataques aleatorios simulados clasificándolos por nivel de riesgo, así se minimiza la superficie de ataque mediante una respuesta automática, permitiendo la priorización de amenazas.

3.2.4. Controles de Red y Protección del Entorno

En la Figura 16 se detalla cómo se configuró el método check_

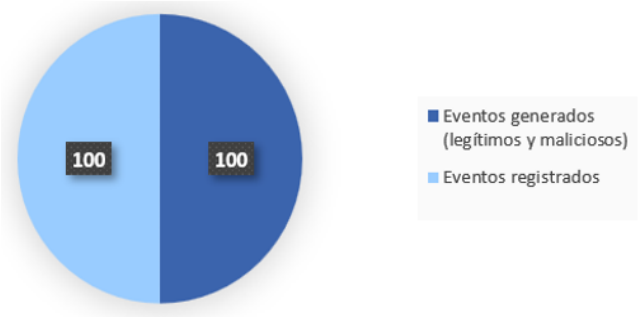


Figura 12. Generación de eventos vs el registro en la base de datos
Fuente: Los Autores

Se puede observar en la Figura 19 la lógica de bloqueo automático por nivel de riesgo y conjuntamente su validación en la Figura 20 con el bloqueo automático de dirección IP por exceder límite

e-ISNN 2550-6730 | **Vol. 9, Núm. 2 (152-163): Julio-Diciembre, 2025** | **DOI: 10.33936/isrtic.v9i2.7687**

INFORMÁTICA Y SISTEMAS
Revista de Tecnologías de la Informática y las Comunicaciones

Facultad de Ciencias Informáticas
Universidad Técnica de Manabí
Av. Urbina y Che Guevara, Portoviejo, Ecuador
revista.iys@utm.edu.ec

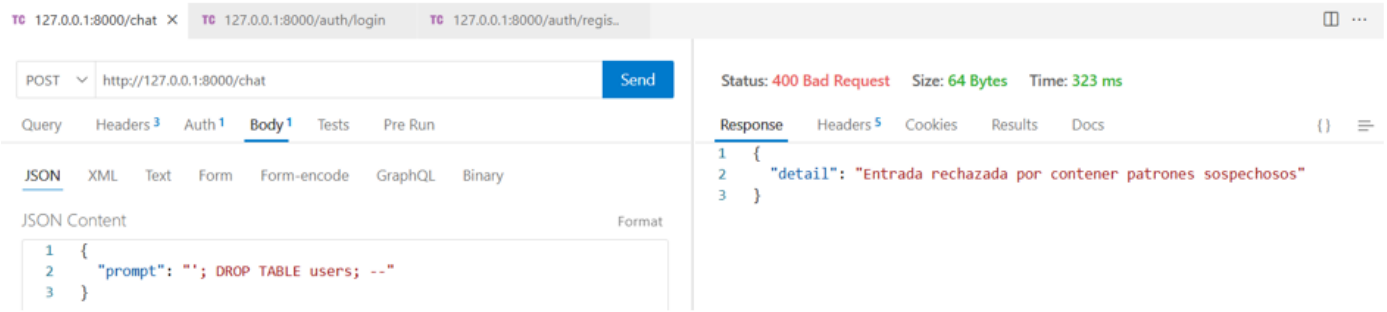


Figura 13. Validación y clasificación de intentos maliciosos en el endpoint (/chat).
Fuente: Los Autores

Figura 14. Registro de inyecciones SQL con patrones maliciosos en PostgreSQL.
Fuente: Los Autores

o generar entradas críticas. En la Figura 21 se puede visualizar que de 100 intentos aleatorios de entradas críticas el sistema bloqueó el 100% de las direcciones IP involucradas. Así mismo, se registró automáticamente en la base de datos, incluyendo detalles como la fecha, hora, dirección de IP del usuario y el contenido de la entrada maliciosa. Esto evidencia la efectividad del mecanismo de detección y respuesta ante amenazas.

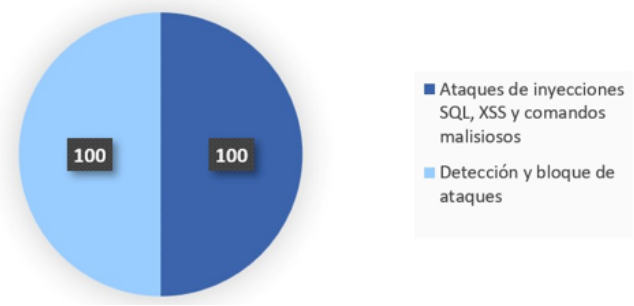


Figura 15. Detección y bloqueo de inyecciones SQL, XSS y comandos maliciosos
Fuente: Los Autores

Los resultados confirman la hipótesis inicial del estudio: la implementación de buenas prácticas de seguridad si permiten fortalecer los sistemas de IA desplegados en la nube. En particular, se evidencia que estas prácticas establecen mecanismos efectivos de control de acceso, detección y mitigación de vulnerabilidades, auditoría de eventos y control del tráfico de red. Estas no solo incrementan la capacidad de los sistemas ante amenazas comunes, sino que también garantizan la integridad, disponibilidad y trazabilidad de los servicios de IA en la nube.

```python
def check_rate_limit(self, ip_address: str, limit: int = 10, window: int = 60) -> bool:
    current_time = time.time()
    if ip_address not in self.rate_limit_storage:
        self.rate_limit_storage[ip_address] = []
    self.rate_limit_storage[ip_address] = [
        ts for ts in self.rate_limit_storage[ip_address]
        if current_time - ts < window
    ]
    if len(self.rate_limit_storage[ip_address]) >= limit:
        return False
    self.rate_limit_storage[ip_address].append(current_time)
    return True
```

Figura 16. Método check_rate_limit.
Fuente: Los Autores

4. Conclusiones

La integración de la inteligencia artificial en entornos de computación en la nube ha abierto nuevas oportunidades para la automatización, el procesamiento inteligente de datos y el desarrollo de soluciones escalables. No obstante, esta convergencia también ha traído amenazas importantes desde el punto de vista de la seguridad, en especial en la gestión de acceso a la información, la exposición de API y la protección de datos sensibles. El estudio, basado en una revisión sistemática de literatura científica reciente (2021-2025), ha determinado un conjunto representativo de buenas prácticas y mecanismos de protección que han sido propuestos para minimizar dichos riesgos, entre los que cabe mencionar: la autenticación fuerte, el cifrado, la auditoría de eventos y el control granular de los permisos. Como aporte práctico, se desarrolló un prototipo basado en inteligencia artificial, al que se le incorporaron algunas de las buenas prácticas identificadas. Este fue desplegado en un entorno Cloud, donde se aplicaron controles como tokens JWT, validación de entradas, cifrado en tránsito y reposo, y registro de eventos críticos. Si bien se trató de una simulación de alcance

e-ISNN 2550-6730 | Vol. 9, Núm. 2 (152-163): Julio-Diciembre, 2025 | DOI: 10.33936/isrtic.v9i2.7687

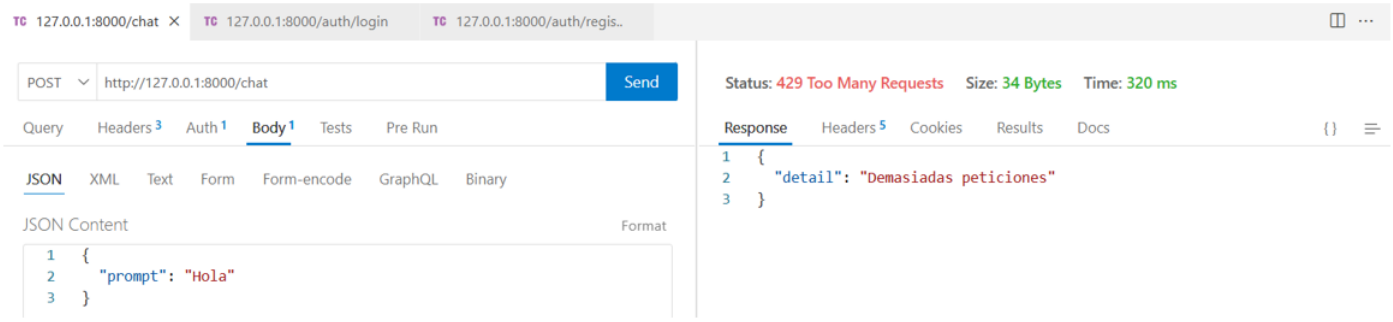


Figura 17. Validación del Método check_rate_limit en el endpoint (/chat).
Fuente: Los Autores

limitado, los resultados obtenidos permitieron comprobar que dichas prácticas son viables técnicamente y efectivas ante amenazas comunes en estos entornos, como inyecciones de comandos, abuso de endpoints o intentos de acceso masivo.

Este estudio pone de manifiesto que aplicar buenas prácticas de seguridad en proyectos de inteligencia artificial en la nube resulta, no sólo viable, sino necesario, incluso en las primeras etapas de prototipado. Además, refuerza la necesidad de contar con una guía fundamentada, basada en evidencia científica, que oriente a desarrolladores, ingenieros de seguridad y decisores en la puesta en práctica de los controles necesarios desde las primeras instancias de desarrollo. No obstante, se da considerando que existen limitaciones propias del uso de entornos simulados, por ello se sugiere que futuras investigaciones apunten a evaluar estos mecanismos en entornos de producción reales, donde juegan un papel relevante factores como la latencia, la escalabilidad, el rendimiento del sistema o la normativa vigente. También se recomienda incluir marcos éticos y de gobernanza para asegurar que se diseñen sistemas de inteligencia artificial seguros, pero también transparentes, auditables y que respeten la privacidad y los derechos de los usuarios. Desde esta perspectiva, el presente artículo supone un primer paso en el diseño de sistemas de inteligencia artificial seguros en la nube, así como una base para futuras investigaciones que pretendan validar, extender o adaptar las buenas prácticas halladas en contextos más complejos, distribuidos y dinámicos.

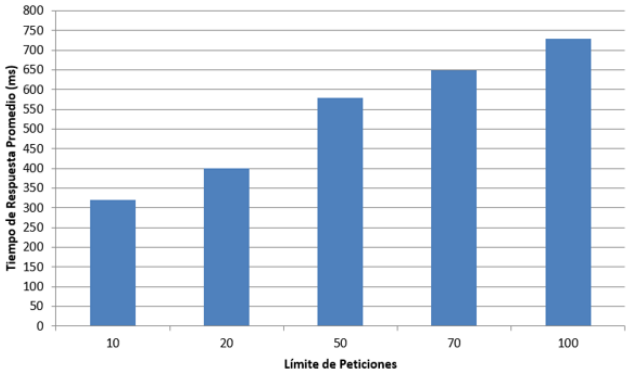


Figura 18. Tiempo de respuesta con base al límite de peticiones.

```
if risk_level == "CRITICAL":
    security_manager.block_ip(client_ip)
    raise HTTPException(status_code=400, detail="Entrada bloqueada por políticas de seguridad")
```

Figura 19. Lógica de Bloqueo Automático por Nivel de Riesgo Crítico.
Fuente: Los Autores

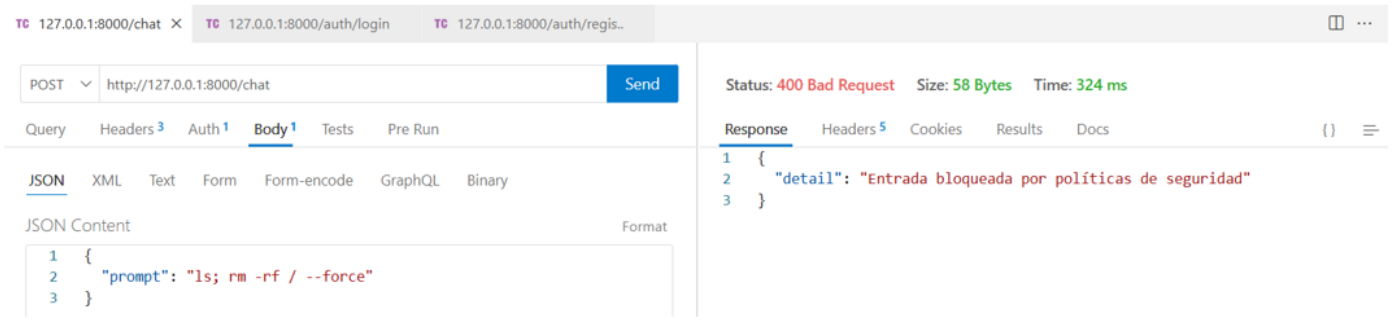


Figura 20. Lógica de Bloqueo Automático por Nivel de Riesgo Crítico.
Fuente: Los Autores

e-ISNN 2550-6730 | **Vol. 9, Núm. 2 (152-163): Julio-Diciembre, 2025** | **DOI: 10.33936/isrtic.v9i2.7687**

INFORMÁTICA Y SISTEMAS
Revista de Tecnologías de la Informática y las Comunicaciones

Facultad de Ciencias Informáticas
Universidad Técnica de Manabí
Av. Urbina y Che Guevara, Portoviejo, Ecuador
revista.iys@utm.edu.ec

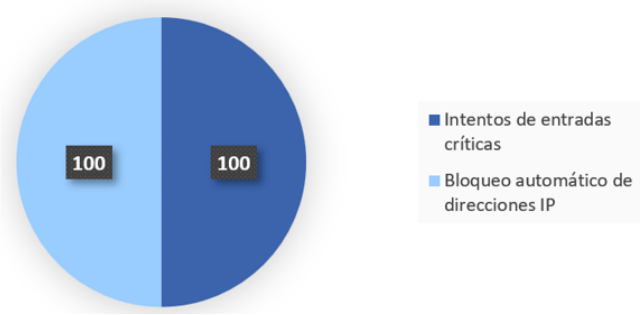


Figura 21. Intentos de entradas críticas vs bloqueo automático de direcciones IP.
Fuente: Los Autores

Contribución de los autores

Yasbeck Jemima Mora Chávez: Administración del proyecto, Investigación, Redacción y Metodología. **Reinaldo Benedicto Fernández Gonzalez:** Investigación, y edición del artículo. **Joofre Antonio Honores Tapia:** Metodología, revisión, redacción y edición del artículo. **Milton Rafael Valarezo Pardo:** Metodología, revisión, redacción y edición del artículo.

Declaración de conflictos de interés

Los autores manifiestan que no existe ningún conflicto de interés relacionado con este trabajo.

Agradecimientos

Agradecemos a nuestros tutores por su acompañamiento académico y sus recomendaciones técnicas, que fueron decisivas en el desarrollo y mejora de esta investigación. También hacemos extensivas las gratitudes a nuestras familias y a las personas más significativas para nosotros, con el aliento que nos brindaron, lo cual tuvo un impacto sin duda considerable para que podamos cumplir con este proceso de investigación.

Referencias bibliográficas

Abioye, S. O., Oyedele, L. O., Akanbi, L., Ajayi, A., Davila Delgado, J. M., Bilal, M., Akinade, O. O., & Ahmed, A. (2021). Inteligencia artificial en el sector de la construcción: Una revisión de la situación actual, oportunidades y retos futuros. *Journal of Building Engineering, 44*, Article 103299. https://doi.org/10.1016/j.jobe.2021.103299

Acosta Cortez, N. N. (2024). Impacto de la inteligencia artificial en la ciberseguridad empresarial: Un análisis crítico de la evolución de amenazas y medidas preventivas [Unpublished undergraduate thesis]. Universidad Técnica de Babahoyo, Facultad de Ciencias Informáticas. https://dspace.utb.edu.ec/handle/49000/15738

Ahmed, W., Iqbal, W., Hassan, A., Ahmad, A., Ullah, F., & Srivastava, G. (2025). Elevating e-health excellence with IOTA distributed ledger technology: Sustaining data integrity in next-gen fog-driven systems. *Future Generation Computer Systems, 168*, Article 107755. https://doi.org/10.1016/j.future.2024.107755

Alnami, H., Mahgoub, I., Al-Najada, H., & Alalwany, E. (2025). A distributed machine learning-based scheme for real-time highway traffic flow prediction in Internet of Vehicles. *Future Internet, 17*(3), Article 131. https://doi.org/10.3390/fi17030131

Bai, Y., Zhao, H., Shi, X., & Chen, L. (2025). Towards practical and privacy-preserving CNN inference service for cloud-based medical imaging analysis: *A homomorphic encryption-based approach. Computer Methods and Programs in Biomedicine, 261*, Article 108599. https://doi.org/10.1016/j.cmpb.2024.108599

Bello, S. A., Oyedele, L. O., Akinade, O. O., Bilal, M., Davila Delgado, J. M., Akanbi, L. A., Ajayi, A. O., & Owolabi, H. A. (2021). Cloud computing in construction industry: Use cases, benefits and challenges. *Automation in Construction, 122*, Article 103441. https://doi.org/10.1016/j.autcon.2020.103441

Cajamarca Sacoto, J. C., Chuquin Machangara, O. M., & Recalde Araujo, J. C. (2025). El papel de la inteligencia artificial en la ciberseguridad de redes empresariales. *Revista Retos Para La Investigación, 4*(1), 65–82. https://doi.org/10.62465/rri.v4n1.2025.126

Calle-Méndez, J. L., & Barriga-Andrade, J. J. (2025). Amenazas de seguridad asociadas con la integración de inteligencia artificial en sistemas de información: *Revisión sistemática. MQRInvestigar, 9*(1), Article e128. https://doi.org/10.56048/MQR20225.9.1.2025.e128

Chen, H.-P., & Ying, K.-C. (2022). Inteligencia artificial en el sector de la construcción: Principales trayectorias de desarrollo y perspectivas de futuro. *Applied Sciences, 12*(12), Article 5832. https://doi.org/10.3390/app12125832

Cui, W., Lin, Q., Shi, J., Zhou, X., Li, Z., Zhan, H., & Zhang, L. (2025). A secure and efficient framework for multimodal biometric authentication in cloud computing. *Applied*

e-ISNN 2550-6730 | Vol. 9, Núm. 2 (152-163): Julio-Diciembre, 2025 | DOI: 10.33936/isrtic.v9i2.7687

Sciences, 15(7), Article 3827. https://doi.org/10.3390/app15073827

Dewangan, N. K., & Chandrakar, P. (2025). TreatChain: A patient-centric treatment cycle security system using blockchain and AI in the cloud. *Peer-to-Peer Networking and Applications, 18*(3), Article 154. https://doi.org/10.1007/s12083-025-01984-z

Flores-Cedeño, P. R., Zambrano-Pilay, E. C., & Chiriboga-Mendoza, F. R. (2024). Seguridad informática e inteligencia artificial en la investigación científica. *Revista Científica INGENIAR: Ingeniería, Tecnología e Investigación, 7*(13), 2–10. https://www.journalingeniar.org/index.php/ingeniar/article/view/177

Grechi, V. L., de Oliveira, A. L., & Braga, R. T. V. (2025). Model-driven safety and security co-analysis: A systematic literature review. *Journal of Systems and Software, 220*, Article 112251. https://doi.org/10.1016/j.jss.2024.112251

Gujar, S. S. (2024, December). AI-enhanced intrusion detection systems for strengthening critical infrastructure security [Conference session]. 2024 Global Conference on Communications and Information Technologies (GCCIT), Bangalore, India. https://doi.org/10.1109/GCCIT63234.2024.10861950

Gutiérrez Rodríguez, J. D., & Castellanos-Sánchez, M. (2023). Transparencia algorítmica y estado abierto en Colombia. *Reflexión Política, 25*(52), 6–21. https://doi.org/10.29375/01240781.4789

Jiang, Y., Ye, Y., Zhao, H., Zhang, S., Cao, Y., & Gu, J. (2021). Analysis of smart water conservancy. *Shuili Xuebao/ Journal of Hydraulic Engineering, 52*(11), 1355–1368. https://doi.org/10.13243/j.cnki.slxb.20210633

Kyivska, K. I., & Tsiutsiura, S. (2021). Implementation of artificial intelligence in the construction industry and analysis of existing technologies. *Technology Audit and Production Reserves, 2*(2), Article 58. https://doi.org/10.15587/2706-5448.2021.229532

Lamar Peña, F. S., Vega Mite, G. A., Honores Tapia, J. A., & Cárdenas Villavicencio, O. E. (2024). Validación y emisión de certificados en educación superior utilizando tecnología blockchain. Informática y Sistemas: *Revista de Tecnologías de la Informática y las Comunicaciones, 8*(1), 36–51. https://doi.org/10.33936/isrtic.v8i1.6535

Lin, L., Chen, C., Pan, L., Zhang, L. Y., Wang, Z., Zhang, J., & Xiang, Y. (2023). A survey of PPG's application in authentication. *Computers & Security, 135*, Article 103488. https://doi.org/10.1016/j.cose.2023.103488

Pan, Y., & Zhang, L. (2021). Roles of artificial intelligence in construction engineering and management: A critical review and future trends. *Automation in Construction, 122*, Article 103517. https://doi.org/10.1016/j.autcon.2020.103517

Rueda-Castañeda, J. E., Gallego-Gómez, N., Estanling-Cárdenas, E., Tello, J. S., & García-Pineda, V. (2024). Identificación de variables relacionadas a la seguridad informática desde la aplicación de la tecnología blockchain. *Revista Politécnica, 20*(40), 9–29. https://doi.org/10.33571/rpolitec.v20n40a1

Sharma, N., & Dhiman, P. (2025). A survey on IoT security: Challenges and their mitigation strategies in cloud-centric architectures. Cluster Computing. Advance online publication. https://doi.org/10.1007/s10586-025-05208-0

Vaca, P. A., & Dulce-Villarreal, E. R. (2024). Blockchain para asegurar la integridad y trazabilidad en la cadena de custodia de evidencia digital en informática forense: *Un estudio de mapeo sistemático. TecnoLógicas, 27*(60), Article e3049. https://doi.org/10.22430/22565337.3049

Wazid, M., Mittal, S., Das, A. K., Islam, S. H., & Alenezi, M. (2025). Designing secure blockchain-based authentication and authorization mechanisms for smart city e-health systems. *Journal of Systems Architecture, 156*, Article 103365. https://doi.org/10.1016/j.sysarc.2025.103365

Zhang, Q., Cheng, L., & Boutaba, R. (2010). Computación en la nube: Estado del arte y desafíos de la investigación. *Journal of Internet Services and Applications, 1*(1), 7–18. https://doi.org/10.1007/s13174-010-0007-6

Zhang, X., Antwi-Afari, M. F., Zhang, Y., & Xing, X. (2024). The impact of artificial intelligence on organizational justice and project performance: A systematic literature and science mapping review. *Buildings, 14*(1), Article 259. https://doi.org/10.3390/buildings14010259

e-ISNN 2550-6730 | **Vol. 9, Núm. 2 (152-163): Julio-Diciembre, 2025** | **DOI: 10.33936/isrtic.v9i2.7687**