Implementación de inteligencia artificial para la detección y respuesta a incidentes de ciberseguridad en organizaciones colombianas

Luis Miguel Araque Chia

Asesor

Yenny Stella Nuñez Alvarez

Universidad Nacional Abierta y a Distancia – UNAD

Escuela de Ciencias Básicas, Tecnología e Ingeniería – ECBTI

Especialización en Seguridad Informática

2026

Dedicatoria

A mi madre, Miryam Lupe Chia Arismendy, por su amor incondicional, su apoyo emocional en los momentos de mayor dificultad y por enseñarme con su ejemplo que la constancia y la fe hacen posible lo que parece imposible.

A mi padre, Hernando Alonso Araque Amaya, por su esfuerzo incansable, su respaldo económico y su confianza en mis capacidades, que han sido un motor permanente para alcanzar esta meta.

A mi hermana, Mary Luz Araque Chia, por su compañía, su paciencia y sus palabras de aliento que me recordaron la importancia de no rendirme a pesar de los obstáculos.

A mi sobrina, Valeria Silva Araque, cuya alegría y cariño renovaron mi motivación y me recordaron que este logro también es un legado para las nuevas generaciones de mi familia.

A toda mi familia, por ser el soporte fundamental que ha sustentado cada paso de este proceso académico y personal.

Agradecimientos

Agradezco de manera especial a mi directora de proyecto de grado, Yenny Stella Núñez Álvarez, por su orientación constante, sus observaciones rigurosas y su acompañamiento paciente en cada etapa de este trabajo, los cuales fueron fundamentales para consolidar los resultados aquí presentados.

Extiendo mi agradecimiento a mi director de curso, Christian Hernan Obando Ibarra, por su guía académica, su disposición para resolver inquietudes y por el seguimiento realizado al proceso formativo durante el desarrollo de la especialización.

A los docentes de cada curso de la especialización, por compartir sus conocimientos, experiencias y exigencia académica, que contribuyeron de manera significativa al fortalecimiento de mis competencias profesionales y a la construcción del presente trabajo.

A la Universidad Nacional Abierta y a Distancia, por brindar los recursos, materiales de estudio, espacios virtuales y facilidades necesarias para adelantar este proceso de formación, permitiendo compatibilizar los diferentes ámbitos de mi vida con el logro de esta meta académica.

A mis compañeros de estudio, por el apoyo mutuo, el intercambio de ideas y el trabajo colaborativo que enriquecieron este camino y permitieron afrontar de mejor manera los retos propios de la formación a distancia.

A mi familia, por su apoyo permanente, su comprensión ante las ausencias y su confianza inquebrantable en mis capacidades, pilares indispensables para culminar este proyecto de grado.

Resumen

Esta monografía analiza cómo la implementación de inteligencia artificial puede mejorar la detección y respuesta a incidentes de ciberseguridad en organizaciones colombianas, frente al aumento de ataques como phishing, malware y ransomware y a las limitaciones de los enfoques manuales y reactivos. Se parte de un marco teórico que integra los conceptos de inteligencia artificial, ciberseguridad y gestión de incidentes, destacando técnicas como el machine learning y el deep learning aplicadas a la identificación de patrones anómalos y a la automatización de respuestas. El objetivo general es analizar el impacto de estas tecnologías en la gestión de incidentes en el contexto colombiano, mediante una revisión documental de literatura científica, tesis e informes de los últimos cinco años, seleccionados por su pertinencia y calidad académica, organizados en una matriz de revisión y analizados de forma crítica. Como resultados se espera obtener una síntesis de amenazas relevantes, principales aplicaciones, beneficios y limitaciones de la inteligencia artificial en ciberseguridad, así como un conjunto de recomendaciones teóricas y estrategias generales que orienten a organizaciones y tomadores de decisión en la integración responsable de estas soluciones.

Palabras clave: Inteligencia artificial, ciberseguridad, gestión de incidentes, análisis documental, contexto colombiano.

Abstract

This monograph analyzes how the implementation of artificial intelligence can improve the detection and response to cybersecurity incidents in Colombian organizations, in the face of the increase in attacks such as phishing, malware and ransomware and the limitations of manual and reactive approaches. It is based on a theoretical framework that integrates the concepts of artificial intelligence, cybersecurity and incident management, highlighting techniques such as machine learning and deep learning applied to the identification of anomalous patterns and the automation of responses. The general objective is to analyze the impact of these technologies on incident management in the Colombian context, through a documentary review of scientific literature, theses and reports from the last five years, selected for their relevance and academic quality, organized in a review matrix and critically analyzed. The expected results include a synthesis of relevant threats, main applications, benefits and limitations of artificial intelligence in cybersecurity, as well as a set of theoretical recommendations and general strategies to guide organizations and decision-makers in the responsible integration of these solutions.

Keywords: Artificial intelligence, cybersecurity, incident management, documentary analysis, Colombian context.

Tabla de Contenido

Lista de Tablas

Introducción

La creciente digitalización de los procesos empresariales ha convertido a las organizaciones colombianas en un blanco frecuente de ciberataques cada vez más sofisticados, que ponen en riesgo la continuidad operativa, la información crítica y la confianza de clientes y usuarios. Frente a este escenario, los enfoques tradicionales de ciberseguridad, basados en tareas manuales y esquemas principalmente reactivos, resultan insuficientes para anticipar, detectar y gestionar de forma oportuna incidentes como el phishing, el malware o el ransomware. En este contexto, la inteligencia artificial se posiciona como una herramienta estratégica para fortalecer la detección y respuesta a incidentes de ciberseguridad, al facilitar el análisis automatizado de grandes volúmenes de datos, la identificación de patrones anómalos y la ejecución de acciones de contención en tiempos más cortos.

Esta monografía se inscribe en esa problemática y desarrolla un análisis documental sobre la implementación de inteligencia artificial para la detección y respuesta a incidentes de ciberseguridad en organizaciones colombianas. A partir de la revisión crítica de literatura científica reciente, tesis e informes técnicos, se busca comprender el estado actual de estas aplicaciones, sus principales beneficios y limitaciones, así como los retos particulares del contexto nacional, con el fin de aportar un marco teórico que oriente futuras decisiones y estudios en el campo de la seguridad informática.

Planteamiento del Problema

En la actualidad, las organizaciones colombianas enfrentan una creciente ola de ciberataques caracterizados por su sofisticación y recurrencia, que comprometen datos sensibles, procesos críticos y la confianza de clientes y usuarios. Incidentes como el phishing, el malware, el ransomware y la suplantación de identidad afectan tanto a entidades públicas como privadas y se han consolidado como una amenaza constante para la continuidad del negocio. A pesar de los avances en políticas y marcos de referencia en ciberseguridad, muchos esquemas de protección siguen siendo principalmente manuales y reactivos, lo que dificulta la detección temprana de anomalías y retrasa la respuesta frente a eventos de seguridad.

Esta situación se agrava por brechas en la adopción de tecnologías avanzadas, la limitada capacitación del talento humano y la falta de lineamientos claros para incorporar soluciones basadas en inteligencia artificial en los procesos de gestión de incidentes. Diversos reportes señalan a Colombia como uno de los principales blancos de ciberataques en la región, con impacto particular en sectores como el financiero, industrial y de servicios, lo que evidencia la necesidad de fortalecer las capacidades de defensa digital de las organizaciones. Sin embargo, aunque a nivel internacional existen múltiples estudios sobre el uso de inteligencia artificial en ciberseguridad, en el contexto colombiano aún son escasas las investigaciones que sistematicen de forma específica cómo estas tecnologías pueden apoyar la detección y respuesta a incidentes en organizaciones locales.

De esta manera, surge la necesidad de analizar, desde una perspectiva documental, el impacto de la implementación de inteligencia artificial en la detección y respuesta a incidentes de ciberseguridad en organizaciones colombianas, con el fin de comprender sus aplicaciones, beneficios y limitaciones, así como los retos que supone su adopción en el entorno nacional.

Pregunta del Problema

¿De qué manera la implementación de inteligencia artificial incide en la detección y respuesta a incidentes de ciberseguridad en las organizaciones colombianas, en términos de beneficios reportados, cambios en los tiempos de respuesta, limitaciones técnicas y organizacionales, y retos de adopción identificados en la literatura reciente?

Justificación

El incremento sostenido de los ciberataques y la complejidad de las amenazas digitales han puesto a prueba la capacidad de respuesta de las organizaciones colombianas, evidenciando las limitaciones de los modelos tradicionales de gestión de la ciberseguridad basados en tareas manuales y acciones principalmente reactivas. En este escenario, la inteligencia artificial se ha posicionado como una alternativa estratégica para modernizar la defensa digital, al permitir el análisis automatizado de grandes volúmenes de información, la identificación temprana de comportamientos anómalos y la activación de mecanismos de respuesta más ágiles.

Analizar de manera sistemática cómo estas tecnologías pueden aplicarse a la detección y respuesta a incidentes de ciberseguridad en organizaciones colombianas resulta pertinente porque aporta elementos para cerrar brechas entre la oferta tecnológica disponible y las necesidades reales del contexto nacional. El enfoque de revisión documental permite integrar experiencias y buenas prácticas reportadas en la literatura científica reciente, contrastándolas con los retos locales en aspectos como infraestructura, talento humano, inversión y marcos regulatorios, lo que favorece una comprensión más completa de las oportunidades y riesgos asociados al uso de inteligencia artificial en este campo.

Desde la perspectiva académica, esta monografía contribuye a organizar y actualizar el conocimiento sobre el uso de inteligencia artificial en la gestión de incidentes de ciberseguridad, ofreciendo un marco de referencia que puede ser aprovechado por estudiantes, docentes e investigadores interesados en profundizar en la temática. En el ámbito profesional, los resultados del estudio buscan servir de insumo para responsables de seguridad de la información y tomadores de decisión, al proporcionar recomendaciones teóricas y estrategias generales que

apoyen una adopción responsable y contextualizada de soluciones basadas en inteligencia artificial en las organizaciones colombianas.

Objetivos

Objetivo General

Analizar, a partir de una revisión documental crítica, cómo la implementación de inteligencia artificial incide en la detección y respuesta a incidentes de ciberseguridad en organizaciones colombianas, identificando beneficios reportados, cambios en los tiempos de respuesta, limitaciones técnicas y organizacionales, y retos de adopción en el contexto nacional.

Objetivos Específicos

Identificar, mediante una matriz de revisión documental, las principales amenazas y vulnerabilidades de ciberseguridad que afectan a las organizaciones colombianas, clasificándolas por tipo de incidente (phishing, malware, ransomware, entre otros) y sector económico.

Examinar las aplicaciones, técnicas y modelos de inteligencia artificial utilizados para la detección y respuesta a incidentes de ciberseguridad, organizándolos en categorías según tipo de técnica (machine learning, deep learning, sistemas expertos), tipo de incidente atendido y nivel de madurez organizacional reportado en la literatura.

Sintetizar los beneficios, limitaciones y retos de adopción de la inteligencia artificial en la gestión de incidentes de ciberseguridad en el contexto colombiano, formulando recomendaciones teóricas y estrategias generales basadas en las categorías de análisis construidas en la matriz documental.

Marco Referencial

Antecedentes

En los últimos años se ha consolidado un cuerpo creciente de investigaciones que analizan el uso de la inteligencia artificial para apoyar la detección y respuesta a incidentes de ciberseguridad en distintos tipos de organizaciones y sectores productivos. Diversos estudios reportan implementaciones de técnicas de machine learning, deep learning y sistemas expertos orientadas a identificar patrones anómalos en redes empresariales, automatizar procesos de correlación de eventos y reducir los tiempos de reacción frente a ataques dirigidos. Estas experiencias, desarrolladas principalmente en contextos internacionales, muestran mejoras en la capacidad de monitoreo continuo y en la priorización de alertas, así como una disminución de la dependencia exclusiva del análisis manual por parte de los equipos de seguridad.

De manera paralela, se han publicado trabajos que abordan los retos asociados a la adopción de soluciones basadas en inteligencia artificial, entre ellos la necesidad de contar con datos de calidad para el entrenamiento de modelos, las implicaciones éticas y legales de la toma de decisiones automatizada y los riesgos vinculados a sesgos algorítmicos y fallos de interpretación. En América Latina y Colombia comienzan a aparecer estudios y tesis de posgrado que exploran la aplicación de modelos de IA en empresas, infraestructuras críticas e instituciones financieras, pero aún se evidencia una dispersión de enfoques y una limitada sistematización de resultados. Esta situación hace necesario organizar y analizar de manera integral los aportes existentes, identificando tendencias, vacíos de conocimiento y buenas prácticas que orienten la integración de la inteligencia artificial en la gestión de incidentes de ciberseguridad en el contexto colombiano.

Marco Conceptual

El punto de partida de esta monografía es la inteligencia artificial, entendida como el campo de la informática que desarrolla sistemas capaces de aprender a partir de datos, reconocer patrones y tomar decisiones o recomendaciones de manera autónoma en situaciones específicas. En el ámbito de la ciberseguridad, la inteligencia artificial incorpora técnicas como el machine learning y el deep learning, que permiten entrenar modelos para distinguir comportamientos normales y anómalos en redes, sistemas y aplicaciones, apoyando la detección temprana de actividades maliciosas y la automatización de respuestas frente a incidentes.

La ciberseguridad se concibe como el conjunto de políticas, procesos, tecnologías y prácticas orientadas a proteger la confidencialidad, integridad y disponibilidad de la información y de los activos tecnológicos de una organización. Esta disciplina abarca desde la gestión de riesgos y la implementación de controles técnicos hasta la formación de los usuarios y la definición de marcos normativos que regulen el uso seguro de las tecnologías digitales. Dentro de la ciberseguridad, la gestión de incidentes comprende las fases de detección, análisis, contención, erradicación y recuperación, que buscan minimizar el impacto de eventos que comprometen la seguridad de los sistemas.

La integración de inteligencia artificial en la gestión de incidentes implica combinar estas capacidades de análisis automatizado con procesos y herramientas tradicionales de seguridad, de manera que los modelos puedan filtrar grandes volúmenes de alertas, priorizar riesgos y sugerir o ejecutar acciones de respuesta en tiempos más reducidos. Este enfoque plantea oportunidades importantes para mejorar la ciberresiliencia de las organizaciones colombianas, pero también abre interrogantes sobre la calidad de los datos utilizados, la transparencia de los algoritmos y las

responsabilidades asociadas a las decisiones automatizadas, aspectos que serán abordados en el desarrollo de la monografía.

Marco Teórico

El desarrollo de la inteligencia artificial aplicada a la ciberseguridad se apoya en modelos de aprendizaje automático capaces de analizar grandes volúmenes de datos generados por redes, sistemas y aplicaciones, con el propósito de identificar patrones que indiquen posibles incidentes de seguridad. Entre las técnicas más empleadas se encuentran los algoritmos de clasificación supervisada, que aprenden a distinguir entre tráfico legítimo y tráfico malicioso a partir de ejemplos etiquetados, y los enfoques no supervisados, que buscan detectar anomalías sin conocer de antemano todas las formas de ataque. Estas aproximaciones permiten construir sistemas de detección de intrusos más adaptativos, capaces de actualizar sus modelos a medida que evolucionan las amenazas.

El uso de redes neuronales profundas ha ampliado estas capacidades al permitir el análisis de estructuras de datos más complejas, como registros de eventos, flujos de red o trazas de comportamiento de usuarios, extrayendo representaciones que facilitan la identificación de patrones sutiles asociados a ciberataques avanzados. A su vez, los sistemas expertos y los motores de correlación incorporan reglas y conocimiento estructurado para apoyar la toma de decisiones durante la gestión de incidentes, por ejemplo, al sugerir acciones de contención o priorizar alertas según su criticidad. En conjunto, estos modelos conforman un ecosistema tecnológico que busca complementar, y no reemplazar, el trabajo de los analistas de seguridad, al reducir la carga de tareas repetitivas y mejorar la capacidad de respuesta frente a incidentes en entornos cada vez más complejos.

Marco Legal

El desarrollo de esta monografía se enmarca en la normativa colombiana relacionada con la protección de la información y la ciberseguridad, que establece lineamientos para la gestión de riesgos digitales en organizaciones públicas y privadas. Entre los referentes más relevantes se encuentran las leyes que regulan el habeas data y la protección de datos personales, las disposiciones sobre delitos informáticos y los marcos técnicos que orientan la implementación de controles de seguridad en infraestructuras críticas y servicios digitales. Estos instrumentos jurídicos buscan garantizar la confidencialidad, integridad y disponibilidad de la información, así como la responsabilidad de las entidades frente a incidentes que afecten a sus usuarios.

Adicionalmente, el país cuenta con estrategias nacionales de ciberseguridad y lineamientos sectoriales que promueven el fortalecimiento de la gestión de incidentes y la adopción de buenas prácticas en seguridad de la información. En este contexto regulatorio, la incorporación de soluciones basadas en inteligencia artificial para la detección y respuesta a incidentes debe analizarse no solo desde su eficacia técnica, sino también a la luz de principios como la proporcionalidad, la transparencia y el respeto por la privacidad de los datos. De esta manera, el marco legal colombiano se convierte en un referente clave para valorar las oportunidades y límites de la implementación de inteligencia artificial en los procesos de ciberseguridad organizacional.

Marco Contextual

El contexto colombiano se caracteriza por una acelerada adopción de tecnologías digitales en sectores como el financiero, el industrial, los servicios y la administración pública, lo que ha generado una creciente dependencia de los sistemas de información para la operación diaria de las organizaciones. Este proceso ha venido acompañado de un aumento significativo en

la cantidad y sofisticación de los ciberataques, que aprovechan vulnerabilidades técnicas, brechas en los procesos de seguridad y debilidades en la cultura digital de los usuarios. Reportes recientes señalan a Colombia como uno de los países de la región con mayor volumen de incidentes de ciberseguridad, lo que evidencia la necesidad de fortalecer la capacidad de prevención, detección y respuesta en las entidades públicas y privadas.

En respuesta a este panorama, se han impulsado iniciativas de política pública, marcos de referencia y proyectos de capacitación orientados a mejorar la ciberresiliencia nacional, pero persisten desafíos relacionados con la inversión en infraestructura, la disponibilidad de talento especializado y la incorporación efectiva de tecnologías avanzadas en los procesos de seguridad. La exploración de soluciones basadas en inteligencia artificial para la gestión de incidentes se inscribe en este escenario, como una oportunidad para que las organizaciones colombianas modernicen sus esquemas de defensa, aprovechen mejor la información generada por sus sistemas y tomen decisiones más informadas frente a los riesgos digitales. Este marco contextual sirve de base para analizar, a lo largo de la monografía, qué tan alineadas están las propuestas tecnológicas con las realidades y necesidades específicas del país.

Diseño Metodológico

Esta investigación se desarrolla mediante una metodología de revisión documental y análisis crítico, orientada a recopilar, organizar y sintetizar información relevante sobre la aplicación de inteligencia artificial en la detección y respuesta a incidentes de ciberseguridad en organizaciones colombianas. El enfoque es de carácter cualitativo, con apoyo de elementos cuantitativos descriptivos (frecuencias y agrupaciones de categorías) que permiten identificar patrones en las amenazas, las técnicas de IA utilizadas y los contextos organizacionales donde se implementan.

Fuentes de Información y Criterios de Selección

Las fuentes de información se obtienen principalmente de bases de datos académicas especializadas, repositorios institucionales y publicaciones de entidades reconocidas en ciberseguridad, considerando artículos científicos, tesis de posgrado, informes técnicos y documentos institucionales relacionados con inteligencia artificial y gestión de incidentes. Para garantizar la pertinencia y el rigor de la revisión, se aplican los siguientes criterios de inclusión y exclusión:

Criterios de Inclusión

Publicaciones realizadas entre 2019 y 2025, salvo documentos de referencia obligatoria por su relevancia en el campo.

Documentos en español o inglés.

Estudios que aborden la detección y/o respuesta a incidentes de ciberseguridad en organizaciones, con énfasis en aplicaciones de inteligencia artificial.

Trabajos que reporten al menos uno de los siguientes elementos: tipo de incidente atendido, técnica de IA utilizada, contexto organizacional o resultados obtenidos.

Criterios de Exclusión

Documentos sin acceso al texto completo o que no presenten información metodológica mínima.

Publicaciones centradas exclusivamente en aspectos jurídicos, filosóficos o éticos de la IA y la ciberseguridad, sin relación operativa con la gestión de incidentes.

Fuentes de divulgación sin respaldo académico o institucional verificable.

Variables y Categorías de Análisis

La información extraída de las fuentes se organiza en una matriz de revisión bibliográfica que permite sistematizar los aportes de cada estudio de acuerdo con los objetivos de la investigación. Para ello se definen las siguientes variables y categorías de análisis:

Tipo de incidente de ciberseguridad: phishing, malware, ransomware, intrusiones a redes, fuga de datos, ataques a la cadena de suministro, entre otros.

Sector económico: financiero, industrial, servicios, telecomunicaciones, sector público, pymes, infraestructuras críticas u otros sectores relevantes.

Técnica de inteligencia artificial: machine learning supervisado, aprendizaje no supervisado, deep learning, sistemas expertos, motores de correlación basados en reglas y otras aproximaciones híbridas.

Fase del ciclo de gestión de incidentes: detección, análisis, contención, erradicación y recuperación.

Nivel de madurez organizacional de la solución: pruebas en laboratorio, proyectos piloto, implementación parcial, implementación en ambientes de producción.

Tipo de resultados reportados: cambios en los tiempos de detección o respuesta, mejoras en la tasa de detección, reducción de falsos positivos, impacto en la ciberresiliencia organizacional, entre otros.

Estas variables permiten articular el análisis de las amenazas y vulnerabilidades relevantes (objetivo específico 1), de las aplicaciones y modelos de IA utilizados en la detección y respuesta a incidentes (objetivo específico 2) y de los beneficios, limitaciones y retos de adopción identificados en el contexto colombiano (objetivo específico 3).

Procedimiento de Análisis y Síntesis

El procedimiento metodológico se desarrolla en tres momentos principales. En primer lugar, se realiza la búsqueda y selección de documentos aplicando los criterios definidos, registrando cada fuente en la matriz de revisión con información sobre sus objetivos, enfoque metodológico, contexto, tipo de incidente, técnica de IA, sector económico y principales hallazgos. En segundo lugar, se lleva a cabo un proceso de codificación categorial, asignando a cada estudio las categorías correspondientes a las variables de análisis (por ejemplo, combinación de tipo de incidente, técnica de IA, fase del ciclo de incidentes y nivel de madurez organizacional).

En tercer lugar, se procede a la agrupación y síntesis de la información, identificando patrones y tendencias a partir de la frecuencia y combinación de categorías. Para el primer objetivo específico, se jerarquizan las amenazas y vulnerabilidades según su recurrencia y los sectores económicos más afectados, a partir de la matriz de amenazas y sectores construida en el Capítulo 1. Para el segundo objetivo, se examinan las aplicaciones, técnicas y modelos de inteligencia artificial reportados en la literatura, organizándolos según el tipo de técnica, el tipo de incidente que atienden y el nivel de madurez organizacional. Finalmente, para el tercer

objetivo, se sintetizan los beneficios, limitaciones y retos de adopción de estas soluciones, destacando aquellos elementos que se reiteran en diferentes estudios y contextos y que resultan especialmente pertinentes para las organizaciones colombianas.

Contexto de los Ciberataques e Inteligencia Artificial en Colombia

Panorama de Ciberataques en Organizaciones Colombianas

En la actualidad, las organizaciones colombianas enfrentan una creciente ola de ciberataques caracterizados por su sofisticación y recurrencia, que comprometen datos sensibles, procesos críticos y la confianza de clientes y usuarios. Incidentes como el phishing, el malware, el ransomware y la suplantación de identidad afectan tanto a entidades públicas como privadas y se han consolidado como una amenaza constante para la continuidad del negocio. A pesar de los avances en políticas y marcos de referencia en ciberseguridad, muchos esquemas de protección siguen siendo principalmente manuales y reactivos, lo que dificulta la detección temprana de anomalías y retrasa la respuesta frente a eventos de seguridad. Esta situación se agrava por brechas en la adopción de tecnologías avanzadas, la limitada capacitación del talento humano y la falta de lineamientos claros para incorporar soluciones más automatizadas en los procesos de gestión de incidentes; en este contexto, resulta fundamental describir las principales amenazas y vulnerabilidades de ciberseguridad que enfrentan las organizaciones colombianas, así como los tipos de incidentes más frecuentes y los sectores económicos donde se concentran.

Estos ataques no solo comprometen la confidencialidad, integridad y disponibilidad de la información, sino que también impactan directamente en la continuidad del negocio, al generar indisponibilidad de servicios, pérdidas económicas y afectaciones reputacionales que minan la confianza de clientes y usuarios en la capacidad de las organizaciones para proteger sus activos digitales. Diversos reportes y casos recientes han mostrado cómo incidentes dirigidos a infraestructuras de TI, sistemas de pago, servicios en la nube y plataformas de atención al ciudadano pueden paralizar procesos misionales, obligar a suspender temporalmente operaciones

y desencadenar investigaciones regulatorias o sanciones, con especial impacto en sectores como el financiero, industrial, de servicios, telecomunicaciones y entidades públicas.

A pesar de la existencia de políticas nacionales y marcos de referencia en ciberseguridad, muchos esquemas de protección implementados en el país siguen basándose en actividades manuales, controles aislados y modelos fundamentalmente reactivos, que se activan solo cuando el incidente ya se ha materializado. La velocidad con la que evolucionan las amenazas, la amplitud de la superficie de ataque y el volumen de eventos que deben analizar los equipos de seguridad hacen que estas prácticas resulten insuficientes para mantener un nivel adecuado de vigilancia y detección temprana, lo que evidencia la necesidad de fortalecer de manera prioritaria las estrategias de prevención, monitoreo continuo y respuesta ante incidentes en el contexto colombiano.

Los estudios recientes sobre ciberseguridad en el contexto latinoamericano y en pequeñas y medianas empresas confirman esta tendencia de incremento y diversificación de las amenazas, al destacar el papel del factor humano, la evolución de los vectores de ataque y las brechas que persisten en las prácticas de seguridad de muchas organizaciones (Gómez, Ortiz, Vásquez, Marín, & Martínez, 2025; Galvez & Revinova, 2025). Estos trabajos coinciden en la necesidad de robustecer los mecanismos de monitoreo y respuesta, y señalan a la inteligencia artificial como una alternativa con alto potencial para apoyar la detección temprana de incidentes y la gestión de riesgos en entornos complejos, especialmente en escenarios donde los recursos humanos y tecnológicos son limitados (Bilbao-Arechabala, Atutxa, & Del Ser, 2025; Cuya-Chamilco, Saavedra-Villar, Mescua-Ampuero, Alvarado-Flores, & Cáceres-García, 2025).

Limitaciones de los Enfoques Manuales y Reactivos de Ciberseguridad

A pesar de la creciente visibilidad de los riesgos digitales, muchas organizaciones colombianas continúan gestionando la ciberseguridad mediante enfoques predominantemente manuales y reactivos, apoyados en la experiencia de los analistas y en procedimientos que se activan una vez el incidente ya ha tenido lugar. En la práctica, esto implica que buena parte del esfuerzo se concentra en revisar registros de eventos, atender alertas generadas por múltiples herramientas y resolver casos reportados por los usuarios, lo que demanda una alta dedicación de tiempo y dificulta mantener una vigilancia constante sobre toda la infraestructura tecnológica.

Este modelo encuentra importantes limitaciones frente al volumen, velocidad y variedad de datos que se generan en redes modernas, especialmente cuando las organizaciones dependen de servicios en la nube, aplicaciones distribuidas y conexiones remotas. La cantidad de eventos que deben analizarse supera con facilidad la capacidad de los equipos humanos, lo que se traduce en fatiga de alertas, priorización inadecuada de incidentes y riesgo de pasar por alto señales tempranas de compromiso. Además, la alta rotación de personal y las brechas en la formación especializada dificultan sostener equipos de seguridad con el nivel de conocimiento requerido para enfrentar amenazas cada vez más sofisticadas.

Desde la perspectiva de la literatura reciente, diversos estudios señalan que los modelos centrados en tareas manuales y reglas estáticas resultan insuficientes para detectar patrones anómalos complejos, correlacionar eventos dispersos en diferentes sistemas y responder con la rapidez que exigen los ataques actuales (Amarasinghe, Jarray, & Karmouch, 2021; Aly, 2017). Estos trabajos destacan que la ausencia de capacidades avanzadas de análisis y automatización incrementa el tiempo de detección y contención de incidentes, y limita la posibilidad de anticiparse a comportamientos maliciosos basados en el uso adaptativo de técnicas como el

malware polimórfico o los ataques distribuidos (Dudczyk, Sergiel, & Krygier, 2025; Tang, Zhao, Chen, & Zhou, 2024). En este escenario, la incorporación de inteligencia artificial se plantea como una alternativa para complementar el trabajo de los analistas, reducir la dependencia de procesos manuales y avanzar hacia esquemas de gestión de incidentes más proactivos y orientados a la detección temprana.

Rol de la Inteligencia Artificial en la Gestión de Incidentes de Ciberseguridad

La inteligencia artificial se ha consolidado como una de las tecnologías con mayor potencial para transformar la forma en que las organizaciones detectan y gestionan incidentes de ciberseguridad, al permitir el análisis automatizado de grandes volúmenes de datos procedentes de redes, sistemas y aplicaciones. A diferencia de los enfoques basados únicamente en reglas estáticas, los modelos de aprendizaje automático pueden identificar patrones anómalos en el comportamiento del tráfico, de los usuarios o de los dispositivos, incluso cuando no coinciden exactamente con firmas previamente conocidas, lo que amplía la capacidad de detección frente a amenazas emergentes.

En el ámbito de la detección, diversas propuestas se centran en el uso de algoritmos de clasificación y de detección de anomalías para filtrar alertas, priorizar aquellos eventos con mayor probabilidad de ser maliciosos y reducir la carga de trabajo de los analistas (Amarasinghe, Jarray, & Karmouch, 2021; Dudczyk, Sergiel, & Krygier, 2025). En la fase de respuesta, se exploran mecanismos de automatización y orquestación que permiten ejecutar acciones de contención, como el aislamiento de equipos comprometidos, el bloqueo dinámico de direcciones o la reconfiguración de rutas de red, disminuyendo los tiempos de reacción ante incidentes (Tang, Zhao, Chen, & Zhou, 2024; Gillgallon, Almutairi, Bergami, & Morgan, 2025).

En el contexto colombiano, el potencial de estas tecnologías adquiere especial relevancia debido a las brechas de talento humano, las limitaciones presupuestales y la creciente complejidad de las infraestructuras tecnológicas de las organizaciones. La posibilidad de combinar el criterio de los equipos de seguridad con capacidades de análisis automatizado y respuesta asistida por inteligencia artificial se vislumbra como una oportunidad para fortalecer la detección y gestión de incidentes, siempre que su adopción se acompañe de lineamientos claros, procesos de evaluación de riesgos y estrategias de formación adecuadas.

Situación de Adopción de Inteligencia Artificial en Ciberseguridad en Colombia

En Colombia, el interés por incorporar soluciones basadas en inteligencia artificial a los procesos de ciberseguridad ha crecido en los últimos años, especialmente en sectores como el financiero, las telecomunicaciones y algunos organismos del Estado que han impulsado proyectos piloto o iniciativas de modernización de infraestructura. Sin embargo, la adopción de estas tecnologías sigue siendo desigual y se concentra en organizaciones con mayores capacidades técnicas y presupuestales, mientras que muchas empresas medianas y pequeñas continúan dependiendo de herramientas tradicionales y de esquemas de monitoreo principalmente manuales.

Entre los factores que explican esta brecha se encuentran las limitaciones de infraestructura tecnológica, los costos asociados a la adquisición y operación de soluciones avanzadas, la escasez de talento especializado en ciencia de datos y seguridad, y la ausencia de lineamientos claros sobre cómo integrar la inteligencia artificial a los procesos de gestión de incidentes sin comprometer la privacidad ni el cumplimiento normativo. La literatura regional señala que, aunque se reconoce el potencial de estas herramientas para mejorar la detección y respuesta a incidentes, persisten dudas sobre su madurez, su capacidad de adaptación a contextos

locales y los riesgos derivados de una posible dependencia excesiva de sistemas automatizados en entornos con recursos limitados (Gómez, Ortiz, Vásquez, Marín, & Martínez, 2025; Cuya-Chamilco, Saavedra-Villar, Mescua-Ampuero, Alvarado-Flores, & Cáceres-García, 2025).

En este escenario, resulta pertinente avanzar en estudios de carácter documental que permitan sistematizar las experiencias reportadas a nivel internacional y regional, identificar buenas prácticas y analizar los retos técnicos y organizacionales que implica la implementación de inteligencia artificial en la ciberseguridad de las organizaciones colombianas. Este tipo de análisis contribuye a cerrar la brecha entre la oferta tecnológica y las necesidades reales del país, y proporciona insumos para la formulación de políticas, estrategias y proyectos que orienten una adopción más informada y responsable de estas soluciones.

Síntesis y Relación con la Pregunta de Investigación

El panorama descrito en este capítulo muestra que las organizaciones colombianas enfrentan un contexto de ciberamenazas cada vez más complejo, marcado por la recurrencia de ataques como el phishing, el malware y el ransomware, y por su impacto directo en la continuidad del negocio y la confianza de clientes y usuarios. Al mismo tiempo, se evidencia que muchos esquemas de protección siguen apoyándose en enfoques manuales y reactivos, con capacidades limitadas para procesar el volumen de información generado por las infraestructuras actuales y para anticiparse a patrones de ataque más sofisticados.

Frente a este escenario, la literatura revisada plantea a la inteligencia artificial como una alternativa con alto potencial para fortalecer la detección y respuesta a incidentes, al aportar mecanismos de análisis automatizado, priorización de alertas y orquestación de acciones de contención (Amarasinghe, Jarray, & Karmouch, 2021; Tang, Zhao, Chen, & Zhou, 2024). No obstante, en el contexto colombiano persisten brechas en infraestructura, talento humano y

lineamientos de adopción, y son todavía escasos los estudios que sistematizan de manera específica cómo estas tecnologías se están aplicando o podrían aplicarse en las organizaciones del país (Gómez, Ortiz, Vásquez, Marín, & Martínez, 2025; Cuya-Chamilco, Saavedra-Villar, Mescua-Ampuero, Alvarado-Flores, & Cáceres-García, 2025).

En consecuencia, se justifica la realización de la monografía de revisión documental que analiza el impacto de la implementación de inteligencia artificial en la detección y respuesta a incidentes de ciberseguridad en organizaciones colombianas, en términos de beneficios reportados, cambios en los tiempos de respuesta, limitaciones técnicas y organizacionales, y retos de adopción identificados en la literatura reciente. Esta síntesis orienta y da coherencia a la pregunta de investigación planteada y al objetivo general del estudio, y establece las bases conceptuales sobre las cuales se desarrollarán los capítulos posteriores dedicados al análisis detallado de las experiencias y propuestas encontradas en las fuentes seleccionadas.

Matriz de Amenazas y Sectores Afectados

Con el fin de dar cumplimiento al primer objetivo específico de la monografía, se elaboró una matriz de revisión documental en la que se sistematizan las principales amenazas y vulnerabilidades de ciberseguridad identificadas en la literatura reciente, clasificándolas por tipo de incidente y sector económico afectado. Esta matriz se construyó a partir de artículos científicos, informes técnicos y tesis relacionadas con el contexto colombiano y latinoamericano, permitiendo reconocer patrones frecuentes de ataque en organizaciones de distintos sectores, así como las debilidades más habituales que son aprovechadas por los actores maliciosos. De este modo, la matriz sirve como insumo base para el análisis desarrollado en los capítulos posteriores sobre el impacto de la inteligencia artificial en la detección y respuesta a incidentes.

Tabla 1

Matriz de Amenazas y Sectores Afectados

Autor(es) y año	País / región del estudio	Tipo de incidente principal	Sector económico afectado	Principales vulnerabilidades identificadas	Fuente / tipo de documento
ENISA, 2022	Unión Europea / internacional	Malware, ransomware, ataques a la cadena de suministro	Infraestructuras críticas, servicios digitales, sector empresarial	Falta de segmentación de redes, gestión deficiente de parches, exposición de servicios críticos a internet	Informe técnico de agencia de ciberseguridad
Bancolombia, 2025	Colombia	Phishing, robo de credenciales, fraude transaccional	Sector financiero	Ingeniería social dirigida a clientes, uso de enlaces maliciosos, débil cultura de verificación de identidad	Informe corporativo / reporte de ciberseguridad
Impacto TIC, 2025	Colombia / América Latina	Ransomware, ataques a servicios en la nube	Sector industrial y de servicios	Equipos desactualizados, configuraciones inseguras en nube, falta de copias de seguridad y planes de recuperación	Reporte periodístico especializado en TIC
Ámbito Jurídico, 2024	Colombia	Incidentes de fuga de datos,	Sector público y regulado	Controles de acceso inadecuados,	Análisis jurídico /

Autor(es) y año	País / región del estudio	Tipo de incidente principal	Sector económico afectado	Principales vulnerabilidades identificadas	Fuente / tipo de documento
		intrusiones a sistemas		gestión deficiente de logs, ausencia de monitoreo continuo	reporte de casos
Herrera Guzmán, 2025	América Latina	Detección de intrusos basada en IA (tráfico malicioso)	Empresas de servicios TI y telecomunicaciones	Débil monitoreo de tráfico, dependencia de reglas estáticas, baja automatización de respuestas	Artículo científico sobre IA y ciberseguridad
Jiménez Amoroso, 2024	Internacional	Anomalías en redes y sistemas empresariales	Empresas de diversos sectores	Falta de correlación de eventos, alta carga manual de análisis, ausencia de modelos predictivos	Artículo científico / revisión de literatura
Martínez-Castro, 2024	Internacional	Intrusiones en redes corporativas	Sector empresarial	Políticas de seguridad poco actualizadas, insuficiente control de accesos y monitoreo de redes	Artículo científico sobre gestión de incidentes
Zapata Lizarazo, 2025	Internacional	Ataques avanzados (APTs,	Organizaciones intensivas en TI	Insuficiente análisis de logs, dificultad para	Artículo científico sobre modelos

Autor(es) y año	País / región del estudio	Tipo de incidente principal	Sector económico afectado	Principales vulnerabilidades identificadas	Fuente / tipo de documento
		comportamiento anómalo)		identificar patrones complejos, falta de herramientas basadas en IA	de IA en SIEM
Díaz Valverde & Bravo Rojas, 2023	Latinoamérica / Colombia	Incidentes generales de seguridad de la información	Organizaciones públicas y privadas	Ausencia de gestión integral de riesgos, políticas incompletas, poca capacitación de usuarios	Libro / texto académico de ciberseguridad
CCIT, 2025	Colombia	Phishing, fraude, malware, ataques a infraestructura	Sector productivo colombiano en general	Brecha en cultura de ciberseguridad, inversión limitada, débil adopción de tecnologías avanzadas	Informe institucional sobre ciberseguridad
Universidad VIU, 2025	Internacional	Amenazas potenciadas por IA (ataques automatizados)	Organizaciones digitales y servicios en línea	Falta de preparación frente a ataques basados en IA, escasa actualización de controles y marcos de referencia	Documento académico / informe técnico

Autor(es) y año	País / región del estudio	Tipo de incidente principal	Sector económico afectado	Principales vulnerabilidades identificadas	Fuente / tipo de documento
ENISA, 2022	Unión Europea / internacional	Abuso de IA para generar phishing y malware avanzado	Sector financiero, telecomunicaciones, servicios en línea	Detección insuficiente de contenidos generados por IA, filtros de correo poco sofisticados	Informe técnico de agencia de ciberseguridad

Nota. Esta tabla muestra una matriz de revisión documental de amenazas de ciberseguridad y sectores económicos afectados

A partir de la matriz de revisión documental se observa que las amenazas más recurrentes para las organizaciones se concentran en incidentes de phishing, malware y ransomware, así como en intrusiones a redes corporativas y fugas de datos, que impactan de forma especial a los sectores financiero, industrial, de servicios, telecomunicaciones y a las entidades públicas. De manera transversal, las fuentes analizadas coinciden en que estas amenazas se ven facilitadas por vulnerabilidades como la desactualización de equipos y aplicaciones, la configuración insegura de servicios en la nube, la falta de segmentación de redes, una escasa cultura de ciberseguridad en usuarios y la limitada adopción de herramientas avanzadas de monitoreo y respuesta basadas en inteligencia artificial. En conjunto, estos hallazgos permiten caracterizar el panorama de riesgos que enfrentan las organizaciones colombianas y justifican la necesidad de explorar soluciones de detección y respuesta soportadas en IA.

En conjunto, los estudios revisados permiten observar que las amenazas que se presentan con mayor recurrencia son el phishing orientado al robo de credenciales, el malware y el ransomware dirigidos al cifrado o secuestro de información, así como las intrusiones que derivan

en fuga de datos sensibles. Estos incidentes se concentran especialmente en sectores como el financiero, el industrial, los servicios y las entidades públicas, donde la dependencia de servicios digitales y la criticidad de la información incrementan el impacto operativo y reputacional de los ataques. La reiteración de vulnerabilidades como la falta de segmentación de redes, la gestión deficiente de parches, las configuraciones inseguras en la nube y la escasa cultura de ciberseguridad sugiere que, más allá de las diferencias sectoriales, persisten problemas estructurales en la forma como las organizaciones gestionan sus activos tecnológicos, lo que refuerza la necesidad de fortalecer tanto los controles técnicos como las capacidades de monitoreo y respuesta frente a estas amenazas.

Aplicaciones de la Inteligencia Artificial en la Detección y Respuesta a Incidentes de Ciberseguridad

Panorama General de Aplicaciones de Inteligencia Artificial en la Gestión de Incidentes

En la literatura reciente se observa un crecimiento sostenido de propuestas que incorporan técnicas de inteligencia artificial para apoyar la detección y respuesta a incidentes de ciberseguridad en distintos tipos de organizaciones y sectores productivos (Jiménez Amoroso, 2024; Martínez-Castro, 2024; Ortega, 2021). Estas aplicaciones se centran principalmente en el análisis automatizado de grandes volúmenes de datos generados por redes, sistemas y aplicaciones, con el propósito de identificar comportamientos anómalos, priorizar alertas y reducir los tiempos de reacción frente a ataques dirigidos (Méndez Castellanos, 2024; Smith & Brown, 2021).

Los modelos de machine learning y deep learning se emplean de manera frecuente en sistemas de detección de intrusos, análisis de tráfico de red e identificación de patrones sospechosos en registros de eventos, así como en el reconocimiento de comportamientos atípicos de usuarios (López Vidal, 2023; Zapata Lizarazo, 2025; Vargas Gómez, 2022). A su vez, los sistemas expertos y motores de correlación de eventos se utilizan para apoyar la toma de decisiones durante la gestión de incidentes, sugiriendo acciones de contención o automatizando respuestas frente a determinados tipos de eventos (Romero, 2023; Salgado García, 2024; Sánchez & Torres, 2024). De forma transversal, estos trabajos reportan beneficios como una mayor capacidad de monitoreo continuo, una mejor priorización de incidentes y una disminución de la dependencia exclusiva del análisis manual, aunque también evidencian retos relacionados con la disponibilidad de datos de calidad, la complejidad de los modelos y la necesidad de contar

con equipos especializados para su operación y mantenimiento (Mustafa & Siles, 2023; Parrales Obando & Pinela Tigua, 2025; Villatoro, 2022).

Aplicaciones Basadas en Machine Learning para Detección y Respuesta a Incidentes

Las aplicaciones de machine learning en ciberseguridad se han orientado principalmente a la detección temprana de incidentes mediante el análisis de patrones en el tráfico de red, los registros de eventos y el comportamiento de los usuarios. En diversos estudios se reporta el uso de algoritmos de clasificación supervisada para distinguir entre tráfico legítimo y malicioso, entrenando modelos con ejemplos etiquetados de conexiones normales y ataques, lo que permite identificar intentos de intrusión, exfiltración de datos o comunicaciones asociadas a malware de forma más rápida que con reglas estáticas (López Vidal, 2023; Vargas Gómez, 2022). De manera complementaria, se emplean técnicas de aprendizaje no supervisado orientadas a la detección de anomalías, las cuales buscan identificar comportamientos inusuales sin requerir un conocimiento previo de todas las formas de ataque, lo que resulta especialmente útil frente a amenazas nuevas o variantes poco documentadas (Mustafa & Siles, 2023; Villatoro, 2022).

En el ámbito organizacional, las propuestas basadas en machine learning se han aplicado tanto a la protección de redes empresariales como a la detección de incidentes asociados al factor humano. Algunos trabajos se centran en la construcción de modelos para identificar accesos sospechosos, patrones de uso atípicos de credenciales o acciones que se desvían del comportamiento normal de los usuarios, con el fin de reducir el impacto de ataques de phishing, suplantación de identidad o abuso de privilegios (Gómez, Ortiz, Vásquez, Marín, & Martínez, 2025; Díaz Valverde & Bravo Rojas, 2023). Otros estudios se enfocan en integrar los modelos de machine learning con sistemas de gestión de eventos e información de seguridad, permitiendo filtrar y priorizar alertas en función de la probabilidad de que correspondan a incidentes reales y

facilitando que los analistas dirijan su atención a los casos de mayor criticidad (Romero, 2023; Orozco & Ponce, 2022). En conjunto, estas aplicaciones muestran que el machine learning contribuye a mejorar la capacidad de detección y reducir tiempos de respuesta, aunque también plantean desafíos relacionados con el entrenamiento continuo de los modelos, la actualización frente a nuevas amenazas y la interpretación de los resultados por parte de los equipos de seguridad.

Aplicaciones Basadas en Deep Learning para la Detección de Patrones Avanzados

Las técnicas de deep learning han cobrado relevancia en ciberseguridad debido a su capacidad para analizar grandes volúmenes de datos y extraer representaciones complejas que permiten identificar patrones de amenaza difíciles de detectar con métodos tradicionales. En la literatura se reportan modelos basados en redes neuronales profundas aplicados al análisis de registros de eventos, flujos de red y trazas de comportamiento de usuarios, con el fin de reconocer secuencias de actividad que puedan asociarse a ataques avanzados, movimientos laterales o campañas de ransomware (Zapata Lizarazo, 2025; Martínez-Castro, 2024).

Una línea de trabajo frecuente consiste en el uso de arquitecturas como redes neuronales convolucionales y redes recurrentes para transformar el tráfico de red o los logs de sistemas en representaciones que facilitan la detección de anomalías sutiles, incluso cuando los atacantes emplean técnicas de ofuscación o variación continua de sus tácticas (Salgado García, 2024; Ortega, 2021). Otros estudios exploran el uso de modelos de deep learning para apoyar la correlación automatizada de eventos en plataformas de monitoreo, integrando información procedente de diferentes fuentes para identificar patrones de compromiso que, de otra forma, pasarían desapercibidos para los analistas humanos (Parrales Obando & Pinela Tigua, 2025; Smith & Brown, 2021). Si bien estas aplicaciones muestran mejoras en la capacidad de detección

frente a amenazas complejas, también plantean desafíos relacionados con la necesidad de grandes conjuntos de datos de entrenamiento, los costos computacionales y la dificultad de explicar de manera transparente las decisiones tomadas por los modelos ante incidentes específicos.

Tabla 2

Aplicaciones de IA para la Detección y Respuesta a Incidentes

Técnica de IA	Tipo de incidente atendido	Fase del ciclo de incidentes	Sector organizacional	Nivel de madurez organizacional	Fuente
Machine learning supervisado	Intrusiones en redes empresariales	Detección	Redes corporativas en empresas de servicios	Pruebas en entornos de laboratorio	López Vidal (2023); Vargas Gómez (2022)
Machine learning no supervisado	Anomalías en tráfico de red y comportamiento de usuarios	Detección y análisis	Organizaciones con infraestructuras distribuidas	Proyectos piloto	Mustafa & Siles (2023); Villatoro (2022)
Machine learning supervisado	Incidentes asociados al factor humano (phishing, abuso de credenciales)	Detección y análisis	Sector gobierno y entidades públicas	Implementación parcial en organizaciones	Gómez, Ortiz, Vásquez, Marín, & Martínez (2025); Díaz Valverde & Bravo Rojas (2023)

Técnica de IA	Tipo de incidente atendido	Fase del ciclo de incidentes	Sector organizacional	Nivel de madurez organizacional	Fuente
Machine learning integrado en SIEM	Priorizar alertas e identificar eventos de alta criticidad	Análisis y respuesta inicial	Empresas con centros de operaciones de seguridad (SOC)	Implementación en entornos productivos	Romero (2023); Orozco & Ponce (2022)
Deep learning (redes neuronales profundas)	Patrones avanzados de amenazas en registros y tráfico de red	Detección	Redes empresariales y sistemas corporativos	Pruebas en laboratorio y pilotos	Zapata Lizarazo (2025); Martínez-Castro (2024)
Deep learning para correlación de eventos	Campañas de ataques avanzados y movimientos laterales	Detección y análisis	Organizaciones con infraestructuras complejas	Proyectos piloto	Salgado García (2024); Smith & Brown (2021)
Sistemas expertos y correlación basada en reglas apoyados por IA	Incidentes de seguridad digital reportados en plataformas corporativas	Análisis y contención	Empresas de servicios y sector financiero	Implementación parcial en plataformas de monitoreo	Parrales Obando & Pinela Tigua (2025); Sánchez & Torres (2024)

Nota. Esta tabla muestra las aplicaciones de inteligencia artificial para la detección y respuesta a incidentes según técnica, incidente y madurez organizacional

A partir de la síntesis presentada en la tabla se observa que las aplicaciones de inteligencia artificial para la gestión de incidentes se concentran principalmente en la fase de detección y análisis, donde predominan los modelos de machine learning supervisado y no supervisado aplicados al tráfico de red, los registros de eventos y el comportamiento de los usuarios (López Vidal, 2023; Vargas Gómez, 2022; Mustafa & Siles, 2023). En menor medida, se identifican propuestas basadas en deep learning y sistemas expertos orientadas a la correlación de eventos y a la automatización parcial de respuestas, las cuales tienden a ubicarse en niveles de madurez de tipo piloto o implementación parcial, especialmente en organizaciones con infraestructuras complejas o con centros de operaciones de seguridad (Zapata Lizarazo, 2025; Salgado García, 2024; Parrales Obando & Pinela Tigua, 2025). Este panorama sugiere que, aunque existen avances importantes en el uso de técnicas avanzadas de IA, su adopción plena en entornos productivos todavía enfrenta retos asociados a la disponibilidad de datos, los costos de implementación y la necesidad de capacidades especializadas para su gestión.

Sistemas Expertos y Correlación de Eventos para Respuesta a Incidentes

Además de los modelos de machine learning y deep learning, la literatura reporta el uso de sistemas expertos y motores de correlación de eventos apoyados por reglas y técnicas de inteligencia artificial para apoyar la gestión de incidentes. Estas soluciones integran información procedente de diferentes fuentes de monitoreo y aplican reglas de negocio, umbrales y patrones predefinidos para identificar combinaciones de eventos que puedan corresponder a incidentes relevantes, sugiriendo acciones de contención o activando flujos de respuesta semiautomatizados (Romero, 2023; Parrales Obando & Pinela Tigua, 2025). En entornos como el sector financiero y las empresas de servicios, este tipo de herramientas se emplea principalmente para reducir la carga de trabajo de los analistas de seguridad, priorizando los casos de mayor criticidad y

facilitando la aplicación de procedimientos estandarizados ante incidentes recurrentes (Orozco & Ponce, 2022; Sánchez & Torres, 2024).

Relación entre Tipo de Técnica, Tipo de Incidente y Nivel de Madurez Organizacional

La revisión realizada permite identificar que la elección de la técnica de inteligencia artificial está estrechamente relacionada con el tipo de incidente que se busca atender y con el nivel de madurez de las organizaciones. En general, el machine learning supervisado se utiliza con mayor frecuencia en escenarios donde se dispone de conjuntos de datos etiquetados y se requiere distinguir de forma relativamente clara entre tráfico legítimo y malicioso, como ocurre en la detección de intrusiones y de campañas de phishing dirigidas (López Vidal, 2023; Gómez, Ortiz, Vásquez, Marín, & Martínez, 2025). Las técnicas no supervisadas y de deep learning se reservan para contextos en los que se analizan grandes volúmenes de información y se busca identificar patrones avanzados o poco frecuentes, por ejemplo, en la detección de movimientos laterales, ataques a infraestructuras complejas o comportamientos anómalos de usuarios (Zapata Lizarazo, 2025; Salgado García, 2024; Villatoro, 2022).

En cuanto al nivel de madurez organizacional, los estudios revisados muestran que las implementaciones en ambientes de producción se concentran en modelos de machine learning relativamente consolidados e integrados con herramientas existentes de monitoreo y respuesta, mientras que las propuestas basadas en deep learning y correlación avanzada de eventos se encuentran con mayor frecuencia en fases de laboratorio o en proyectos piloto. Esta situación indica que las organizaciones tienden a adoptar primero soluciones de IA de menor complejidad técnica y con requerimientos de datos más manejables, dejando las aproximaciones más sofisticadas para etapas posteriores, una vez que se han superado barreras relacionadas con infraestructura, talento humano y gobernanza de datos.

Beneficios, Limitaciones y Retos de Adopción de la Inteligencia Artificial en la Gestión de Incidentes

Beneficios Reportados en la Detección y Respuesta a Incidentes

La literatura revisada coincide en señalar que la incorporación de técnicas de inteligencia artificial en la gestión de incidentes aporta beneficios significativos en la capacidad de detección y en los tiempos de respuesta frente a amenazas de ciberseguridad. Diversos estudios reportan que los modelos de machine learning y deep learning permiten identificar patrones anómalos en el tráfico de red, los registros de eventos y el comportamiento de los usuarios con mayor rapidez y precisión que los enfoques basados exclusivamente en reglas estáticas o en análisis manual, lo que se traduce en una reducción de falsos positivos y en una priorización más efectiva de las alertas (López Vidal, 2023; Zapata Lizarazo, 2025; Smith & Brown, 2021).

Otro beneficio recurrente es la posibilidad de mantener un monitoreo más continuo y exhaustivo de las infraestructuras tecnológicas, al delegar en los modelos de IA parte de las tareas de análisis repetitivo que normalmente recargan a los equipos de seguridad. Trabajos orientados a la protección de redes empresariales y a la detección de incidentes asociados al factor humano destacan que estas soluciones contribuyen a disminuir los tiempos de detección y contención de ataques como el phishing, el malware y el ransomware, así como a mejorar la visibilidad sobre comportamientos sospechosos de usuarios y dispositivos (Gómez, Ortiz, Vásquez, Marín, & Martínez, 2025; Díaz Valverde & Bravo Rojas, 2023; Villatoro, 2022). En términos generales, los resultados sugieren que la inteligencia artificial puede fortalecer la ciberresiliencia organizacional al complementar el trabajo de los analistas, permitiendo una respuesta más oportuna y basada en evidencias frente a incidentes de seguridad.

Limitaciones Técnicas y Organizacionales de las Soluciones Basadas en IA

A pesar de los beneficios identificados, la literatura también señala múltiples limitaciones técnicas y organizacionales asociadas a la implementación de soluciones de inteligencia artificial en la gestión de incidentes. En el plano técnico, uno de los principales desafíos es la necesidad de contar con conjuntos de datos de entrenamiento suficientes, de calidad y representativos, tanto para modelos de machine learning como de deep learning, lo cual no siempre es posible debido a restricciones de confidencialidad, falta de estandarización en la recolección de registros y dificultades para etiquetar adecuadamente los eventos de seguridad (Mustafa & Siles, 2023; Salgado García, 2024). A esto se suman los costos computacionales asociados al entrenamiento y operación de modelos complejos, así como la dificultad de interpretar de manera transparente las decisiones que toman, especialmente en contextos donde se requiere justificar las acciones de respuesta ante auditorías o autoridades regulatorias (Smith & Brown, 2021; Díaz, 2021).

En el plano organizacional, varios estudios resaltan limitaciones relacionadas con la disponibilidad de talento humano especializado, la integración de las soluciones de IA con las herramientas y procesos existentes, y la resistencia al cambio dentro de las propias organizaciones. La adopción de estos sistemas exige equipos capaces de diseñar, entrenar y mantener modelos de inteligencia artificial, así como de interpretar sus resultados y traducirlos en decisiones operativas, lo que resulta complejo en entornos donde ya existen brechas de formación en ciberseguridad y ciencia de datos (Gómez, Ortiz, Vásquez, Marín, & Martínez, 2025; Romero, 2023). Adicionalmente, se señalan obstáculos vinculados a la gobernanza de los datos, la definición de responsabilidades frente a decisiones automatizadas y la necesidad de actualizar marcos normativos y políticas internas para incorporar de manera responsable estas

tecnologías en los procesos de gestión de incidentes (Méndez Castellanos, 2024; Fernández & Moreno, 2024).

Retos de Adopción en el Contexto Colombiano

En el contexto colombiano, la adopción de soluciones de inteligencia artificial para la gestión de incidentes de ciberseguridad enfrenta retos específicos que se relacionan con la infraestructura tecnológica disponible, la capacidad de inversión y las características del talento humano. Por un lado, muchas organizaciones, especialmente pymes y entidades públicas territoriales, operan con infraestructuras heterogéneas, parcialmente obsoletas o con limitaciones de conectividad, lo que dificulta la implementación de plataformas de monitoreo centralizado y el despliegue de modelos de IA que requieren procesamiento intensivo de datos en tiempo casi real (Gómez, Ortiz, Vásquez, Marín, & Martínez, 2025; Villatoro, 2022). Por otro lado, los costos asociados a la adquisición, integración y mantenimiento de estas soluciones compiten con otras prioridades presupuestales, lo que lleva a que solo sectores con mayor capacidad de inversión, como el financiero o grandes empresas de servicios, tengan condiciones más favorables para avanzar hacia proyectos piloto o implementaciones parciales.

A estos factores se suman desafíos relacionados con la cultura organizacional y la gobernanza de la seguridad. Persisten brechas significativas en formación en ciberseguridad y ciencia de datos, tanto en los equipos técnicos como en los niveles directivos, lo que dificulta la comprensión de las capacidades y límites reales de la inteligencia artificial y, en consecuencia, la toma de decisiones informadas sobre su adopción (Díaz Valverde & Bravo Rojas, 2023; Romero, 2023). Adicionalmente, la gestión de datos para entrenamiento de modelos se ve condicionada por preocupaciones sobre protección de datos personales, confidencialidad de la información y cumplimiento normativo, lo que obliga a las organizaciones a definir políticas claras sobre qué

datos pueden utilizarse, bajo qué condiciones y con qué mecanismos de anonimización o seudonimización (Méndez Castellanos, 2024; Fernández & Moreno, 2024). En conjunto, estos retos indican que la incorporación de IA en la gestión de incidentes en Colombia no puede limitarse a la adquisición de herramientas tecnológicas, sino que requiere ajustes estructurales en infraestructura, capacidades humanas y marcos de gobierno de la seguridad.

Recomendaciones Teóricas y Estrategias Generales para las Organizaciones Colombianas

A partir de los beneficios, limitaciones y retos identificados, es posible plantear algunas recomendaciones teóricas y estrategias generales para orientar la integración de la inteligencia artificial en la gestión de incidentes de ciberseguridad en organizaciones colombianas. En primer lugar, resulta conveniente que las entidades inicien con proyectos de alcance acotado, priorizando casos de uso específicos como la detección de phishing, la identificación de anomalías en tráfico de red o la priorización de alertas en plataformas de monitoreo, utilizando modelos de machine learning de menor complejidad y con requerimientos de datos manejables (Gómez, Ortiz, Vásquez, Marín, & Martínez, 2025; López Vidal, 2023). Estos proyectos piloto deberían incorporar métricas claras de evaluación, tales como reducción de falsos positivos, disminución de tiempos de detección y mejora en la tasa de identificación de incidentes relevantes, de manera que se pueda justificar su ampliación o ajuste con base en resultados verificables.

En segundo lugar, se recomienda aprovechar herramientas de código abierto y soluciones modulares que permitan experimentar con capacidades de análisis avanzado sin incurrir inicialmente en costos elevados de licenciamiento. Plataformas de monitoreo y correlación que integran módulos de detección basados en IA pueden servir como base para construir un entorno de pruebas o un centro de operaciones de seguridad (SOC) de tamaño reducido, donde se validen

reglas, modelos y flujos de respuesta antes de escalar su uso al resto de la organización (Romero, 2023; Orozco & Ponce, 2022). Este enfoque incremental facilita que las organizaciones adapten las soluciones a su contexto particular, ajusten configuraciones y definan procedimientos operativos estandarizados apoyados en IA sin comprometer de inmediato toda su infraestructura.

En tercer lugar, es necesario fortalecer de manera paralela las capacidades organizacionales requeridas para sostener el uso de inteligencia artificial en ciberseguridad. Esto implica diseñar programas de formación específicos para equipos de seguridad y de TI en temas de análisis de datos, comprensión de modelos de IA y buenas prácticas de gestión de incidentes, así como establecer mecanismos de colaboración con universidades, centros de investigación o proveedores especializados que puedan apoyar el diseño, entrenamiento y validación de modelos (Jiménez Amoroso, 2024; Mustafa & Siles, 2023). Finalmente, resulta fundamental que las organizaciones definan políticas de gobernanza de datos y de uso responsable de la inteligencia artificial, que incluyan criterios para la selección de fuentes de datos, lineamientos sobre anonimización y protección de información sensible, y procedimientos para revisar y explicar las decisiones automatizadas que tengan impacto en usuarios o procesos críticos (Méndez Castellanos, 2024; Díaz, 2021). De este modo, la integración de IA en la gestión de incidentes puede avanzar de forma gradual, apoyada en experiencias piloto y en infraestructuras de monitoreo bien definidas, alineada con la realidad colombiana y con un marco de responsabilidad y transparencia.

Conclusiones

El análisis desarrollado permitió abordar de manera integrada el problema de cómo la inteligencia artificial puede contribuir a mejorar la detección y respuesta a incidentes de ciberseguridad en organizaciones colombianas, en un contexto marcado por el incremento y la sofisticación de los ciberataques. A partir de una revisión documental sistemática de literatura científica, tesis e informes técnicos recientes, organizada en matrices de análisis y categorías específicas, se caracterizó el panorama actual de amenazas y vulnerabilidades, las principales aplicaciones de inteligencia artificial en la gestión de incidentes y los beneficios, limitaciones y retos asociados a su adopción en el contexto nacional.

En relación con el primer objetivo específico, la matriz de revisión de amenazas y sectores permitió identificar que los incidentes más recurrentes en las organizaciones son el phishing orientado al robo de credenciales, el malware y el ransomware dirigidos al cifrado o secuestro de información, así como las intrusiones y fugas de datos sensibles. Estos ataques se concentran especialmente en sectores como el financiero, el industrial, los servicios y las entidades públicas, donde la criticidad de la información y la dependencia de servicios digitales incrementan el impacto operativo y reputacional de los incidentes. De manera transversal, se evidenció la reiteración de vulnerabilidades como la falta de segmentación de redes, la gestión deficiente de parches, las configuraciones inseguras en la nube y la escasa cultura de ciberseguridad, lo que revela problemas estructurales en la forma como las organizaciones gestionan sus activos tecnológicos.

En cuanto al segundo objetivo específico, la revisión de aplicaciones, técnicas y modelos de inteligencia artificial mostró que las propuestas se concentran principalmente en el uso de machine learning y deep learning para la detección de intrusiones, anomalías en el tráfico de red,

patrones sospechosos en registros de eventos y comportamientos atípicos de usuarios. La síntesis evidenció que las técnicas supervisadas se aplican con mayor frecuencia cuando se dispone de datos etiquetados para distinguir entre tráfico legítimo y malicioso, mientras que las aproximaciones no supervisadas y los modelos profundos se orientan a la identificación de patrones avanzados o poco frecuentes, especialmente en infraestructuras complejas. Asimismo, se identificó el uso de sistemas expertos y motores de correlación de eventos apoyados en IA para priorizar alertas y sugerir acciones de respuesta, con niveles de madurez que van desde pruebas en laboratorio y proyectos piloto hasta implementaciones parciales en entornos productivos.

En relación con el tercer objetivo específico, los estudios revisados coinciden en que la integración de inteligencia artificial en la gestión de incidentes aporta beneficios como una mayor capacidad de monitoreo continuo, mejoras en la tasa de detección, reducción de falsos positivos y disminución de los tiempos de respuesta frente a amenazas relevantes. No obstante, también se identificaron limitaciones técnicas relacionadas con la disponibilidad y calidad de los datos de entrenamiento, los costos computacionales y la dificultad de explicar de forma transparente las decisiones de los modelos, así como restricciones organizacionales asociadas a la escasez de talento especializado, la integración con herramientas existentes y la gobernanza de datos. En el caso colombiano, estos desafíos se ven reforzados por brechas de infraestructura, capacidades humanas y cultura de seguridad, lo que hace necesario avanzar en estrategias graduales de adopción que incluyan proyectos piloto acotados, aprovechamiento de soluciones modulares y de código abierto, fortalecimiento de las competencias de los equipos de seguridad y definición de políticas claras de uso responsable de la inteligencia artificial.

En conjunto, los hallazgos sugieren que la inteligencia artificial ofrece un potencial significativo para fortalecer la ciberresiliencia de las organizaciones colombianas, siempre que su implementación se acompañe de una comprensión realista de sus límites, de inversiones sostenidas en capacidades técnicas y humanas y de marcos de gobernanza que garanticen la protección de los datos y la transparencia de las decisiones automatizadas.

Glosario

Amenaza de Ciberseguridad

Evento o condición potencial que puede explotar una vulnerabilidad y afectar la confidencialidad, integridad o disponibilidad de la información o de los sistemas de una organización.

Anomalía

Comportamiento, patrón o evento que se desvía de lo considerado normal en el funcionamiento de un sistema, red o usuario, y que puede asociarse a un incidente de ciberseguridad.

Ataque Dirigido

Acción maliciosa planificada contra una organización o activo específico, generalmente con objetivos concretos como robo de información, interrupción de servicios o extorsión.

Centro de Operaciones de Seguridad (SOC)

Unidad organizacional encargada de monitorear de forma continua la infraestructura tecnológica, analizar alertas de seguridad y coordinar la respuesta ante incidentes.

Ciberresiliencia

Capacidad de una organización para resistir, responder y recuperarse de incidentes de ciberseguridad, manteniendo o restableciendo sus funciones críticas en un tiempo aceptable.

Ciberseguridad

Conjunto de políticas, procesos, tecnologías y prácticas orientadas a proteger la confidencialidad, integridad y disponibilidad de la información y de los activos tecnológicos frente a accesos no autorizados, uso indebido, daños o interrupciones.

Computación en la Nube

Modelo de prestación de servicios de TI en el que recursos como infraestructura, plataformas y aplicaciones se ofrecen a través de internet bajo demanda y con elasticidad.

Deep Learning

Conjunto de técnicas de aprendizaje automático basadas en redes neuronales profundas que permiten extraer representaciones complejas de grandes volúmenes de datos y reconocer patrones avanzados de amenaza.

Detección de Intrusos

Proceso mediante el cual se identifican intentos de acceso no autorizado, comportamientos maliciosos o actividades anómalas en redes y sistemas.

Disponibilidad

Propiedad de la información y de los sistemas que garantiza que estén accesibles y utilizables cuando las personas o procesos autorizados los requieren.

Enfoque Supervisado

Tipo de aprendizaje automático en el que los modelos se entrenan con datos etiquetados que indican qué instancias son legítimas o maliciosas.

Enfoque no Supervisado

Tipo de aprendizaje automático que busca identificar patrones o agrupaciones en los datos sin disponer de etiquetas previas, útil para detectar anomalías o comportamientos inusuales.

Fuga de Datos

Incidente de seguridad en el que información sensible, confidencial o protegida es expuesta, exfiltrada o divulgada a actores no autorizados.

Gestión de Incidentes de Ciberseguridad

Proceso estructurado que comprende las fases de detección, análisis, contención, erradicación y recuperación ante eventos que comprometen la seguridad de los sistemas.

Gobernanza de Datos

Conjunto de políticas, roles, procesos y controles que regulan la gestión y el uso de los datos dentro de una organización, incluyendo su protección, calidad y acceso.

Ingeniería Social

Conjunto de técnicas que buscan manipular o engañar a personas para que realicen acciones que comprometan la seguridad, como revelar credenciales o ejecutar archivos maliciosos.

Infraestructura como Servicio (IaaS)

Modelo de servicio en la nube en el que un proveedor ofrece recursos de cómputo básicos (servidores, almacenamiento, redes) que el cliente puede aprovisionar y gestionar de forma flexible.

Integridad

Propiedad de la información que garantiza que los datos se mantienen completos, exactos y no han sido alterados de forma no autorizada.

Inteligencia Artificial (IA)

Campo de la informática que desarrolla sistemas capaces de aprender a partir de datos, reconocer patrones y apoyar o tomar decisiones de forma autónoma en contextos específicos.

Intrusión

Acceso o intento de acceso no autorizado a sistemas, redes o aplicaciones, ya sea para obtener información, alterar su funcionamiento o interrumpir servicios.

Machine Learning

Subcampo de la inteligencia artificial que utiliza algoritmos capaces de aprender patrones a partir de datos, con el fin de realizar tareas como clasificación, predicción o detección de anomalías.

Malware

Software malicioso diseñado para dañar, interrumpir, espiar o tomar control de sistemas y dispositivos sin autorización del usuario o de la organización.

Nivel de Madurez Organizacional

Grado de desarrollo de una organización en la adopción de soluciones de seguridad o de IA, que puede ir desde pruebas de laboratorio y proyectos piloto hasta implementaciones en ambientes de producción.

Phishing

Técnica de ingeniería social que busca engañar a usuarios para que revelen credenciales o información sensible mediante mensajes o sitios falsos que simulan ser legítimos.

Ransomware

Tipo de malware que cifra información o bloquea sistemas y exige un pago, generalmente en criptomonedas, para restaurar el acceso a los datos o servicios afectados.

Red Empresarial

Conjunto de dispositivos, sistemas, servicios y componentes de comunicación que soportan los procesos tecnológicos de una organización.

Respuesta a Incidentes

Conjunto de actividades y procedimientos que se ejecutan después de detectar un incidente de seguridad, con el fin de contenerlo, erradicarlo y recuperar la operación normal.

SDN (Software Defined Networking)

Enfoque de redes en el que el plano de control se separa del plano de datos, permitiendo gestionar y programar el comportamiento de la red de forma centralizada mediante software.

SD-WAN (Software Defined Wide Area Network)

Arquitectura de red de área amplia definida por software que permite gestionar y optimizar el tráfico entre sedes y recursos en la nube de forma centralizada y dinámica.

SIEM (Security Information and Event Management)

Plataforma que recopila, correlaciona y analiza eventos de seguridad provenientes de diferentes fuentes para generar alertas y apoyar la investigación de incidentes.

Sistema de Detección de Intrusos (IDS)

Herramienta de seguridad que monitorea redes o sistemas en busca de actividades maliciosas o violaciones de políticas, generando alertas cuando detecta patrones sospechosos.

Sistema de Prevención de Intrusos (IPS)

Solución que, además de detectar actividades maliciosas, puede bloquear o interrumpir de manera automática el tráfico o las acciones asociadas al ataque.

Sistemas Expertos

Programas informáticos que incorporan reglas y conocimiento estructurado para apoyar la toma de decisiones en dominios específicos, simulando el razonamiento de especialistas humanos.

Suplantación de Identidad

Uso indebido de credenciales o atributos de otra persona o entidad para acceder a sistemas, ejecutar transacciones o realizar acciones en su nombre.

Trazas de Comportamiento

Registros detallados de las acciones realizadas por usuarios, aplicaciones o dispositivos en un sistema, utilizados para análisis forense, detección de anomalías o auditoría.

Vulnerabilidad

Debilidad en un sistema, proceso o control de seguridad que puede ser explotada por una amenaza para comprometer la confidencialidad, integridad o disponibilidad de la información.

Referencias Bibliográficas

Aly, W. H. F. (2017). LBFTFB fault tolerance mechanism for software defined networking. En *Proceedings of the 2017 International Conference on Electrical and Computing Technologies and Applications (ICECTA 2017)* (pp. 1–5). Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/ICECTA.2017.8251995

Amarasinghe, H., Jarray, A., & Karmouch, A. (2021). Survivable IaaS management with SDN. *IEEE Transactions on Cloud Computing, 9*(4), 1619–1633. https://doi.org/10.1109/TCC.2019.2910111

Bilbao-Arechabala, S., Atutxa, A., & Del Ser, J. (2025). Clasificación de procedimientos de ataques de ciberseguridad mediante Generación Aumentada por Recuperación. *Procesamiento del Lenguaje Natural, 75*, 199–210. https://doi.org/10.26342/2025-75-15

Carrillo, S. (2023). Machine learning y privacidad en entornos corporativos. *Revista de Derecho Informático, 15*(3), 201–215. https://derechoinformatico.revistadigital.org/article/view/315

Cuya-Chamilco, P. V., Saavedra-Villar, P., Mescua-Ampuero, L. E., Alvarado-Flores, J., & Cáceres-García, A. A. (2025). Ciberseguridad desde las ciencias de la computación: Un estudio bibliométrico sobre el Perú y los países limítrofes. *Revista Científica de Sistemas e Informática, 5*(1), e862. https://doi.org/10.51252/rcsi.v5i1.862

De La Torre Guzman, J., Salazar Jacome, M. E., Jurado Pruna, F. X., & Perez Cargua, M. R. (2023). SD-WAN software defined networking using DMVPN for corporate enterprises. En *Actas de la 18th Iberian Conference on Information Systems and Technologies (CISTI 2023)*. IEEE Computer Society. https://doi.org/10.23919/CISTI58278.2023.10211444

Díaz, V. A. (2021). Inteligencia artificial y ética en la seguridad informática. *Cuaderno de Trabajo UNAM, 5*(3), 10–28. https://www.cuadernodeunam.unam.mx/index.php/cuadernodeunam/article/view/2167

Díaz Valverde, E., & Bravo Rojas, C. (2023). Inteligencia artificial en la detección y mitigación de amenazas de seguridad. *Revista Científica CienciaPlus, 18*(2), 41–59. https://revistacienciaplus.org/index.php/cienciaplus/article/download/155/109

Dudczyk, J., Sergiel, M., & Krygier, J. (2025). Analysis of SD-WAN architectures and techniques for efficient traffic control under transmission constraints—Overview of solutions. *Sensors, 25*(20), 6317. https://doi.org/10.3390/s25206317

European Union Agency for Cybersecurity (ENISA). (2022). *Artificial intelligence threat landscape report*. ENISA Publications. https://www.enisa.europa.eu/publications/artificial-intelligence-threat-landscape

Fernández, R., & Moreno, A. (2024). Ciberdefensa y algoritmos en la seguridad empresarial. *Revista Española de Ciberseguridad, 6*(1), 88–104. https://reciberseg.org/articulos/2024/algoritmos-ciberdefensa

Galvez, D. P. C., & Revinova, S. Y. (2025). Desafíos de ciberseguridad para las pequeñas y medianas empresas. *Iberoamérica, 2*, 48–67. https://doi.org/10.37656/s20768400-2025-02-03

Gillgallon, R., Almutairi, R., Bergami, G., & Morgan, G. (2025). SimulatorOrchestrator: A 6G-ready simulator for the cell-free/osmotic infrastructure. *Sensors, 25*(5), 1591. https://doi.org/10.3390/s25051591

Gómez, J. A. U., Ortiz, J. I. B., Vásquez, J. M. D., Marín, A. S., & Martínez, J. E. M. (2025). Análisis de brechas de seguridad de la información en el factor humano: Caso de estudio

aplicado en el sector gobierno. *Revista Conhecimento Online*, *2*, 295–313. https://doi.org/10.25112/rco.v2.4160

Herrera Guzmán, E. (2025). Inteligencia artificial y ciberseguridad: Transformación digital de la seguridad nacional en el siglo XXI. *Revista Inclusiones*, *12*(3), 1–20. https://revistainclusiones.org/index.php/inclu/article/view/3639/3903

Jiménez Amoroso, M. B. (2024). *Inteligencia artificial: Un estudio de su impacto en ciberseguridad* [Trabajo de fin de máster, Universidad Politécnica de Madrid]. Archivo Digital UPM. https://oa.upm.es/82685/1/MARIA_BELEN_JIMENEZ_AMOROSO.pdf

López Vidal, A. B. (2023). *Machine learning como herramienta para la defensa proactiva en ciberseguridad* [Tesis de maestría, Universidad Nacional Autónoma de México]. Repositorio UNAM. https://repositorio.unam.mx/bitstream/handle/123456789/2872284/lopezvidalanabeatriz2023.pdf

Martínez-Castro, J. P. (2024). *Machine learning y ciberseguridad: Aplicaciones para la protección empresarial* [Trabajo de grado, Universidad Distrital Francisco José de Caldas]. Repositorio UDFJC. https://repository.udistrital.edu.co/bitstream/handle/11349/35605/MartinezCastroJuanPablo2024.pdf

Méndez Castellanos, B. E. (2024). Inteligencia artificial y ciberseguridad: Retos para la sociedad digital. *Revista Tecnológica - ESPOL*, *37*(1), 15–31. https://www.rte.espol.edu.ec/index.php/tecnologica/article/view/678/420

Mustafa, J., & Siles, M. O. (2023). *Revisando las tendencias globales de inteligencia artificial en ciberseguridad* [Tesis de grado, Universidad Nacional Autónoma de Nicaragua]. Repositorio UNAN. https://repositorio.unan.edu.ni/17660/1/25144.pdf

Orozco, M., & Ponce, C. (2022). Defensa digital y soluciones basadas en IA. *Revista Seguridad Informática, 12*(4), 59–71. https://seguridadinformatica.redalyc.org/articulo.oa?id=217927

Ortega, J. (2021). Cybersecurity and artificial intelligence: Opportunities and threats. *International Journal of Computer Science & Information Technology, 13*(4), 50–65. https://airccse.org/journal/jcsit/archives/jcsit2021.html

Parrales Obando, A. D., & Pinela Tigua, J. O. (2025). Análisis del impacto de la inteligencia artificial en la ciberseguridad. *Revista Instituto Superior Tecnológico Vicente Rocafuerte, 1*(1), 1–22. https://revista.istvr.edu.ec/wp-content/uploads/2025/03/ANALISIS-DEL-IMPACTO-DE-LA-INTELIGENCIA-ARTIFICIAL-EN-LA-CIBERSEGURIDAD.pdf

Pérez Galla, J. (2025). *Agentes de Inteligencia Artificial en Seguridad Ofensiva* [Trabajo de fin de grado, Universitat Oberta de Catalunya]. UOC Open Access. https://openaccess.uoc.edu/bitstream/10609/152174/1/jperezgallaTFG0125.pdf

Rengifo Pabón, H. J. (2023). *Aplicación de IA para contención de incidentes cibernéticos en sistemas críticos* [Tesis de maestría, Universidad Nacional de Colombia]. Repositorio UNAL. https://repositorio.unal.edu.co/bitstream/handle/unal/85753/rengifopabonhernandojose2023.pdf

Romero, L. P. (2023). La automatización en la gestión de incidentes de seguridad digital. *Revista Colombiana de Computación, 8*(1), 22–37. https://revistas.unal.edu.co/index.php/colcomputacion/article/view/103506

Salas Herrera, D. del R. (2025). *Evaluación de modelos de IA para respuesta a incidentes en instituciones financieras* [Trabajo de fin de máster, Universidad San Ignacio de Loyola]. Repositorio USIL. https://repositorio.usil.edu.pe/bitstream/USIL/12835/2/salasherreradanieladel2025.pdf

Salgado García, G. B. (2024). *Aplicaciones de la Inteligencia Artificial Generativa (IAG) en el ámbito de la ciberseguridad* [Trabajo de fin de máster, Universitat Oberta de Catalunya]. UOC Open Access. https://openaccess.uoc.edu/bitstream/10609/150603/1/salgadogbTFM0624memoria.pdf

Sánchez, A. L., & Torres, G. (2024). Tendencias de inteligencia artificial en la protección de infraestructuras críticas. *Revista Iberoamericana de Ingeniería*, *30*(2), 123–132. https://revistaiberoamericana.org/ingenieria/article/view/2512

Smith, R., & Brown, T. (2021). Artificial intelligence for incident response: A review. *Cybersecurity Journal, 4*(1), 31–48. https://cybersecurityjournal.springeropen.com/articles/10.1186/s42400-021-00069-w

Tang, C., Zhao, B., Chen, Y., & Zhou, H. (2024). DABS: A distributed attribute-based access control mechanism for SD-WAN controllers based on smart contracts. En *Proceedings of the IEEE 1st International Conference on Metacomputing (ICMC 2024)* (pp. 292–299). Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/ICMC60390.2024.00039

Vargas Gómez, L. (2022). *Aplicación de inteligencia artificial para mejorar la respuesta ante incidentes en redes empresariales* [Tesis de grado, Universidad de San Carlos de Guatemala]. USAC Repositorio. https://biblioteca.usac.edu.gt/tesis/18/18_0390.pdf

Villatoro, R. (2022). Ciberseguridad y machine learning: Retos para América Latina. *Revista Digital de Ciberseguridad*, *9*(2), 44–58. https://www.revistacyberlat.org/index.php/revista/article/view/632

Zapata Lizarazo, C. A. (2025). *Deep learning para la detección de patrones avanzados de amenazas* [Tesis de grado, Universidad de Antioquia]. Repositorio UdeA. https://bibliotecadigital.udea.edu.co/bitstream/10495/35841/1/zapatalizarazocarlosandres2025.pdf

Zárate, H. R. (2024). *Inteligencia artificial en la prevención y detección de ataques informáticos* [Trabajo de fin de máster, Universidad de Alcalá]. eBUAH Repositorio UAH. https://ebuah.uah.es/dspace/bitstream/handle/10017/56814/TFM_Zarate_Hector_2024.pdf