

## Generative artificial intelligence in cybersecurity: a systematic literature review

### Inteligencia artificial generativa en el ámbito de la ciberseguridad: una revisión sistemática de literatura

#### Autores:

Cordova-Alvarado, Robin Lenin  
UNIVERSIDAD CATÓLICA DE CUENCA - UCACUE  
Estudiante de Maestría en Ciberseguridad  
Cuenca – Ecuador



[robin.cordova.76@est.ucacue.edu.ec](mailto:robin.cordova.76@est.ucacue.edu.ec)



<https://orcid.org/0000-0002-7067-2238>

Andrade-López, Miguel Santiago  
UNIVERSIDAD CATÓLICA DE CUENCA - UCACUE  
Docente de Maestría en Ciberseguridad  
Cuenca – Ecuador



[msandradel@ucacue.edu.ec](mailto:msandradel@ucacue.edu.ec)



<https://orcid.org/0000-0002-6882-4204>

Álvarez-Vera, Manuel Salvador  
UNIVERSIDAD CATÓLICA DE CUENCA - UCACUE  
Docente de Maestría en Ciberseguridad  
Cuenca – Ecuador



[malvarezv@ucacue.edu.ec](mailto:malvarezv@ucacue.edu.ec)



<https://orcid.org/0000-0002-2521-0042>

Fechas de recepción: 03-JUN-2024 aceptación: 03-JUL-2024 publicación: 15-SEP-2024



<https://orcid.org/0000-0002-8695-5005>

<http://mqrinvestigar.com/>



## Resumen

La presente Revisión Sistemática de Literatura (RSL) tiene por objetivo identificar el estado actual de la Inteligencia Artificial Generativa (Gen IA) en el ámbito de la Ciberseguridad, investigando también sus aplicaciones, técnicas, desafíos éticos, riesgos y limitaciones. Se utilizó la metodología propuesta por Barbara Kitchenham, dejando como resultado 31 trabajos de investigación distribuidas en seis bases de datos de prestigio investigativo en el área, que permitieron identificar información relevante y ayudaron a responder las preguntas de investigación planteadas. La tecnología de Inteligencia Artificial Generativa ofrece diversas aplicaciones prometedoras en el ámbito de la ciberseguridad, dentro de los principales descubrimientos de la Gen IA, facilita la identificación de vulnerabilidades y la prevención de amenazas, analiza código malicioso, mejora la seguridad de la red, simula escenarios de ciberseguridad, genera datos sintéticos para entrenar modelos de IA y fomenta la educación. Sin embargo, a los ciberdelincuentes les ayuda a perfeccionar algunas habilidades como la generación de phishing, ransomware, creación de deepfakes, generación de código malicioso y facilita la desinformación. Como conclusión para el presente trabajo de investigación, podemos decir que la Gen IA, se está utilizando tanto para los equipos de defensa o personal de seguridad informática, como también los ciberdelincuentes para llevar a cabo sus ataques en ciberseguridad, esto presenta algunas limitaciones, riesgos y desafíos éticos.

**Palabras clave:** inteligencia artificial generativa; gen IA; ciberseguridad; detección de amenazas; privacidad

## Abstract

This Systematic Literature Review (SLR) aims to identify the current state of Generative Artificial Intelligence (Gen AI) in the field of Cybersecurity, investigating its applications, techniques, ethical challenges, risks, and limitations. The methodology proposed by Barbara Kitchenham was employed, resulting in 31 research papers sourced from six prestigious research databases in the area. These papers provided relevant information and helped address the research questions. Generative Artificial Intelligence technology offers various promising applications in cybersecurity. Among the key findings of Gen AI, it facilitates vulnerability identification and threat prevention, analyzes malicious code, enhances network security, simulates cybersecurity scenarios, generates synthetic data for AI model training, and promotes education. However, it also assists cybercriminals in refining skills such as phishing generation, ransomware, deepfake creation, malicious code generation, and misinformation dissemination. In conclusion, this research work highlights that Gen AI is being utilized by both defense teams and cybercriminals in cybersecurity operations, presenting inherent limitations, risks, and ethical challenges.

**Keywords:** generative artificial intelligence; gen AI; cybersecurity; threat detection; privacy

## Introducción

La convergencia entre la inteligencia artificial y la ciberseguridad ha emergido como un tema imperativo, generando debates y estudios esenciales (Truong et al., 2020). Si bien existen RSL que ayudan a comprender los temas y proponen lineamientos para nuevos trabajos investigativos, su impacto se va reduciendo en base al tiempo de su publicación.

La Inteligencia Artificial Generativa se refiere a sistemas de inteligencia artificial con la capacidad de producir contenido, como texto, imágenes y audio, a partir de vastos conjuntos de datos existentes. Esta capacidad se logra mediante el uso de algoritmos y modelos complejos que aprenden de grandes volúmenes de datos, reconocen las estructuras subyacentes y las emulan de manera única (Cao et al., 2023).

La democratización de la IA Generativa ha transformado la tecnología de inteligencia artificial, con avances significativos en algoritmos de aprendizaje automático y redes neuronales a principios de la década de 2020. Estos avances han permitido la creación de modelos más sofisticados y eficientes. Al mismo tiempo, los costos de desarrollo y despliegue han disminuido debido a la caída de los precios del poder de cómputo y la disponibilidad de herramientas de código abierto (Ferrara, 2024). La accesibilidad de la IA Generativa se ha ampliado gracias a interfaces fáciles de usar y servicios en la nube, fomentando la innovación y creatividad en diversos sectores, permitiendo a personas y organizaciones aprovechar esta tecnología.

El impacto que tiene la inteligencia artificial en el ámbito de la ciberseguridad va en aumento con los crecientes avances tecnológicos y con la llegada de la computación cuántica (Wiafe et al., 2020), estos aspectos mejorarían significativamente los algoritmos implementados en la rama de la inteligencia artificial.

La razón fundamental que respalda la realización de revisiones sistemáticas es la identificación de posibles brechas en la investigación y la ampliación de los límites del conocimiento en un área específica. Este procedimiento se ejecuta con la intención de ofrecer una revisión de elevada calidad, caracterizada por su transparencia y capacidad de reproducibilidad (Manterola et al., 2013). El propósito subyacente es condensar los estudios de investigación disponibles, aportando así una contribución valiosa al cuerpo de conocimientos existente.

El trabajo de investigación se limitará únicamente a responder a las preguntas de investigación planteadas dentro de un protocolo mediante la elaboración de una revisión sistemática de literatura basado en “Guidelines for performing Systematic Literature Reviews in Software Engineering” propuesto por Barbara Kitchenham (Kitchenham, 2007), que nos

proporciona un protocolo para realizar Revisiones Sistemáticas de Literatura (RSL). Dentro de la RSL, se realizará una revisión exhaustiva de documentos que cumplan con los criterios de inclusión y exclusión que se detallarán en el protocolo del trabajo investigativo a realizar, adicionalmente serán filtrados solo los que logren responder a los trabajos de investigación.

El objetivo es ofrecer una visión integral y actualizada del estado actual de la investigación en este campo dinámico y crítico para la protección de sistemas informáticos y que nuevos investigadores tracen una ruta hacia estudios más específicos que con la finalidad que mejores y reduzcan tiempo en su investigación.

Con esta RSL se obtendrá una mejor perspectiva de qué existe en materia de IA Generativa en el ámbito de la ciberseguridad, y de esta manera identificar y tener presente el impacto que esta rama de las Ciencias de la Computación tiene dentro de las mismas, especialmente en su aplicación.

## Material y métodos

La metodología expuesta en este trabajo se rigió principalmente por el protocolo de Barbara Kitchenham para la elaboración de RSL (Kitchenham, 2007). No obstante, los criterios de calidad presentados fueron procesos híbridos definidos por los autores.

### Planificación

#### *Formulación de las preguntas de investigación*

A continuación, se presentan las preguntas de investigación, las cuales rige y está enfocada la presente RSL:

- RQ1: ¿Cuál es el estado actual sobre aplicación de Inteligencia Artificial Generativa en Ciberseguridad?
- RQ2: ¿Cuáles son las aplicaciones o técnicas más prometedores de la Inteligencia Artificial Generativa para detectar y mitigar ataques cibernéticos?
- RQ3: ¿Cuáles son los desafíos éticos, riesgos y limitaciones asociados con la implementación de inteligencia Artificial Generativa en Ciberseguridad?

#### *Definir las palabras claves*

Para la realización de la presente RSL se dispuso la siguiente lista de palabras claves que reflejan específicamente los intereses del objeto de estudio, basándonos en el Thesaurus de la IEEE (Moscara, 2019), como base de buen fundamento para poder referenciar correctamente la terminología dentro de la investigación.

- Generative Artificial Intelligence, Gen AI, Generative AI (definido por los autores)
- Ciberseguridad, Computer security (cybersecurity)
- Técnicas, techniques; Aplicaciones; applications

**Fuentes de datos y estrategia de búsqueda**

***Definir las bases de datos***

Las fuentes bibliográficas para la realización de la RSL son las descritas en la Tabla 1:

**Tabla 1**  
*Listado de Base de datos científicas*

Base de datos	URL
WoS	<a href="http://apps.webofknowledge.com/">http://apps.webofknowledge.com/</a>
Scopus	<a href="http://scopus.com/">http://scopus.com/</a>
ACM	<a href="https://dl.acm.org/">https://dl.acm.org/</a>
IEEE Xplore	<a href="https://ieeexplore.ieee.org/">https://ieeexplore.ieee.org/</a>
Springer Link	<a href="https://link.springer.com/">https://link.springer.com/</a>
Science Direct	<a href="https://www.sciencedirect.com/">https://www.sciencedirect.com/</a>

Nota. Bases de datos científicas definidas por los autores.

***Definir y comprobar los scripts de búsqueda***

En la Tabla 2, se muestra los scripts de búsqueda para poder implementarlos en la sección de ejecución, para ello se generó los scripts para cada base de datos de manera individual.

**Tabla 2**  
*Definición de scripts de búsqueda por base de datos*

N	Bases de datos	Script
1	WoS	Refine results for Generative Artificial Intelligence (Resumen) AND cybersecurity (Resumen) and Artículo or articulo de revision (Tipos de documentos)
2	Scopus	TITLE-ABS-KEY ( "Generative Artificial Intelligence" AND " cybersecurity " ) AND PUBYEAR > 2019 AND PUBYEAR < 2025 AND ( LIMIT-TO ( DOCTYPE , "ar" ) OR LIMIT-TO ( DOCTYPE , "cp" ) ) AND ( LIMIT-TO ( LANGUAGE , "English" ) )
3	ACM	{Abstract:("Generative Artificial Intelligence") AND AllField:(cybersecurity)} "filter": {E-Publication Date: Past 5 years},{ACM Content: DL}
4	IEEE	("Abstract": "Generative Artificial Intelligence") AND ("Full Text Only":cybersecurity) AND PUBYEAR > 2019 ("Full Text Only": "Generative Artificial Intelligence") AND ("Abstract":cybersecurity)
5	Springer	"Generative Artificial Intelligence" AND cybersecurity
6	Science Direct	"Generative Artificial Intelligence" AND cybersecurity Refine by: years 2019,2020,2021,2022,2023,2024 Article type Review articles,



Nota. Scripts de búsqueda desarrollados por los autores.

### ***Ejecutar los scripts de búsqueda***

En la Tabla 3 se presenta los resultados en número de documentos tras ejecutar los scripts de búsqueda establecidos en la sección anterior.

**Tabla 3**

*Resultados de la ejecución de scripts de búsqueda por base de datos*

Bases de Datos						Cantidad	
WoS	Scopus	ACM	IEEE Xplore	Springer	Science Direct	Total Documentos	Documentos únicos
6	16	7	30	16	26	101	93

### ***Descargar y guardar los resultados de los scripts de búsqueda***

Se utilizó la herramienta Mendeley para gestionar las referencias bibliográficas, de la ejecución de los scripts se ha logrado identificar un total de 93 trabajos que serán evaluados con los criterios de inclusión y exclusión.

### ***Definir los criterios de inclusión y exclusión***

#### **Criterios de inclusión.**

Para la RSL, se consideran los criterios de inclusión de la siguiente tabla:

**Tabla 4**

*Criterios de inclusión para la RSL*

Criterio	Descripción
Contenido	Contener al menos una cadena de búsqueda en su título o resumen.
Fecha de publicación	Estudios publicados desde el año 2019.
Motores de búsqueda	WoS, Scopus, ACM, IEEE Xplore, Springer Link, ScienceDirect.
Idioma	Se consideran artículos estrictamente en inglés.
Tipos de estudios	artículos y artículos de conferencias.

#### **Criterios de exclusión.**

Los estudios no serán considerados si no contienen información que aporten a responder las preguntas de investigación y que no cumplan con los criterios de calidad e inclusión.

### **Selección de estudios y evaluación de la calidad**

Ejecutando la fase de Selección de estudios y evaluación de la calidad, se lograron los siguientes resultados que se detallan en la Figura 1.

**Figura 1**

*Ejecución de fases de análisis estudios y evaluación de la calidad*



Fuente: Elaboración propia

Aquellos trabajos que superaron los análisis y el control de calidad establecido por el equipo investigador se describen en la Tabla 5.

**Tabla 5**

*Estudios que culminaron exitosamente las fases de selección y criterios de calidad*

ID	Base de datos	Estudio	Año
ES01	WoS	AIGC challenges and opportunities related to public safety: A case study of ChatGPT (Guo et al., 2023).	2023
ES02	WoS	Investigating ChatGPT and cybersecurity: A perspective on topic modeling and sentiment analysis (Okey et al., 2023).	2023
ES03	WoS	A Machine-Learning-Based Cyberattack Detector for a Cloud-Based SDN Controller (Mozo et al., 2023).	2023
ES04	Scopus	Generative AI for pentesting: the good, the bad, the ugly (Hilario et al., 2024).	2024
ES05	Scopus	Phishing to improve detection (Zheng & Becker, 2023).	2023
ES06	Scopus	GenAI in the Cyber Kill Chain: A Comprehensive Review of Risks, Threat Operative Strategies and Adaptive Defense Approaches (Deshpande & Gupta, 2023).	2023
ES07	Scopus	Cyber Security Issues and Challenges Related to Generative AI and ChatGPT (Pasupuleti et al., 2023).	2023
ES08	Scopus	Change Management using Generative Modeling on Digital Twins (Das et al., 2023).	2023
ES09	Scopus	A Dimensional Perspective Analysis on the Cybersecurity Risks and Opportunities of ChatGPT-Like Information Systems (Hu &	2023

Chen, 2023).

ES10	Scopus	Scamming the Scammers: Using ChatGPT to Reply Mails for Wasting Time and Resources (Cambiaso & Caviglione, 2023).	2023
ES11	Scopus	Generative AI for Cyber Threat-Hunting in 6G-enabled IoT Networks (Ferrag et al., 2023).	2023
ES12	Scopus	ChatGPT: Vision and challenges (Gill & Kaur, n.d.).	2023
ES13	Scopus	Examine the Role of Generative AI in Enhancing Threat Intelligence and Cyber Security Measures (Saddi et al., 2024).	2024
ES14	Scopus	Privacy and Security Concerns in Generative AI: A Comprehensive Survey (Golda et al., 2024).	2024
ES15	ACM	A Lightweight Generative Adversarial Network for Imbalanced Malware Image Classification (Chui, 2023).	2023
ES16	IEEE	Advancements in Generative AI: A Comprehensive Review of GANs, GPT, Autoencoders, Diffusion Model, and Transformers (Bengesi et al., n.d.).	2024
ES17	IEEE	Generative AI for Physical Layer Communications: A Survey (Van Huynh et al., 2023).	2024
ES18	IEEE	Generative AI for Industry 5.0: Analyzing the impact of ChatGPT, DALLE, and Other Models (Sai et al., n.d.).	2024
ES19	IEEE	Security Risks Concerns of Generative AI in the IoT (Xu et al., n.d.).	2024
ES20	IEEE	Can Large Language Models Better Predict Software Vulnerability? (Katsadourous et al., 2023).	2023
ES21	IEEE	Disinfecting AI: Mitigating Generative AI's Top Risks (Campbell & Jovanovic, 2024).	2024
ES22	IEEE	The Impact of Generative AI and LLMs on the Cybersecurity Profession (Capodiecici et al., 2024).	2024
ES23	IEEE	ChatGPT's Security Risks and Benefits: Offensive and Defensive Use-Cases, Mitigation Measures, and Future Implications (Charfeddine et al., 2024).	2024
ES24	IEEE	From ChatGPT to ThreatGPT: Impact of Generative AI in Cybersecurity and Privacy (Gupta et al., 2023).	2023
ES25	IEEE	Generative AI for Cyber Security: Analyzing the Potential of ChatGPT, DALL-E, and Other Models for Enhancing the Security Space (Sai et al., 2024).	2024
ES26	Springer	Ransomware attacks in the context of generative artificial intelligence—an experimental study (Teichmann, 2023).	2023
ES27	Springer	GenAI against humanity: nefarious applications of generative artificial intelligence and large language models (Ferrara, 2023).	2024
ES28	Springer	A Comprehensive Survey of Attack Techniques,	2024

		Implementation, and Mitigation Strategies in Large Language Models (Esmradi et al., n.d.).	
ES29	Science Direct	China's Interim Measures on generative AI: Origin, content and significance (Migliorini, 2024).	2024
ES30	Science Direct	Ethical and regulatory challenges of large language models in medicine (Ong et al., 2024).	2024
ES31	Science Direct	ChatGPT for digital forensic investigation: The good, the bad, and the unknown (Scanlon et al., 2023).	2023

Ejecutando la etapa de selección de artículos mediante la lectura crítica y criterio de calidad, se obtuvieron 31 estudios de los cuales se ha observado que 16 trabajos, es decir más del 50% están indexados en revistas clasificadas en cuartiles Q1 según el SJR, mientras que 3 se encuentran en Q2, y uno en Q4. No obstante, el resto de las investigaciones, debido a su contemporaneidad, aún no han sido evaluadas según las métricas del SJR, sin embargo, han demostrado contener excelente información para responder a las preguntas de investigación.

### Extracción y síntesis de datos

#### *Extraer las respuestas a las preguntas de investigación*

En la fase anterior se define los estudios que serán utilizados para responder a las preguntas de investigación, un total de 31 estudios como base para la presente RSL, estos estudios culminaron exitosamente todas las fases de selección y criterios de calidad. A continuación, se realiza la síntesis de datos de tipo cualitativo por la naturaleza de los datos, donde se responden a las 3 preguntas de investigación planteadas en la RSL.

**Tabla 6**

*Extracción de respuestas para la pregunta de investigación 1*

<b>Pregunta</b>	<b>RQ1: ¿Cuál es el estado actual sobre aplicación de Inteligencia Artificial Generativa en Ciberseguridad?</b>
<b>Estudios</b>	<i>ES01, ES02, ES03, ES04, ES05, ES06, ES07, ES08, ES09, ES10, ES11, ES12, ES13, ES14, ES15, ES16, ES17, ES18, ES19, ES20, ES21, ES22, ES23, ES24, ES25, ES26, ES27, ES31.</i>
<b>Resultados</b>	<i>La aplicación de la Inteligencia Artificial Generativa en ciberseguridad ha mostrado avances significativos y preocupaciones críticas. Herramientas como ChatGPT han logrado mejorar la seguridad de la red al simular escenarios y fomentar la educación en ciberseguridad, sin embargo, pueden ser explotadas por ciberdelincuentes para desarrollar software de suplantación de identidad, crear sitios web falsos y generar código malicioso, con ello se ha facilitado la generación de ataques de ingeniería social, comprometiendo la privacidad de los datos. La supervisión humana</i>

*es esencial para garantizar resultados precisos y abordar falsos positivos o negativos generados por la IA.*

*Las capacidades de la Gen AI para reproducir escenarios reales facilitan la detección de vulnerabilidades de día cero, los métodos convencionales de detección de phishing están quedando obsoletos ante la creciente sofisticación de las estafas de phishing alimentadas por Gen AI. Además, la Gen AI se utiliza para crear correos electrónicos de phishing convincentes, vídeos falsos y distribuir información falsa en redes sociales, haciendo que los ataques sean más difíciles de detectar. La Gen AI también es empleada para generar deepfakes, mejorar estrategias de control y explotación, y desarrollar ciberataques autónomos que se adaptan y evolucionan automáticamente.*

*La tecnología de Gen IA se utiliza para analizar anomalías en redes, proteger contra malware y supervisar el cumplimiento de la privacidad, pero también para descubrir y explotar vulnerabilidades, escribir ransomware y generar correos electrónicos maliciosos La IA generativa también se aplica en la generación de datos sintéticos para entrenar modelos de aprendizaje automático. Estas capacidades permiten identificar y clasificar con precisión diferentes tipos de ciberamenazas.*

*Existen herramientas avanzadas de ciberamenazas, como WormGPT y FraudGPT, estas presentan un potencial amenazante debido a su sofisticación y capacidad de evasión, comprometiendo los métodos de cifrado existentes y planteando riesgos significativos para las infraestructuras críticas. El uso de IA generativa ha revelado preocupaciones significativas sobre la exposición de información sensible y la creación de contenido de ataque de alta calidad. Algunos estudios indican que empleados han introducido datos internos en plataformas como ChatGPT, aumentando los riesgos de ciberseguridad.*

**Tabla 7**

*Extracción de respuestas para la pregunta de investigación 2*

<b>Pregunta</b>	<b>RQ2: ¿Cuáles son las aplicaciones o técnicas más prometedoras de la Inteligencia Artificial Generativa para detectar y mitigar ataques cibernéticos?</b>
<b>Estudios</b>	<i>ES01, ES02, ES03, ES04, ES06, ES07, ES09, ES10, ES11, ES13, ES15, ES17, ES20, ES21, ES23, ES24, ES25, ES28.</i>
<b>Resultados</b>	<i>La tecnología de Inteligencia Artificial Generativa ofrece diversas aplicaciones prometedoras en el ámbito de la ciberseguridad. Entre sus capacidades destacan: ChatGPT puede analizar código malicioso, detectar comportamientos sospechosos y clasificar códigos maliciosos, mejorando</i>

*así la detección y prevención de amenazas. Además, puede analizar grandes volúmenes de datos de inteligencia sobre amenazas, facilitando la identificación de vulnerabilidades y la generación de nuevos escenarios de ataque para fortalecer las defensas cibernéticas.*

*Los modelos de lenguaje de gran escala (LLM) permiten la identificación rápida de vulnerabilidades y la automatización de pruebas, reduciendo la intervención manual y mejorando la evaluación de la seguridad. Estos modelos pueden simular comportamientos de atacantes reales, aprender de patrones históricos y adaptarse a nuevas tácticas, proporcionando una mejor comprensión de las amenazas. Además, pueden ajustarse a los requisitos específicos de una organización, permitiendo un enfoque personalizado en la detección y mitigación de vulnerabilidades.*

*La Inteligencia Artificial Generativa también se utiliza para examinar correos electrónicos y detectar mensajes sospechosos, simular escenarios de ataque hipotéticos, y mejorar la identificación y mitigación de malware mediante metodologías basadas en Redes Generativas Antagónicas (GAN). Además, las plataformas de IA generativa, como Microsoft SecurityCopilot, ofrecen servicios automatizados de IA para la seguridad de red, protección contra malware y supervisión del cumplimiento de la privacidad.*

*En el ámbito de la ciberdefensa, la Gen IA ayuda a generar informes de seguridad, analizar incidentes y hacer recomendaciones estratégicas. Además, puede procesar grandes cantidades de datos para identificar amenazas potenciales, detectar fallos de seguridad y ayudar a generar código seguro. La integración de IA generativa con sistemas de detección de intrusos y la creación de honeypots son otras aplicaciones clave que mejoran la capacidad de respuesta ante amenazas.*

*En la seguridad del Internet de las Cosas (IoT), la IA generativa puede realizar análisis de comportamiento y detección de anomalías, mejorando la protección de dispositivos IoT. Además, puede detectar y prevenir deepfakes, analizar contratos inteligentes en blockchain y cazar amenazas en redes sociales. Otras herramientas específicas como Google Cloud Security AI Workbench y SentinelOne Purple AI aprovechan la IA generativa para mejorar la seguridad de datos, defenderse contra ataques avanzados y analizar amenazas cibernéticas.*

---

Fuente: Elaboración propia

**Tabla 8**

*Extracción de respuestas para la pregunta de investigación 3*

<b>Pregunta</b>	<b>RQ3: ¿Cuáles son los desafíos éticos, riesgos y limitaciones asociados con la implementación de inteligencia Artificial Generativa en Ciberseguridad?</b>
<b>Estudios</b>	<i>ES01, ES02, ES04, ES06, ES07, ES08, ES09, ES10, ES11, ES12, ES14, ES16, ES17, ES18, ES19, ES21, ES22, ES23, ES24, ES25, ES26, ES27, ES28, ES29, E30, ES31.</i>
<b>Resultados</b>	<p><i>La implementación de tecnologías de Inteligencia Artificial Generativa en ciberseguridad plantea desafíos éticos, riesgos y limitaciones. Uno de los principales problemas es la exposición de información personal y las amenazas a la seguridad. Los modelos generativos pueden ser comprometidos para manipular información, difundir desinformación y afectar procesos electorales, generar códigos de ataque dirigidos a infraestructuras críticas y comprometer la seguridad nacional.</i></p> <p><i>Un riesgo latente de Gen IA es la generación de falsos positivos y negativos, lo que conlleva a la identificación de vulnerabilidades inexistentes o la omisión de amenazas reales. Además, el uso de Gen AI plantea problemas éticos y legales, incluyendo sesgos potenciales en los modelos y la complejidad de cumplir con normativas de protección de datos. La supervisión humana es crucial para el despliegue responsable de esta tecnología.</i></p> <p><i>La adopción de IA generativa también presenta dificultades en la privacidad y la seguridad de los datos, ya que los modelos requieren grandes volúmenes de datos para su entrenamiento, incluyendo información sensible. Esto plantea riesgos significativos como el acceso no autorizado y la violación a los datos. Los adversarios pueden utilizar técnicas avanzadas, como las Redes Generativas Antagónicas (GAN), para evadir sistemas de detección de intrusos y generar tráfico hostil. Además, los modelos de IA generativa pueden ser aprovechados para generar material malicioso o fraudulento. Los modelos generativos también pueden tener dificultades para comprender el contexto y generar respuestas coherentes, lo que puede resultar en resultados sin sentido o irrelevantes. La calidad y la integridad de los datos de entrenamiento son cruciales, ya que problemas como el sesgo y el ruido pueden afectar negativamente el rendimiento del modelo.</i></p>

*La implementación de Gen AI enfrenta problemas de interpretación debido a la complejidad de sus arquitecturas y operaciones de caja negra. Algunas de las limitaciones que tienen las tecnologías basadas en Gen IA se representan en problemas de escalabilidad, altos costos computacionales y desafíos en la preservación de la privacidad. Además, la implementación y el mantenimiento de sistemas Gen AI pueden ser costosos, lo que limita su accesibilidad a empresas con suficientes recursos financieros e intelectuales.*

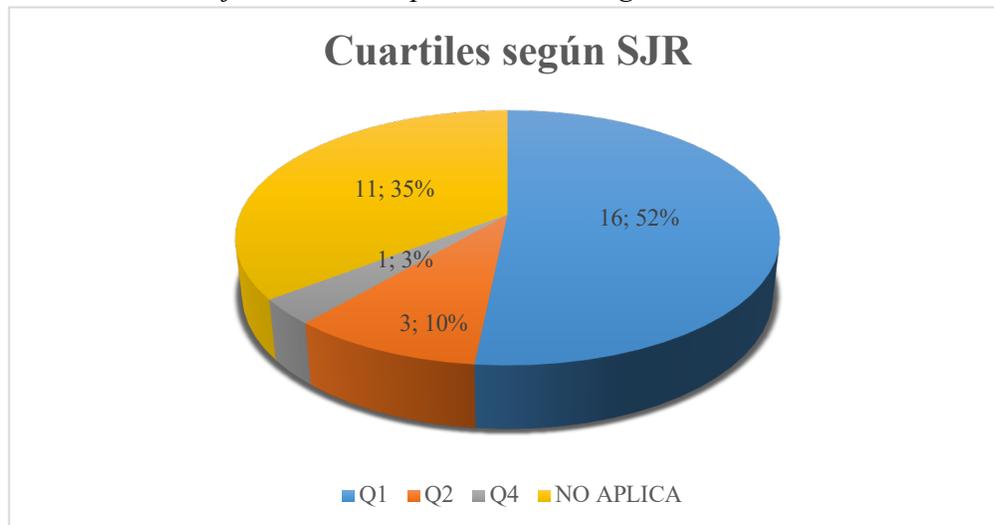
*Por último, la IA generativa plantea preocupaciones sobre la privacidad de los datos, la perpetuación de sesgos presentes en los datos de entrenamiento, la falta de transparencia en la recopilación de datos, y el control limitado de los usuarios sobre los resultados generados. Estos desafíos subrayan la necesidad de estrategias de mitigación sólidas, directrices éticas y una supervisión continua para garantizar el uso responsable de la IA generativa en ciberseguridad.*

## Resultados

La RSL obtuvo 31 estudios para responder las preguntas de investigación, de los cuales 16 encuentran calificados con cuartiles según el SJR en nivel Q1, tres artículos en nivel Q2 y uno en Q4. Respecto al resto de estudios recopilados no se encontraban calificados según el SJR. A continuación, se muestra gráficamente los artículos y su nivel de cuartil y en la Tabla 8 se detalla cuantitativamente los procesos de análisis y selección de estudios.

**Gráfico 1**

*Porcentaje de métricas para estudios según cuartiles del SJR*



Fuente: Elaboración propia

Según el gráfico 1, es evidente que los artículos seleccionados para esta revisión sistemática de la literatura (RSL) tienen altos estándares de calidad. Esto se demuestra al observar que más del 50% de los artículos están clasificados en los niveles Q1 y Q2 del índice SJR, lo cual indica que han sido publicados en revistas de prestigio y alto impacto en el ámbito académico.

**Tabla 9**

*Resultado del análisis de estudios para responder a las preguntas de investigación*

<b>Base de datos</b>	<b>Análisis 1</b>	<b>Análisis 2</b>	<b>Análisis 3</b>
WoS	6	6	3
Scopus	16	15	11
ACM	7	2	1
IEEE Xplore	30	17	10
Springer Link	16	4	3
Science Direct	26	8	3
<b>Total</b>	<b>101</b>	<b>52</b>	<b>31</b>

La metodología adoptada en este estudio refleja un proceso sistemático, esencial para garantizar la validez de la RSL. Este enfoque metódico incluye etapas clave como la planificación, la definición de criterios de inclusión y exclusión, la búsqueda, análisis y selección rigurosa de estudios, y la síntesis de datos obtenidos. Cada una de estas fases contribuye significativamente a la integridad y sistematización del trabajo, proporcionando una base sólida para encontrar los hallazgos en el campo de estudio y responder a las preguntas de investigación definidas.

## Discusión

### Estado actual de la Gen IA en ciberseguridad

La aplicación de la Gen AI en ciberseguridad ha demostrado avances significativos y ha generado preocupaciones críticas. Según (Guo et al., 2023), las herramientas como ChatGPT pueden generar contenido que mejora la seguridad de la red, simula escenarios de ciberseguridad y fomenta la educación en este ámbito. Sin embargo, esta misma capacidad puede ser aprovechada por ciberdelincuentes para desarrollar software malicioso, crear sitios web falsos y llevar a cabo ataques de ingeniería social más sofisticados. Esta dualidad también se menciona en (Deshpande & Gupta, 2023), donde se resalta que la Gen IA puede distribuir información falsa en redes sociales y manipular la opinión pública, incrementando los desafíos de ciberseguridad.

(Hilario et al., 2024) indica que los modelos de Gen IA pueden reproducir escenarios reales, facilitando el desarrollo de herramientas avanzadas para detectar vulnerabilidades de día cero y realizar pentesting. Sin embargo, la supervisión humana sigue siendo esencial para garantizar la precisión y abordar falsos positivos y negativos (Hilario et al., 2024) y (Ferrag

et al., 2023). Adicionalmente, la creciente sofisticación de la Gen IA en la creación de ransomware (Teichmann, 2023), así como también estafas de phishing y deepfakes plantea un desafío significativo, ya que las técnicas convencionales de detección están quedando obsoletas, tal como se discute en (Cambiaso & Caviglione, 2023; Pasupuleti et al., 2023; Zheng & Becker, 2023).

Además, la capacidad de la Gen IA para analizar código malicioso, detectando comportamientos y características maliciosas, incluso cuando emplean técnicas de ofuscación y cifrado, es una aplicación prometedora (Guo et al., 2023). La Gen IA también puede generar automáticamente nuevas vulnerabilidades de seguridad y escenarios de ataque, facilitando la simulación y mejorando las capacidades de defensa y respuesta (Deshpande & Gupta, 2023; Guo et al., 2023; Hilario et al., 2024).

### **Aplicaciones y técnicas**

Las aplicaciones de Gen IA en ciberseguridad son variadas y prometedoras. Según (Guo et al., 2023), utilizando capacidades de análisis y generación de texto y código, se puede lograr la encriptación automática y la transmisión segura de información. Además, ChatGPT puede analizar grandes volúmenes de datos de inteligencia amenazadora, ayudando a prevenir amenazas futuras a la seguridad de la red (Charfeddine et al., 2024).

La GenAI permite la identificación más rápida de vulnerabilidades y la automatización de la generación de escenarios de prueba, reduciendo la necesidad de intervención manual y permitiendo una evaluación más exhaustiva (Hilario et al., 2024). Puede simular el comportamiento de atacantes reales aprendiendo de patrones y adaptándose a nuevas tácticas, proporcionando una comprensión más realista de cómo pueden actuar los adversarios (Hilario et al., 2024; Okey et al., 2023; Xu et al., n.d.).

La identificación de malware se ha mejorado mediante metodologías de investigación basadas en Redes Generativas Antagónicas (Chui, 2023; Deshpande & Gupta, 2023). Herramientas como Microsoft Security Copilot proporcionan servicios automatizados de seguridad de red, protección contra malware y supervisión del cumplimiento de la privacidad (Hu & Chen, 2023). ChatGPT también se utiliza para generar mensajes de correo electrónico que engañan a los estafadores, malgastando sus recursos y prolongando la interacción con ellos (Cambiaso & Caviglione, 2023).

La IA generativa pueden analizar grandes cantidades de datos para identificar patrones y anomalías que indiquen la presencia de amenazas, generando informes y alertas basados en estos análisis (Ferrag et al., 2023). Gen AI permite a los defensores acelerar y automatizar el proceso de respuesta a incidentes, analizar incidentes de ciberseguridad, generar informes sobre incidentes y amenazas, y hacer recomendaciones estratégicas (Gupta et al., 2023; Sai et al., 2024). La generación de honeypots persuasivos para atraer a los atacantes y analizar sus métodos es otra aplicación significativa de Gen IA (Sai et al., 2024).

### **Desafíos éticos, riesgos y limitaciones**

La implementación de Inteligencia Artificial Generativa en ciberseguridad enfrenta numerosos desafíos éticos, riesgos y limitaciones. La exposición de información personal y las amenazas a la seguridad social son preocupaciones críticas. Los modelos generativos pueden aumentar los riesgos de manipulación informativa y política, afectando procesos electorales (Guo et al., 2023). El uso de Gen AI plantea problemas éticos y legales, incluyendo la posibilidad de resultados sesgados si los modelos se entrenan con datos no representativos, lo que puede generar escenarios de prueba injustos o no detectar vulnerabilidades específicas (Hilario et al., 2024).

Además, el uso de grandes cantidades de datos para entrenar modelos de IA genera preocupaciones sobre la privacidad y la seguridad de los datos, ya que la información sensible puede ser expuesta (Gill & Kaur, n.d.; Hu & Chen, 2023; Pasupuleti et al., 2023). La recopilación de grandes volúmenes de datos para entrenar modelos de IA genera preocupaciones, ya que la información sensible puede ser expuesta (Charfeddine et al., 2024; Deshpande & Gupta, 2023; Gill & Kaur, n.d.; Guo et al., 2023; Hilario et al., 2024; Hu & Chen, 2023; Pasupuleti et al., 2023). Los sistemas de IAG requieren una gran cantidad de recursos computacionales, lo que puede tener un efecto adverso en el ecosistema (Gill & Kaur, n.d.). Los desafíos también incluyen la posibilidad de que los adversarios inyecten perturbaciones en los datos de entrada para replicar modelos o perjudicar el rendimiento (Esmradi et al., n.d.; Van Huynh et al., 2023).

Las preocupaciones sobre la privacidad incluyen el robo y envenenamiento de modelos, así como la persistencia de sesgos presentes en los datos de entrenamiento (Charfeddine et al., 2024; Esmradi et al., n.d.; Ferrag et al., 2023; Golda et al., 2024; Gupta et al., 2023; Hu & Chen, 2023; Pasupuleti et al., 2023; Sai et al., n.d.; Xu et al., n.d.). La falta de transparencia y control por parte de las empresas y los usuarios también es un desafío significativo, lo que puede resultar en decisiones injustas o discriminatorias (Gupta et al., 2023; Migliorini, 2024; Ong et al., 2024; Sai et al., 2024). Los modelos de lenguaje de gran escala (LLM) presentan varios desafíos y limitaciones en ciberseguridad. Estos modelos se centran en generar respuestas, pero no siempre priorizan la exactitud, lo que puede resultar en la generación de respuestas incorrectas o engañosas, conocidas como "alucinaciones" (Scanlon et al., 2023).

La protección de la Información de Identificación Personal (PII) y otros datos sensibles es vital para mantener la seguridad y la confianza de los usuarios (Charfeddine et al., 2024). Finalmente, es crucial que investigadores, desarrolladores y responsables políticos trabajen juntos para abordar estos desafíos y garantizar la integridad y seguridad de los sistemas basados en IA. La implementación de marcos de seguridad de IA emergentes, como el "Top 10 for LLM Applications" del Open Web Application Security Project, el marco de gestión de riesgos de IA del National Institute of Standards and Technology, y el marco MITRE

Adversarial Threat Landscape for Artificial-Intelligence Systems, proporcionan orientación a los equipos de seguridad sobre mejores prácticas, controles, recomendaciones y procedimientos (Campbell & Jovanovic, 2024).

Nuestro trabajo es uno de los pioneros en realizar una revisión sistemática del estado actual de la inteligencia artificial generativa en ciberseguridad. Aunque estudios previos como (Guo et al., 2023; Van Huynh et al., 2023) han llevado a cabo revisiones exhaustivas, estos se han centrado exclusivamente en la seguridad y privacidad de la IA generativa. Otros estudios, como (Charfeddine et al., 2024; Dwivedi et al., 2023; Hu & Chen, 2023; Pasupuleti et al., 2023), se han enfocado en herramientas específicas como ChatGPT, mientras que investigaciones como (Deshpande & Gupta, 2023; Hilario et al., 2024; Ong et al., 2024; Scanlon et al., 2023; Teichmann, 2023), han explorado aplicaciones en campos específicos, como amenazas en ciberseguridad, pentesting, ransomware e investigación forense digital.

Nuestra investigación abarca una perspectiva más amplia, proporcionando una visión integral de las aplicaciones, técnicas, desafíos éticos, riesgos y limitaciones de la inteligencia artificial generativa en ciberseguridad.

Este trabajo se limita a responder las preguntas de investigación planteadas, proporcionando una visión general del estado actual, aplicaciones, técnicas, desafíos éticos, riesgos y limitaciones de la inteligencia artificial generativa en ciberseguridad. Sin embargo, existen áreas muy específicas donde la IA generativa está en su apogeo, como el Internet de las Cosas (IoT), la capa física de red, pentesting, ransomware e investigación forense digital, que no han sido exploradas en profundidad en este estudio.

Para futuras investigaciones, se recomienda elegir una de estas aplicaciones específicas detalladas en este estudio para una exploración más profunda. Particularmente, investigaciones futuras podrían enfocarse en desarrollar y evaluar marcos o normativas para el uso de la IA generativa en estos campos específicos. Además, se podrían realizar estudios empíricos para validar la eficacia de las técnicas propuestas y abordar los desafíos éticos y de privacidad de manera más exhaustiva, asegurando así un uso seguro y responsable de la inteligencia artificial generativa en ciberseguridad.

## Conclusiones

Este estudio ha realizado una revisión sistemática de la literatura sobre el estado actual en la aplicación de inteligencia artificial generativa en el ámbito de la ciberseguridad, proporcionando una visión comprensiva del estado actual, las aplicaciones prometedoras, los desafíos éticos, riesgos y limitaciones asociadas. Los hallazgos principales destacan que la Gen IA, se está utilizando tanto en la implementación de estrategias defensivas como en la ejecución de acciones ofensivas en el ámbito de la ciberseguridad. Su capacidad para generar

contenido complejo mejora la simulación de escenarios de seguridad y la detección de vulnerabilidades, pero también puede ser explotada por ciberdelincuentes para desarrollar técnicas avanzadas de ingeniería social y generar código malicioso, aumentando los riesgos de privacidad y seguridad de los datos.

La presente Revisión Sistemática de Literatura e investigación elaborada, contribuye a la comprensión de cómo la Gen IA puede transformar el campo de la ciberseguridad. Al abordar tanto las aplicaciones beneficiosas como los riesgos y desafíos, se proporciona una base sólida para futuras investigaciones que pueden explorar áreas más específicas y desarrollar marcos teóricos y normativos para el uso responsable de la Gen IA.

### Referencias bibliográficas

- Bengesí, S., El-Sayed, H., Sarker, K., Houkpati, Y., Irungu, J., & Oladunni, T. (n.d.). *Advancements in Generative AI: A Comprehensive Review of GANs, GPT, Autoencoders, Diffusion Model, and Transformers*.  
<https://towardsdatascience.com/applied-deep-learning-part-3->
- Cambiaso, E., & Caviglione, L. (2023). *Scamming the Scammers: Using ChatGPT to Reply Mails for Wasting Time and Resources*. <http://ceur-ws.org>
- Campbell, M., & Jovanovic, M. (2024). Disinfecting AI: Mitigating Generative AI's Top Risks. In *Computer* (Vol. 57, Issue 5, pp. 111–116). IEEE Computer Society.  
<https://doi.org/10.1109/MC.2024.3374433>
- Cao, Y., Li, S., Liu, Y., Yan, Z., Dai, Y., Yu, P. S., & Sun, L. (2023). *A Comprehensive Survey of AI-Generated Content (AIGC): A History of Generative AI from GAN to ChatGPT*. <http://arxiv.org/abs/2303.04226>
- Capodieci, N., Sanchez-Adames, C., Harris, J., & Tatar, U. (2024). The Impact of Generative AI and LLMs on the Cybersecurity Profession. *2024 Systems and Information Engineering Design Symposium (SIEDS)*, 448–453.  
<https://doi.org/10.1109/SIEDS61124.2024.10534674>
- Charfeddine, M., Kammoun, H. M., Hamdaoui, B., & Guizani, M. (2024). ChatGPT's Security Risks and Benefits: Offensive and Defensive Use-Cases, Mitigation Measures, and Future Implications. *IEEE Access*, 12, 30263–30310.  
<https://doi.org/10.1109/ACCESS.2024.3367792>
- Chui, K. T. (2023). *A Lightweight Generative Adversarial Network for Imbalanced Malware Image Classification*. 1–4. <https://doi.org/10.1145/3647444.3652455>
- Das, N., Kotal, A., Roseberry, D., & Joshi, A. (2023). *Change Management using Generative Modeling on Digital Twins*. <http://arxiv.org/abs/2309.12421>
- Deshpande, A. S., & Gupta, S. (2023). GenAI in the Cyber Kill Chain: A Comprehensive Review of Risks, Threat Operative Strategies and Adaptive Defense Approaches. *3rd IEEE International Conference on ICT in Business Industry and Government, ICTBIG 2023*. <https://doi.org/10.1109/ICTBIG59752.2023.10456106>



- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., ... Wright, R. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 71. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Esmradi, A., Yip, D. W., & Chan, C. F. (n.d.). *A Comprehensive Survey of Attack Techniques, Implementation, and Mitigation Strategies in Large Language Models*.
- Ferrag, M. A., Debbah, M., & Al-Hawawreh, M. (2023). *Generative AI for Cyber Threat-Hunting in 6G-enabled IoT Networks*. <http://arxiv.org/abs/2303.11751>
- Ferrara, E. (2023). *GenAI Against Humanity: Nefarious Applications of Generative Artificial Intelligence and Large Language Models*. <https://doi.org/10.1007/s42001-024-00250-1>
- Ferrara, E. (2024). GenAI against humanity: nefarious applications of generative artificial intelligence and large language models. *Journal of Computational Social Science*. <https://doi.org/10.1007/s42001-024-00250-1>
- Gill, S. S., & Kaur, R. (n.d.). *ChatGPT: Vision and Challenges*.
- Golda, A., Mekonen, K., Pandey, A., Singh, A., Hassija, V., Chamola, V., & Sikdar, B. (2024). Privacy and Security Concerns in Generative AI: A Comprehensive Survey. *IEEE Access*, 12, 48126–48144. <https://doi.org/10.1109/ACCESS.2024.3381611>
- Guo, D., Chen, H., Wu, R., & Wang, Y. (2023). AIGC challenges and opportunities related to public safety: A case study of ChatGPT. In *Journal of Safety Science and Resilience* (Vol. 4, Issue 4, pp. 329–339). KeAi Communications Co. <https://doi.org/10.1016/j.jnlssr.2023.08.001>
- Gupta, M., Akiri, C., Aryal, K., Parker, E., & Praharaj, L. (2023). From ChatGPT to ThreatGPT: Impact of Generative AI in Cybersecurity and Privacy. In *IEEE Access* (Vol. 11, pp. 80218–80245). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/ACCESS.2023.3300381>
- Hilario, E., Azam, S., Sundaram, J., Imran Mohammed, K., & Shanmugam, B. (2024). Generative AI for pentesting: the good, the bad, the ugly. *International Journal of Information Security*. <https://doi.org/10.1007/s10207-024-00835-x>
- Hu, C., & Chen, J. (2023). A Dimensional Perspective Analysis on the Cybersecurity Risks and Opportunities of ChatGPT-Like Information Systems. *Proceedings - 2023 International Conference on Networking and Network Applications, NaNA 2023*, 324–331. <https://doi.org/10.1109/NaNA60121.2023.00061>
- Katsadouros, E., Patrikakis, C. Z., & Hurlburt, G. (2023). Can Large Language Models Better Predict Software Vulnerability? In *IT Professional* (Vol. 25, Issue 3, pp. 4–8). IEEE Computer Society. <https://doi.org/10.1109/MITP.2023.3284628>

- Kitchenham, B. (2007). *Guidelines for performing Systematic Literature Reviews in Software Engineering*.
- Manterola, C., Astudillo, P., Arias, E., & Claros, N. (2013). Revisión sistemática de la literatura. Qué se debe saber acerca de ellas. *Cirugía Española*, 91(3), 149–155. <https://doi.org/10.1016/j.ciresp.2011.07.009>
- Migliorini, S. (2024). China's Interim Measures on generative AI: Origin, content and significance. *Computer Law and Security Review*, 53. <https://doi.org/10.1016/j.clsr.2024.105985>
- Moscara, E. (2019). *2019 IEEE Thesaurus Version 1.0 Created by The Institute of Electrical and Electronics Engineers (IEEE)*. <http://www.niso.org/kst/reports/standards>
- Mozo, A., Karamchandani, A., de la Cal, L., Gómez-Canaval, S., Pastor, A., & Gifre, L. (2023). A Machine-Learning-Based Cyberattack Detector for a Cloud-Based SDN Controller. *Applied Sciences (Switzerland)*, 13(8). <https://doi.org/10.3390/app13084914>
- Okey, O. D., Udo, E. U., Rosa, R. L., Rodríguez, D. Z., & Kleinschmidt, J. H. (2023). Investigating ChatGPT and cybersecurity: A perspective on topic modeling and sentiment analysis. *Computers and Security*, 135. <https://doi.org/10.1016/j.cose.2023.103476>
- Ong, J. C. L., Chang, S. Y. H., William, W., Butte, A. J., Shah, N. H., Chew, L. S. T., Liu, N., Doshi-Velez, F., Lu, W., Savulescu, J., & Ting, D. S. W. (2024). Ethical and regulatory challenges of large language models in medicine. In *The Lancet Digital Health*. Elsevier Ltd. [https://doi.org/10.1016/S2589-7500\(24\)00061-X](https://doi.org/10.1016/S2589-7500(24)00061-X)
- Pasupuleti, R., Vadapalli, R., & Mader, C. (2023). Cyber Security Issues and Challenges Related to Generative AI and ChatGPT. *Proceedings - 2023 10th International Conference on Social Networks Analysis, Management and Security, SNAMS 2023*. <https://doi.org/10.1109/SNAMS60348.2023.10375472>
- Saddi, V. R., Gopal, S. K., Mohammed, A. S., Dhanasekaran, S., & Naruka, M. S. (2024). Examine the Role of Generative AI in Enhancing Threat Intelligence and Cyber Security Measures. *2024 2nd International Conference on Disruptive Technologies, ICDT 2024*, 537–542. <https://doi.org/10.1109/ICDT61202.2024.10489766>
- Sai, S., Sai, R., & Chamola, V. (n.d.). *Received XX Month, XXXX; revised XX Month, XXXX; accepted XX Month, XXXX; Date of publication XX Month Generative AI for Industry 5.0: Analyzing the impact of ChatGPT, DALLE, and Other Models*. <https://doi.org/10.1109/OJCOMS.2024.011100>
- Sai, S., Yashvardhan, U., Chamola, V., & Sikdar, B. (2024). Generative AI for Cyber Security: Analyzing the Potential of ChatGPT, DALL-E, and Other Models for Enhancing the Security Space. *IEEE Access*, 12, 53497–53516. <https://doi.org/10.1109/ACCESS.2024.3385107>
- Scanlon, M., Breiting, F., Hargreaves, C., Hilgert, J. N., & Sheppard, J. (2023). ChatGPT for digital forensic investigation: The good, the bad, and the unknown. *Forensic*

- Teichmann, F. (2023). Ransomware attacks in the context of generative artificial intelligence—an experimental study. *International Cybersecurity Law Review*, 4(4), 399–414. <https://doi.org/10.1365/s43439-023-00094-x>
- Truong, T. C., Zelinka, I., Plucar, J., Čandik, M., & Šulc, V. (2020). Artificial Intelligence and Cybersecurity: Past, Presence, and Future. *Advances in Intelligent Systems and Computing*, 1056, 351–363. [https://doi.org/10.1007/978-981-15-0199-9\\_30](https://doi.org/10.1007/978-981-15-0199-9_30)
- Van Huynh, N., Wang, J., Du, H., Hoang, D. T., Niyato, D., Nguyen, D. N., Kim, D. I., & Letaief, K. B. (2023). *Generative AI for Physical Layer Communications: A Survey*. <http://arxiv.org/abs/2312.05594>
- Wiafe, I., Koranteng, F. N., Obeng, E. N., Assyne, N., Wiafe, A., & Gulliver, S. R. (2020). Artificial Intelligence for Cybersecurity: A Systematic Mapping of Literature. *IEEE Access*, 8, 146598–146612. <https://doi.org/10.1109/ACCESS.2020.3013145>
- Xu, H., Li, Y., Balogun, O., Wu, S., Wang, Y., & Cai, Z. (n.d.). *Security Risks Concerns of Generative AI in the IoT*.
- Zheng, S. Y., & Becker, I. (2023). Phishing to improve detection. *ACM International Conference Proceeding Series*, 334–343. <https://doi.org/10.1145/3617072.3617121>

**Conflicto de intereses:**

Los autores declaran que no existe conflicto de interés posible.

**Financiamiento:**

No existió asistencia financiera de partes externas al presente artículo.

**Agradecimiento:**

N/A

**Nota:**

El artículo no es producto de una publicación anterior.