UNIVERSIDAD NACIONAL DE UCAYALI
Revista De Investigación Universitaria
Versión electrónica 2664 – 8423
ARTICULO DE REVISIÓN

Vol. 13 N° 2, pp. 1046 – 1063, Julio/diciembre 2023
Recibido 26/09/2023
Aceptado 19/12//2023
Publicado 31/12/2023

Inteligencia artificial en la seguridad de la información en una organización

Artificial intelligence in information security in an organization

Cristhian Aldair Villacorta Vidal
T013300620@unitru.edu.pe
https://orcid.org/0009-0008-2504-7354
Universidad Nacional de Trujillo, Trujillo, Perú

Elvis Steve Ortiz Centurion
T023300220@unitru.edu.pe
https://orcid.org/0009-0001-6369-2919
Universidad Nacional de Trujillo, Trujillo, Perú

Alberto Carlos Mendoza de los Santos
amendozad@unitru.edu.pe
https://orcid.org/0000-0002-0469-915X
Universidad Nacional de Trujillo, Trujillo, Perú.

Resumen

La inteligencia artificial ha demostrado ser beneficiosa en diversos aspectos, desde la optimización de tareas hasta la sustitución de funciones profesionales. Es por ello que las organizaciones en su mayoría están siendo encaminadas en el uso de la IA para la seguridad de su información, la cuál es muy sensible y requiere un trato especial. Por otro lado, este trabajo describe un estudio basado en la metodología PRISMA, cuyo objetivo es identificar las diversas implicaciones positivas y negativas de la incorporación de la inteligencia artificial en la seguridad de la información en una compañía. Dicho esto, es necesario buscar una respuesta a: ¿Cuáles son las implicaciones de la incorporación de la inteligencia artificial en la seguridad de la información dentro de las organizaciones? Esta investigación es relevante porque enumera las siguientes virtudes: Viabilización de la automatización de respuestas, potenciación de la automatización, prevención de pérdida de datos, identificación de amenazas avanzada, privacidad y ética vinculadas a la aplicación de la IA, evaluación del comportamiento, asimismo presenta desafíos tales como: malware basado en IA, ataques de ingeniería social mejorados, ataques a la autenticación, ocultamiento del código malicioso de la detección los cuales incentivan a tener una mayor privacidad y seguridad. Dichos hallazgos serán de ayuda a diversas organizaciones para tener una visión general de aquello a lo que pueden llegar a enfrentarse con el fin de resguardar la información de la empresa.

Palabras Claves: Seguridad de la información; Protección de datos; Inteligencia artificial; Control de acceso, Aprendizaje automático

Abstract

Artificial intelligence has proven to be beneficial in various aspects, from task optimization to replacing professional functions. That is why most organizations are being guided in the use of AI for the security of their information, which is very sensitive and requires special treatment. On the other hand, this work describes a study based on the PRISMA methodology, whose objective is to identify the various positive and negative implications of the incorporation of artificial intelligence in information security in a company. That said, it is necessary to seek an answer to: What are the implications of incorporating artificial intelligence in information security within organizations? This research is relevant because it lists the following virtues: Enabling the automation of responses, enhancing automation, preventing data loss, advanced threat identification, privacy and ethics linked to the application of AI, behavioral evaluation, and also

Vol. 13 N° 2, pp. 1046 – 1063, Julio/diciembre 2023
Recibido 26/09/2023
Aceptado 19/12//2023
Publicado 31/12/2023

presents challenges such as: AI-based malware, enhanced social engineering attacks, authentication attacks, hiding malicious code from detection which encourage greater privacy and security. These findings will be helpful to various organizations to have an overview of what they may face in order to protect company information.

Keywords: Security of the information; Data Protection; Artificial intelligence; Access control, Machine learning

Introducción

La evolución de la IA ha generado una gran repercusión en la organización y en su manera de tratar su información privada, publica y confidencial. Este avance tecnológico ha permitido a las empresas recopilar, analizar y comprender grandes cantidades de información de manera eficiente y precisa. Gracias al aprendizaje automático y a la comprensión de lenguaje natural, las organizaciones extraen conocimientos valiosos de los datos, lo que les ayuda a tomar decisiones informadas y a prever tendencias emergentes en sus respectivas industrias, Raimundo & Rosário (2021).

Asimismo, la inteligencia artificial ha mejorado significativamente la protección de la información ante los elevados ataques cibernéticos a los que las empresas y los individuos están expuestos Kaur et al. (2023). Los sistemas de seguridad tradicionales a menudo se ven superados por las tácticas cada vez más sofisticadas de los ciberdelincuentes. Sin embargo, la inteligencia artificial ha evidenciado su eficacia como una herramienta efectiva en la identificación y reducción de riesgos de manera instantánea, Liu & Zhang (2023).

En este contexto, de acuerdo con Solís C. (2023), se sostiene que una de las principales ventajas de la IA en el campo de la seguridad de la información reside en su capacidad para analizar volúmenes extensos de datos en tiempo real. La mayoría de los sistemas de seguridad convencionales se fundamentan en normas preestablecidas y patrones conocidos para identificar eventuales riesgos. No obstante, los actores maliciosos en línea están constantemente ideando nuevas estrategias y tácticas para eludir estas medidas, lo que implica que las pautas predefinidas podrían ser insuficientes para salvaguardar contra las amenazas más avanzadas.

El propósito de esta investigación es analizar la inclusión de la inteligencia artificial, en consonancia con su influencia en la sociedad contemporánea, en relación con la salvaguarda de la información en contextos organizativos. La intención fue reunir datos pertinentes sobre cómo la inclusión de la inteligencia artificial ha repercutido en la seguridad de la información, estableciendo conexiones con el uso del blockchain y otras soluciones. Se tuvo en cuenta tanto los beneficios como los obstáculos que surgen al implementar esta tecnología. El propósito

Vol. 13 N° 2, pp. 1046 – 1063, Julio/diciembre 2023
Recibido 26/09/2023
Aceptado 19/12//2023
Publicado 31/12/2023

principal fue obtener una perspectiva integral sobre cómo la inteligencia artificial está remodelando la seguridad de la información en las organizaciones, en línea con sus amplias influencias en diversas áreas.

Materiales y métodos

Tipo de estudio: Se empleó la metodología PRISMA, desarrollada por Urrútia y Bonfill en 2010, al abordar la siguiente interrogante: ¿Cuáles son las implicaciones de la incorporación de la inteligencia artificial en la seguridad de la información dentro de las organizaciones?

Fundamentación de la Metodología: Mediante un exhaustivo examen de la literatura científica siguiendo el enfoque metodológico PRISMA (Preferential Reporting Items for Systematic Reviews and Meta-Analyses), según lo documentado por Beal V. (2021). La metodología PRISMA posibilita la realización de un análisis sistemático de datos respaldado por trabajos previos de otros investigadores, con el objetivo de abordar la pregunta planteada y proporcionar una respuesta fundamentada.

Asimismo, Cabrera E. et al. (2004) sostiene que el enfoque PRISMA se presenta como un método interactivo de condensación enfocado en la síntesis informativa. Aquí, el usuario se involucra con el sistema a través de listas de palabras que denotan ideas centrales, generadas automáticamente mediante técnicas de análisis sintáctico básico. PRISMA facilita la exhibición de estas ideas centrales, el acceso a detalles vinculados a ellas, la presentación completa de documentos y la revelación de las principales proposiciones de cada oración en el texto.

De acuerdo con lo mencionado por Carmen P. (2012), en el mes siete del año 2009 se lanzó la declaración PRISMA, que representa una actualización diseñada con fines educativos para complementar la lista de verificación. Esta declaración respalda los 27 componentes de la lista de verificación y también incluye siete tablas que detallan aspectos esenciales de la metodología. Además, es importante destacar que PRISMA no solo se limita a los metaanálisis de ensayos clínicos, sino que se puede adaptar a una variedad de revisiones sistemáticas. Según la autora, PRISMA se concibe como una herramienta beneficiosa para mejorar los estudios en términos de contextos, intervenciones y otros aspectos relevantes.

Considerando las definiciones, a continuación, se describirán las fases posteriores del proceso de investigación, que comprenden la formulación de la estrategia de búsqueda, la selección de la literatura pertinente, el registro de los hallazgos y, finalmente, la interpretación de los resultados obtenidos.

Vol. 13 N° 2, pp. 1046 – 1063, Julio/diciembre 2023
Recibido 26/09/2023
Aceptado 19/12//2023
Publicado 31/12/2023

Criterios de Inclusión, exclusión y calidad: Los criterios de inclusión comprenden la incorporación de información en los idiomas inglés, portugués y español que esté vinculada con los resultados de la inteligencia artificial en el ámbito de la seguridad de información durante el periodo más reciente de los últimos cuatro años (2019-2023). Asimismo, se incluirán estudios empíricos, revisiones sistemáticas y artículos de opinión de fuentes confiables y reconocidas en el ámbito de la IA y la seguridad en la información.

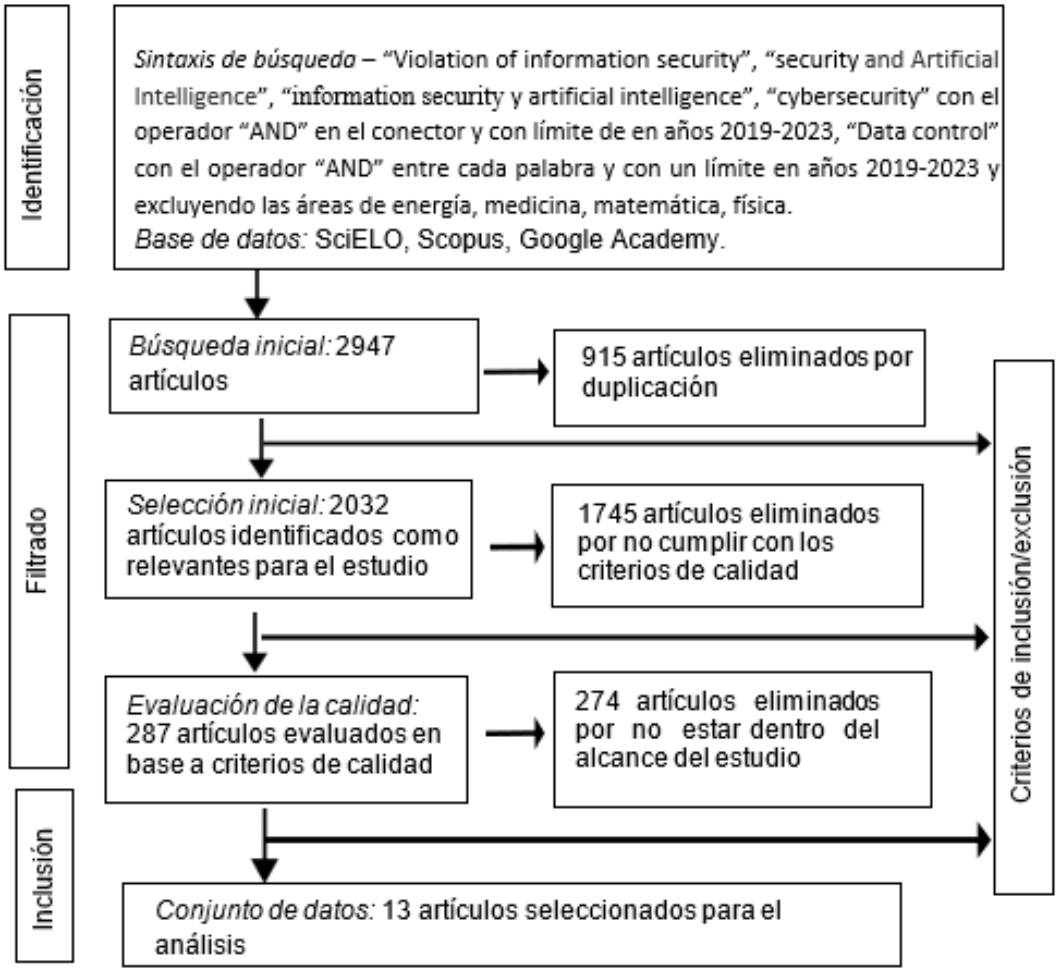

***Figura 01*:** Proceso de Extracción de Datos

Los criterios de exclusión son: Información relacionada indirectamente con la incorporación de la IA en la seguridad de información, así como publicaciones anteriores al período especificado (antes de 2019). Se excluirán también fuentes no confiables, como sitios web no verificables y blogs sin respaldo científico.

UNIVERSIDAD NACIONAL DE UCAYALI
Revista De Investigación Universitaria
Versión electrónica 2664 – 8423
ARTICULO DE REVISIÓN

Vol. 13 N° 2, pp. 1046 – 1063, Julio/diciembre 2023
Recibido 26/09/2023
Aceptado 19/12//2023
Publicado 31/12/2023

Criterios de calidad: Se considerarán para la selección aquellos estudios que hayan sido publicados en revistas científicas de renombre, conferencias académicas reconocidas o plataformas oficiales de instituciones de investigación. Además, se llevará a cabo una evaluación exhaustiva de la pertinencia de los hallazgos y las conclusiones proporcionadas respecto a las diversas implicaciones de la inteligencia artificial en la seguridad de la información. Lo mencionado anteriormente fortalecen la robustez y la confiabilidad de los resultados alcanzados. Estos estándares de calidad refuerzan la solidez y credibilidad de los resultados conseguidos.

La exploración y obtención de datos se realizaron de manera personal, y cualquier divergencia identificada entre los participantes fue solventada a través de un acuerdo mutuo. Esto tuvo como propósito realizar una revisión sistemática.

Tabla 01

La adquisición de las investigaciones se llevó a cabo a través de la exploración en buscadores académicos y sus herramientas de búsqueda correspondientes.

Motor Búsqueda	Términos de Búsqueda	Resultados	Seleccionados
Scielo	(Violation of information security) AND (artificial intelligence) AND (Security) AND year_cluster:("2022" OR "2021" OR "2020" OR "2019")	7	1
Scopus	(TITLE-ABS-KEY (security AND cybersecurity) AND TITLE-ABS-KEY (security AND monitoring) AND TITLE-ABS-KEY (Information AND science) AND TITLE-ABS-KEY (Protection AND Detection) AND TITLE-ABS-KEY (Data AND control) AND TITLE-ABS-KEY (Internet AND of AND things)) AND PUBYEAR > 2018 AND PUBYEAR < 2024 AND (LIMIT-TO (LANGUAGE, "English") OR LIMIT-TO (LANGUAGE, "Spanish")) AND (EXCLUDE (SUBJAREA, "MEDI") OR EXCLUDE (SUBJAREA, "MATH") OR EXCLUDE (SUBJAREA, "ENER")) AND (LIMIT-TO (EXACTKEYWORD, "Artificial Intelligence") OR LIMIT-TO (EXACTKEYWORD, "Security Of Data") OR LIMIT-TO (EXACTKEYWORD, "Network Security"))	2105	9
Google Academy	("artificial intelligence in security" AND "cybersecurity" AND "security information" AND “ Data control ” Chemistry-Medicine -Energy)	835	3

*Proceso de recolección de información:*Con el propósito de iniciar la resolución de la pregunta propuesta, en primer lugar, se identificó algunos términos de importancia. Como instancias: "Artificial intelligence”, “IT security”, “cybersecurity”, “information science”, “Data control”.

UNIVERSIDAD NACIONAL DE UCAYALI
Revista De Investigación Universitaria
Versión electrónica 2664 – 8423
ARTICULO DE REVISIÓN

Vol. 13 N° 2, pp. 1046 – 1063, Julio/diciembre 2023
Recibido 26/09/2023
Aceptado 19/12//2023
Publicado 31/12/2023

En otro aspecto, el procedimiento prisma se inicia mediante la indagación de registros en distintas bases de datos, luego sigue con la no consideración de repeticiones y concluye con investigaciones que abarcan análisis cualitativos y cuantitativos (análisis exhaustivos y síntesis de datos). Urrútia G. y Bonfill X. (2010). Las bases de datos fueron seleccionadas para esta minuciosa revisión en virtud de su extendido empleo en diversas indagaciones sistemáticas, así como por la copiosa cantidad de textos académicos en los que se hallan consignadas. En el siguiente apartado, se procederá a exponer las bases de datos seleccionadas:

Resultados y discusión

Tabla 02

Contribución de cada artículo seleccionado.

N°	Título	Autor(es) y año	Aporte
1	Antecedent factors of violation of information security rules	Alexandre Cappellozza, Gustavo Hermínio Salati Marcondes de Moraes, Gilberto Perez y Alessandra Lourenço Simões (2021)	Cappellozza et al. (2021) resaltan los efectos positivos de la Inteligencia Artificial (IA) en varios aspectos cruciales de la seguridad de la información. La IA se destaca al identificar comportamientos éticos, salvaguardar datos críticos, proporcionar formación en ciberseguridad, supervisar el cumplimiento de políticas, evaluar riesgos y garantizar la privacidad. Estas contribuciones de la IA tienen el potencial de fortalecer significativamente la seguridad de la información, lo que se traduce en mejoras sustanciales en la ciberseguridad en el entorno empresarial digital.
2	Deep learning technology of computer network security detection based on artificial intelligence	Qinghui Liu y Tianping Zhang (2023)	Liu & Zhang (2023) plantean que la inteligencia artificial se utiliza para administrar la seguridad de las redes de computadoras mediante la implementación de métodos de aprendizaje profundo, la identificación de intrusiones y el análisis de medidas de seguridad. Esta implementación ha mejorado la protección de las redes, permitiendo la efectiva prevención y solución de diversas amenazas.
3	An artifcial intelligence lightweight blockchain security model for security and privacy in IIoT systems	Shitharth Selvarajan, Gautam Srivastava, Alaa O. Khadidos, Adil O. Khadidos, Mohamed Baza,Ali Alshehri y Jerry Chun-Wei Lin	(Shitharth et al., 2023) plantea un enfoque de seguridad basado en un modelo de cadena de bloques liviano impulsado por IA (AILBSM). Este método se respalda en el uso de sistemas de Inteligencia Artificial (IA) para proteger la privacidad y mantener la integridad de los sistemas de la Internet Industrial de las Cosas (IIoT). En esencia, este modelo fusiona los beneficios de las técnicas de IA apoyadas en cadenas de bloques livianas junto con la red neuronal optimizada para coexistencia conocida como COSNN (Convivencia Optimizada de Red Neuronal Sprinter) para optimizar las funciones de seguridad.

Vol. 13 N° 2, pp. 1046 – 1063, Julio/diciembre 2023
Recibido 26/09/2023
Aceptado 19/12//2023
Publicado 31/12/2023

		(2023)	
4	Artificial intelligence for cybersecurity: Literature review and future research directions	Ramanpreet Kaur, Duˇsan Gabrijelčič y Tomaˇz Klobučar (2023)	Kaur et al. (2023) destacan que la inteligencia artificial (IA) puede aportar ventajas significativas en áreas clave de la ciberseguridad, incluyendo el aprendizaje automático, la minería de datos para tomar decisiones basadas en patrones de comportamiento, el análisis predictivo, la categorización de información, la mejora de la inteligencia de amenazas y la colaboración, así como la prevención y detección de ataques cibernéticos en constante evolución mediante técnicas avanzadas de IA.
5	Analysis of security and data control in smart personal assistants from the user's perspective	Cayetano Valero , Jaime Pérez, Sonia Solera-Cotanilla, Mario Vega-Barbas, Guillermo Suarez-Tangi, Manuel Álvarez-Campana y Gregorio López (2023)	El incremento en la utilización de asistentes personales inteligentes (IPA) ha generado inquietudes en relación a la seguridad de la información. Investigaciones recientes señalan cuestiones de seguridad en las interfaces de programación de aplicaciones (API) de estos sistemas, que incluyen la susceptibilidad a ataques de suplantación de voz, deficiencias en el control de admisión, la capacidad de acceder a información personal sin la debida autorización y la carencia de ajustes específicos para usuarios jóvenes. Los IPA son vulnerables a ataques que involucren la falsificación de voces, lo que facilita a individuos con tonos similares acceder a datos personales o hacerse pasar por el usuario. Las debilidades en los mecanismos de activación y en los procedimientos de autorización inseguros pueden resultar en accesos no autorizados, permitiendo acciones fraudulentas (López G. et al., 2023).
6	The Impact and Limitations of Artificial Intelligence in Cybersecurity: A Literature Review	Meraj Farheen Ansari, Bibhu Dash, Pawankumar Sharma y Nikhitha Yathiraju (2022)	(Ansari et al., 2022) sugieren que las empresas pueden enfrentar el reto de las falsas alarmas generadas por sistemas de Inteligencia Artificial en el campo de la seguridad de la información mediante la utilización de enfoques basados en firmas y técnicas de aprendizaje automático. El enfoque basado en firmas implica que la Inteligencia Artificial detecta ciberataques y software malicioso al comparar códigos de firmas, permitiendo al equipo de seguridad cibernética contrarrestar los ataques mediante la coincidencia de firmas con ataques anteriores o bases de datos existentes. Por otro lado, el enfoque de aprendizaje automático habilita a la Inteligencia Artificial para analizar registros y paquetes de datos en redes, identificando ataques de manera eficaz y rápida. Esto le permite anticipar comportamientos sospechosos y responder eficientemente a amenazas potenciales. Además, los algoritmos de aprendizaje automático utilizados para detectar anomalías en las cuentas de usuarios contribuyen a proteger los sistemas contra amenazas internas, lo que también favorece una respuesta más eficaz ante incidentes de seguridad.
7	Role of Artificial Intelligence in Smart Cities for	Bibhu Dash y Pawankumar Sharma	La gestión de riesgos en sistemas de IA en entornos urbanos avanzados es esencial para salvaguardar la seguridad y proteger datos e infraestructura, implica realizar evaluaciones de impacto en la protección de información para identificar riesgos y desarrollar

Vol. 13 N° 2, pp. 1046 – 1063, Julio/diciembre 2023
Recibido 26/09/2023
Aceptado 19/12//2023
Publicado 31/12/2023

	Information Gathering and Dissemination (A Review)	(2022)	estrategias de mitigación, implementar medidas de ciberseguridad como firewalls asimismo programas antivirus, establecer supervisión humana para garantizar la transparencia y detectar manipulaciones indebidas, cumplir con regulaciones como el GDPR además definir políticas de manejo seguro de la información, proporcionando formación a los usuarios para concienciar sobre los riesgos de manipulación de datos (Dash & Sharma, 2022).
8	The Impact of Artificial Intelligence on Data System Security: A Literature Review	Ricardo Raimundo y Alberico Rosário (2021)	Raimundo & Rosário (2021) abordan que IA desempeña un papel crucial en diversas áreas de la seguridad, incluyendo la capacitación mediante simulaciones interactivas y entrenamientos virtuales, donde los empleados pueden practicar situaciones de seguridad. Además, la IA se utiliza para analizar grandes conjuntos de datos, identificando patrones y tendencias que revelan áreas de mejora asimismo riesgos potenciales. También se implementa en la asistencia virtual y chatbots para responder a preguntas sobre seguridad, proporcionando acceso instantáneo a información.
9	Knowledge in the grey zone: AI and cybersecurity	Tim Stevens (2022)	La inteligencia artificial, especialmente el aprendizaje automático, ha demostrado ser esencial en esta transición al permitir la identificación de amenazas desconocidas y la detección de patrones que no pueden ser capturados mediante análisis basados en firmas, fortaleciendo así la capacidad de las organizaciones para proteger sus activos digitales en un entorno en constante evolución.
10	Threats and Opportunities with AI-based Cyber Security Intrusion Detection: A Review	Bibhu Dash, Meraj Farheen Ansari, Pawankumar Sharma and Azad Al (2022)	Resalta la importancia de la detección temprana de amenazas de seguridad cibernética y cómo la inteligencia artificial puede ser utilizada para mejorar la eficiencia de los sistemas. de detección de intrusiones.
11	The State-of-the-Art in AI-Based Malware Detection Techniques: A Review	Adam Wolsey (2022)	La seguridad cibernética ha avanzado significativamente, pero al mismo tiempo, el desarrollo de malware también ha evolucionado. Se ha observado un aumento en la adopción rápida de la inteligencia artificial por parte de los ciberdelincuentes como una herramienta para crear malware más sofisticado, desafiando así los algoritmos de inteligencia artificial diseñados para proteger contra estas amenazas.
12	Artificial Intelligence Crime: An Overview of Malicious Use and Abuse of AI	Oskar Josef Gstrein; Andrej Zwitter (2022)	Sugiere cuatro tipos de uso malicioso de la IA: Ingeniería social Desinformación/noticias falsas Piratería informática Sistemas de armas autónomos
13	An Overview of Artificial Intelligence Used in Malware	Lothar Fritsch, Aws Jaber & Anis Yazidi (2022)	El malware potenciado por inteligencia artificial ha avanzado en su capacidad para eludir y atacar sistemas de seguridad respaldados por IA. También se ha desarrollado malware orientado al sabotaje ciber físico utilizando técnicas de aprendizaje no supervisado. Además, se ha observado la existencia de malware que se aloja dentro de redes

Vol. 13 N° 2, pp. 1046 – 1063, Julio/diciembre 2023
Recibido 26/09/2023
Aceptado 19/12//2023
Publicado 31/12/2023

			neuronales con objetivos distintos.

Implicaciones de la IA en la seguridad de la información

Se exponen las consecuencias favorables resultantes de la ejecución de la Inteligencia Artificial en la seguridad de la información dentro de una organización, organizadas según los artículos examinados:

Viabilización de la automatización de respuestas:

Esto lleva consigo la idea de que las tareas monótonas y recurrentes pueden ser manejadas por sistemas de IA, permitiendo que los profesionales en seguridad dirijan su atención hacia actividades más elaboradas y estratégicas. (Cappellozza, et al., 2021) (Kaur, et al., 2023) (Ansari, et al., 2022).

Potenciación de la automatización:

Esto conlleva la idea de que la IA tiene la capacidad de examinar pautas de conducta y datos biométricos para fortificar la confirmación de identidad, lo que a su vez reduce las vulnerabilidades originadas por contraseñas poco seguras o que han sido sustraídas. (Liu & Zhang, 2023) (Shitharth et al., 2023), (Raimundo & Rosário, 2021)

Prevención de pérdida de datos:

Esto implica que la IA es capaz de supervisar y examinar el flujo de datos para identificar pautas anómalas y potenciales filtraciones de información, lo que contribuye a evitar la pérdida de datos de importancia crucial. (Kaur et al., 2023), (Ansari et al., 2022).

La identificación de amenazas avanzada:

Conlleva la idea de que la IA tiene la capacidad de reconocer pautas de conducta maliciosa que podrían no ser detectadas por enfoques de seguridad convencionales, lo que resulta en una mejora en la pronta detección de ataques. (Liu & Zhang, 2023) (Kaur et al., 2023), Stevens, 2022), (Dash et al., 2022)

Privacidad y ética vinculadas a la aplicación de la IA:

Esto se reflejan en el hecho de que, a pesar de los beneficios que la IA puede ofrecer, también plantea preocupaciones relacionadas con la sensibilidad de la información y la posible opción de

Vol. 13 N° 2, pp. 1046 – 1063, Julio/diciembre 2023
Recibido 26/09/2023
Aceptado 19/12//2023
Publicado 31/12/2023

que se produzcan sesgos y discriminación en los sistemas automatizados. (Cappellozza et al., 2021), (López G. et al., 2023), (Raimundo & Rosário, 2021).

Evaluación del comportamiento:

Gracias a la IA es factible realizar un análisis de los comportamientos de usuarios y sistemas con el fin de detectar acciones. inusuales o sospechosas, lo cual amplía la aptitud para detectar y reaccionar frente a posibles amenazas. (Liu & Zhang, 2023), (Ansari et al., 2022)

Por otro lado, a medida que las medidas de protección contra los ataques de software malicioso se tornan más avanzadas, inteligentes y eficaces, los ciberdelincuentes continúan ideando métodos cada vez más ingeniosos para eludirlas. Esto conlleva diversas implicaciones negativas, entre las que se destacan:

Malware basado en IA:

En el presente momento, una de las principales inquietudes del malware es mantenerse oculto para evitar ser detectado y burlar las soluciones de seguridad informática.

Con la implementación de la IA en las medidas de seguridad contra ataques de malware, los criminales cibernéticos pueden utilizar cualquier método a su disposición para sabotear estos esfuerzos. Tal como menciona (Wosley, 2022) métodos como la manipulación de conjuntos de datos empleados por sistemas de inteligencia artificial o la ingeniería inversa de modelos de aprendizaje automático pueden ser utilizados para eludir las soluciones de seguridad.

No obstante (S. Chen et al., 2018) señala que existen diversas técnicas de aprendizaje profundo, como las redes generativas adversarias (GAN), para generar malware malicioso que logra evadir la detección al simular el comportamiento de aplicaciones legítimas. Asimismo, otra técnica de aprendizaje profundo en la que el malware tenía la capacidad de identificar su objetivo mediante el reconocimiento de voz, la ubicación geográfica y el reconocimiento facial antes de llevar a cabo un ataque.

Además, como una mejora adicional, combinan una red neuronal con una red basada en enjambres de inteligencia artificial para crear un virus de enjambre neuronal con la capacidad de eludir la detección.

Ataques de Ingeniería Social Mejorados:

UNIVERSIDAD NACIONAL DE UCAYALI
Revista De Investigación Universitaria
Versión electrónica 2664 – 8423
ARTICULO DE REVISIÓN

Vol. 13 N° 2, pp. 1046 – 1063, Julio/diciembre 2023
Recibido 26/09/2023
Aceptado 19/12//2023
Publicado 31/12/2023

Los ataques de ingeniería social emplean tácticas engañosas con el fin de influir en personas y lograr que divulguen información confidencial o personal (Mouton ntain, 2016).

Asimismo (Blauth, 2022), menciona que, de forma análoga, los agresores pueden simular ser individuos o compañías de confianza para persuadir a la víctima a abrir un correo electrónico o hacer clic en un enlace con el propósito de obtener información. Esta táctica, denominada "phishing", también puede ser refinada mediante el uso de inteligencia artificial para ampliar al máximo su impacto y alcance.

Tal como indica (Seymour, 2016) Según un estudio realizado, los resultados revelan que la naturaleza de estas plataformas puede favorecer la utilización de texto generado de manera automática para llevar a cabo ataques de phishing. La inteligencia artificial puede contribuir, en el contexto actual, a una de las fundamentales preocupaciones relacionadas con este tipo de ataques en plataformas de redes sociales se debe al hecho de que las publicaciones. Suelen redactarse de manera informal, con errores ortográficos y gramaticales ocasionales, además de utilizar enlaces acortados.

Ataques a la autenticación:

(Fritsch, 2022) indica que los sensores de los dispositivos (micrófono, acelerómetro) se utilizaron en combinación con modelos de IA con la intención de extraer PIN, contraseñas y patrones. Ejecutando la extracción de contraseña a través del móvil empleando algoritmos de inteligencia artificial tales como clasificador, bosque aleatorio. Asimismo, la extracción de PIN y números de tarjetas de crédito a través del móvil empleando micrófonos de teléfono con ayuda de análisis de voz y tonos basado en un modelo.

Ocultamiento del código malicioso de la detección:

Ocultar código malware como carga útil dentro de modelos de IA que cumplen otras funciones. Por ejemplo, redes neuronales para reconocimiento facial (Fritsch, 2022). Una de las ventajas de esta estrategia es que los modelos de inteligencia artificial suelen ser difíciles de examinar en busca de código malicioso debido a su complejidad interna. Esto podría permitir que el malware pase inadvertido para las soluciones de seguridad.

Por ejemplo, en el contexto del reconocimiento facial, el modelo de IA podría continuar desempeñando su función principal, que es identificar rostros de manera normal, pero al mismo tiempo podría llevar a cabo acciones no autorizadas, como el robo de datos biométricos o la recopilación de información confidencial.

Vol. 13 N° 2, pp. 1046 – 1063, Julio/diciembre 2023
Recibido 26/09/2023
Aceptado 19/12//2023
Publicado 31/12/2023

Principales hallazgos:

Los hallazgos derivados de los estudios escogidos ofrecen una perspectiva acerca de la incorporación de la IA en la salvaguardia de datos en una entidad organizativa. En las siguientes líneas, se expondrá un panorama general y categorización abarcadora de todos ellos:

Mejora en la Detección de Amenazas:

La IA puede identificar patrones de comportamiento y amenazas cibernéticas de manera más efectiva que los métodos tradicionales.

Stevens (2022) menciona que la IA y algoritmos están siendo utilizados para identificar amenazas cibernéticas, marcando un cambio de enfoque de firmas a la detección de anomalías. Este enfoque busca patrones y relaciones que el análisis de firmas no puede identificar por sí solo. Asimismo, (Dash et al., 2022) señala que el reconocer patrones de tráfico de red mediante algoritmos de aprendizaje automático genera seguridad y estabilidad.

Dichos patrones si se consideran fuera de lo normal, puede alertar a los administradores para investigar a fondo el posible incidente.

Colaboración e Inteligencia de Amenazas:

Kaur et al. (2023) investigaron el potencial de la IA en el contexto de las plataformas de inteligencia de amenazas, destacando cómo esta tecnología agiliza la colaboración entre profesionales y el intercambio puntual de información sobre amenazas. Asimismo, detallaron cómo la inteligencia artificial refuerza la capacidad para identificar y prevenir ataques cibernéticos en constante cambio.

Gestión de Riesgos y Cumplimiento Normativo:

Dash y Sharma (2022) investigaron la administración de riesgos en sistemas de inteligencia artificial dentro de contextos urbanos avanzados. Describieron la relevancia de llevar a cabo evaluaciones de impacto para salvaguardar los datos, la implementación de medidas de ciberseguridad y la introducción de supervisión humana como garantía de seguridad para los datos e infraestructuras.

Formación y Concientización en Seguridad:

Raimundo y Rosário (2021) propusieron estrategias para mejorar la formación en seguridad entre los empleados mediante la implementación de simulaciones interactivas, entrenamientos virtuales

Vol. 13 N° 2, pp. 1046 – 1063, Julio/diciembre 2023
Recibido 26/09/2023
Aceptado 19/12//2023
Publicado 31/12/2023

y asistentes digitales. Asimismo detallaron cómo la inteligencia artificial puede personalizar la formación según las necesidades individuales, brindando orientación efectiva para reforzar la comprensión en seguridad.

Identificación y Prevención de Amenazas:

Liu & Zhang (2023) presentaron una estrategia orientada a identificar amenazas de seguridad en redes informáticas mediante la implementación de tecnologías de aprendizaje profundo e inteligencia artificial. En su estudio, detallaron cómo estas herramientas son efectivas en la detección de actividades maliciosas, fortaleciendo así la protección contra ciberataques en una variedad de escenarios, que incluyen drones, plataformas de redes sociales y dispositivos del Internet de las Cosas (IoT).

Kaur et al. (2023) introdujeron una perspectiva que enfatiza múltiples dominios en el ámbito de la ciberseguridad que pueden obtener ventajas sustanciales de la aplicación de la inteligencia artificial. Entre estos beneficios se incluyen el análisis de pautas de conducta y la identificación de amenazas en evolución constante. Además, sugirieron que la inteligencia artificial posee la aptitud de contribuir a la prevención y detección de ataques cibernéticos al emplear enfoques avanzados asimismo presentar una colaboración en plataformas de inteligencia de amenazas.

Aplicaciones en Blockchain y IIoT:

Shitharth et al. (2023) introdujeron un método de seguridad que se fundamenta en un modelo de cadena de bloques ligero impulsado por inteligencia artificial, elaborado de manera particular con el propósito de. el entorno del Internet Industrial de las Cosas (IIoT). En su análisis, explicaron cómo esta propuesta combina técnicas de inteligencia artificial y cadenas de bloques para fortalecer la protección de la confidencialidad asimismo la integridad de los sistemas IIoT.

Mejora de la Respuesta a Incidentes:

Ansari et al. (2022) detallaron cómo la inteligencia artificial optimiza la capacidad de respuesta ante incidentes de seguridad al administrar la base de datos de vulnerabilidades, emitir alertas ante intentos de ataque en tiempo real e identificar irregularidades en las cuentas de usuarios para resguardar sistemas de amenazas internas.

En cuanto a las diversas implicaciones obtenidas, producto de la integración de la IA en la protección de información, se observa que existen diversos enfoques abordados por los investigadores.

Vol. 13 N° 2, pp. 1046 – 1063, Julio/diciembre 2023
Recibido 26/09/2023
Aceptado 19/12//2023
Publicado 31/12/2023

Según los diversos aportes, es evidente que la integración de la Inteligencia Artificial (IA) en la protección de la información dentro de las organizaciones abarca una serie de implicaciones. En palabras de Kaur et al. (2023), se destaca el papel crucial de la IA en agilizar la colaboración entre profesionales de seguridad y el intercambio de información sobre amenazas, permitiendo una respuesta coordinada ante ataques cibernéticos en constante evolución. Estos autores también enfatizan cómo la IA refuerza la capacidad para identificar y prevenir dichos ataques.

A esto se suma, Dash y Sharma (2022) resaltan la relevancia de la administración de riesgos y el acatamiento de regulaciones., y cómo la IA contribuye a la identificación y mitigación de riesgos mediante evaluaciones de impacto y medidas de ciberseguridad. En este contexto, la implementación de supervisión humana más aún el cumplimiento de regulaciones son aspectos críticos con el fin de mantener la fidelidad de los datos y la infraestructura., como lo plantean los autores.

Por otra parte, el estudio de Raimundo y Rosário (2021) aborda la mejora de la formación y la concientización en seguridad gracias a la IA. Estos autores sugieren estrategias para optimizar la formación en seguridad a través de simulaciones interactivas, entrenamientos virtuales y asistentes digitales personalizados. No obstante, (Dash et al., 2022), menciona el uso de algoritmos de aprendizaje automático con el propósito de fortalecer

Estas herramientas, según estos autores, fortalecen la comprensión de los conceptos de seguridad asimismo aumentan la conciencia sobre los riesgos.

Por consiguiente, Liu y Zhang (2023) y Kaur et al. (2023) concuerdan en la capacidad de la IA para identificar y prevenir amenazas de seguridad. Liu y Zhang presentan un enfoque basado en técnicas de aprendizaje profundo e IA para identificar amenazas en diversos escenarios, mientras que Kaur et al. sugieren que la IA contribuye a la prevención al igual que la detección de ataques cibernéticos mediante enfoques avanzados y colaboración en plataformas de inteligencia de amenazas. Por otro lado, Stevens (2022) señala que un cambio de enfoque de firmas a la detección de anomalías ayuda a identificar amenazas de seguridad de la información implica el reconocimiento y análisis de posibles riesgos y vulnerabilidades que podrían afectar la integridad, disponibilidad y confidencialidad de la información y sistemas de una organización.

Por otro lado, en el ámbito de las aplicaciones en blockchain y el Internet Industrial de las Cosas (IIoT), Shitharth et al. (2023) proponen un modelo de cadena de bloques impulsado por IA para fortalecer la protección de la confidencialidad e integridad en sistemas IIoT. Estos autores

Vol. 13 N° 2, pp. 1046 – 1063, Julio/diciembre 2023
Recibido 26/09/2023
Aceptado 19/12//2023
Publicado 31/12/2023

destacan cómo esta combinación de técnicas mejora la seguridad en un entorno donde la confianza en los datos es crucial.

En consonancia con estos beneficios, también se presentan desafíos éticos y de privacidad, como señalan Cappellozza et al. (2021), López G. et al. (2023) y Raimundo y Rosário (2021). Estos autores subrayan la importancia de abordar inquietudes sobre la confidencialidad de los datos y la posible discriminación en sistemas automatizados, a medida que la IA se convierte en una parte integral de la seguridad cibernética.

Sin embargo, esta revisión tiene limitaciones, como la inclinación hacia países desarrollados y la falta de acceso abierto a algunos artículos, destaca la relevancia actual al explorar el impacto de la IA en la seguridad de la información. Esto resalta la importancia de la IA en la seguridad de datos y la gestión de riesgos en el entorno digital actual, lo que sugiere la necesidad de investigaciones continuas en este campo.

Conclusiones

La incorporación de la inteligencia artificial en la seguridad de información en las organizaciones se ha revelado como un factor de gran relevancia en el ámbito de la seguridad cibernética y la gestión de riesgos. La IA no solo optimiza la detección de amenazas, sino también facilita la colaboración entre expertos en seguridad reforzando la capacitación en aspectos relacionados con la seguridad, asimismo desempeña un papel crucial en la identificación mejorando la prevención de ataques cibernéticos. Además, en contextos específicos como blockchain y el Internet Industrial de las Cosas (IIoT), la IA consolida la preservación de la confidencialidad y la integridad de los sistemas.

Sin embargo, la creciente sofisticación de los ciberataques basados en IA plantea preocupaciones graves, ya que los ciberdelincuentes aprovechan las capacidades de la IA para desarrollar malware más sigiloso y recurrir a tácticas de ingeniería social cada vez más persuasivas. Además, la recolección y el análisis masivo de datos por parte de sistemas de IA ponen en riesgo la confidencialidad de la información ya que dar lugar a la discriminación en proceso de decisiones automatizada.

En este contexto de constante evolución, la integración de la IA en la seguridad de la información representa un avance prometedor, pero que necesita una supervisión y regulación adecuada para garantizar su uso beneficioso y ético.

Vol. 13 N° 2, pp. 1046 – 1063, Julio/diciembre 2023
Recibido 26/09/2023
Aceptado 19/12//2023
Publicado 31/12/2023

No obstante, se presentaron algunas limitaciones como el no acceso a artículos, ya que no eran de libre acceso, asimismo los siguientes trabajos de investigación tendrán como base los diversos hallazgos para a partir de ellos revisar algoritmos para fortalecer las virtudes y prevenir aquellos desafíos en lo que respecta a la seguridad de la información. Del mismo modo incentivamos a investigaciones futuras abordar la gestión de riesgos en un entorno de IA, cuyo fin es investigar cómo las organizaciones pueden gestionar de manera efectiva los riesgos asociados con la implementación de soluciones de IA en la seguridad de la información.

Referencias bibliográficas

Amores, C. V., Pérez, J. G. B., Solera-Cotanilla, S., Vega-Barbas, M., Suarez-Tangil, G., Álvarez-Campana, M., & López G. (2023). Analysis of security and data control in smart personal assistants from the users perspective. *Future Generation Computer Systems*, *144*, 12–23. https://doi.org/10.1016/j.future.2023.02.009

Ansari, M. F. (2022, September 1). *The Impact and Limitations of Artificial Intelligence in Cybersecurity: A literature review*. https://ssrn.com/abstract=4323317

Beal, V. (2021). What is SSADM? | Webopedia. Webopedia. https://www.webopedia.com/definitions/ssadm/

Blauth, T. F., Gstrein, O. J., & Zwitter, A. (2022). Artificial Intelligence Crime: An Overview of Malicious Use and Abuse of AI. IEEE Access, 10, 77110–77122. https://doi.org/10.1109/ACCESS.2022.3191790

Cabrera E., Gonzalo J., Peinado V., Penas A. & Verdejo F. (2004). *PRISMA: un modelo interactivo de Síntesis de Información*. Rua.ua.es. Recuperado el 11 de agosto de 2023, de https://rua.ua.es/dspace/bitstream/10045/1449/1/PLN_33_02.pdf

Cappellozza, A., De Moraes, G. H. S. M., Perez, G., & Simões, A. (2021). Antecedent factors of violation of information security rules. *RAUSP Management Journal, ahead-of-print*(ahead-of-print). https://doi.org/10.1108/rausp-02-2021-0022

Carmen Pérez Rodrigo (2012).“ Las revisiones sistemáticas: declaración PRISMA”. Consultado de https://renc.es/imagenes/auxiliar/files/Nutr_1-2012%20Taller%20escritura.pdf

Dash, Bibhu and Ansari, Meraj Farheen and Sharma, Pawankumar and Ali, Azad, Threats and Opportunities with AI-Based Cyber Security Intrusion Detection: A Review (September

Vol. 13 N° 2, pp. 1046 – 1063, Julio/diciembre 2023
Recibido 26/09/2023
Aceptado 19/12//2023
Publicado 31/12/2023

2022). International Journal of Software Engineering & Applications (IJSEA), Vol.13, No.5, September 2022, https://ssrn.com/abstract=4323258

Fritsch, L., Jaber, A., Yazidi, A. (2022). An Overview of Artificial Intelligence Used in Malware. In: Zouganeli, E., Yazidi, A., Mello, G., Lind, P. (eds) Nordic Artificial Intelligence Research and Development. NAIS 2022. Communications in Computer and Information Science, vol 1650. Springer, Cham. https://doi.org/10.1007/978-3-031-17030-0_4

J. Seymour and P. Tully, 'Weaponizing data science for social engineering: Automated E2E spear phishing on Twitter'. 2016. [Online]. Available: https://www.blackhat.com/docs/us16/materials/us-16-Seymour-TullyWeaponizing-Data-Science-For-SocialEngineering-Automated-E2E-SpearPhishing-On-Twitter-wp.pdf

Kaur, R., Gabrijelčič, D., & Klobučar, T. (2023). Artificial intelligence for cybersecurity: Literature review and future research directions. Information Fusion, 97, 101804. https://doi.org/10.1016/j.inffus.2023.101804

Liu, Q., & Zhang, T. (2023). Deep learning technology of computer network security detection based on artificial intelligence. Computers & Electrical Engineering, 110, 108813. https://doi.org/10.1016/j.compeleceng.2023.108813

Mouton F., Leenen L., H.S. (2016) Venter Social engineering attack examples, templates and scenarios https://doi.org/10.1016/j.cose.2016.03.004

Raimundo, R., & Rosário, A. T. (2021). The Impact of Artificial intelligence on data system Security: A literature review. Sensors, 21(21), 7029. https://doi.org/10.3390/s21217029

Shitharth, S., Srivastava, G., Khadidos, A. O., Khadidos, A. O., Baza, M., Alshehri, A., & Lin, J. C. (2023). An artificial intelligence lightweight blockchain security model for security and privacy in IIoT systems. Journal of Cloud Computing, 12(1). https://doi.org/10.1186/s13677-023-00412-y

Solís, C. R. M. (2023). La inteligencia artificial en la seguridad informática. *Revista Empresarial & Laboral*. https://revistaempresarial.com/tecnologia/seguridad-informatica/la-inteligencia-artificial-en-la-seguridad-informatica/

Vol. 13 N° 2, pp. 1046 – 1063, Julio/diciembre 2023
Recibido 26/09/2023
Aceptado 19/12//2023
Publicado 31/12/2023

S. Chen et al., "Automated poisoning attacks and defenses in malware detection systems: An adversarial machine learning approach," computers & security, vol. 73, pp. 326–344, 2018.

Stevens, Tim, Knowledge in the grey zone: AI and cybersecurity (2020). Stevens, T. (2020) Knowledge in the grey zone: AI and cybersecurity. Digital War 1(1): 164-170., SSRN: https://ssrn.com/abstract=4031502

Urrútia G. & Bonfill X.(2010). Declaración PRISMA: una propuesta para mejorar la publicación de revisiones sistemáticas y metaanálisis. Medicina clínica, 135(11),507–511. https://doi.org/10.1016/j.medcli.2010.01.015

Wolsey A. (2022) The State-of-the-Art in AI-Based Malware Detection Techniques: A Review arXiv https://doi.org/10.48550/arXiv.2210.11239