

Solución autonómica para la gestión integral de la ciberseguridad basado en aprendizaje por refuerzo

Autonomic solution for cybersecurity integral management based on reinforcement learning

Ing. Ariel Baloirá Reyes^{*}, Dr.C. Walter Baluja García²,
Dr.C. Caridad E. Anías Calderón³

Recibido: 11/2025 | Aceptado: 11/2025 | Publicado: 12/2025

Resumen

La ciberseguridad se ha convertido en una carrera contrareloj. Los ciberdelincuentes mejoran sus herramientas cada día y lanzan múltiples ataques en cortos intervalos de tiempo. Este trabajo presenta los componentes funcionales de un sistema autónomo para la gestión integral de la ciberseguridad. Sus objetivos principales son tres: (1) disminuir los tiempos de detección, análisis y respuesta ante amenazas y vulnerabilidades; (2) automatizar los procesos de gestión de ciberseguridad en entidades cubanas; y (3) mantener la infraestructura segura y estable con la mínima intervención humana posible. Para la creación del sistema se utilizaron soluciones de *software* libre

^{*} Universidad Tecnológica de La Habana “José Antonio Echeverría”, CUJAE. Calle 114 e/ Ciclovía y Rotonda, Marianao, La Habana, Cuba. ariel@tele.cujae.edu.cu

² Ministerio de Educación Superior. Calle 23 e/ F y G, Vedado, La Habana, Cuba. walter@mes.gob.cu

³ Universidad Tecnológica de La Habana “José Antonio Echeverría”, CUJAE. Calle 114 e/ Ciclovía y Rotonda, La Habana, Cuba. cacha@tesla.cujae.edu.cu

y código abierto. Para dotar de autonomía al sistema se desplegaron dos modelos de inteligencia artificial empleando Python y PyTorch. El desarrollo de este sistema constituye un paso de avance para Cuba hacia la soberanía tecnológica en el área de la ciberseguridad.

Palabras clave: gestión de ciberseguridad, gestión autónoma, aprendizaje por refuerzo, LLM, ciberseguridad

Abstract

Cybersecurity has become a race against time. Cybercriminals improve their tools every day and perpetuate a large number of attacks in a very short period of time. This work presents the functional components of an autonomous system for integrated cybersecurity management. It seeks to (I) reduce detection, analysis, and response times to threats and vulnerabilities; (II) automate cybersecurity management processes in Cuban entities; and (III) maintain secure and stable the infrastructure with minimal human intervention. Free software and open-source solutions were used to create the system. Two artificial intelligence models using Python and PyTorch were deployed to make the system autonomous. The development of this system represents a step forward for Cuba toward technological sovereignty in the area of cybersecurity.

Keywords: cybersecurity management; autonomous management; reinforcement learning; LLM; cybersecurity

Introducción

La ciberseguridad actual se ha convertido en una carrera contrarreloj entre los atacantes y los operadores de las infraestructuras y sistemas de TI (Tecnologías de la Información). Las amenazas han aumentado de manera exponencial con el pasar de los años y su control representa un desgaste continuo para los especialistas de ciberseguridad (2024 SonicWall Cyber Threat Report, 2024; CrowdStrike, 2024; European Union Agency for Cybersecurity., 2024; Fabio Assolini, 2024). Los ciberatacantes han escalado sus operaciones por medio de la automatización y el empleo de la inteligencia artificial (IA). Esto obliga a los especialistas de ciberseguridad a desarrollar soluciones automatizadas que permitan hacerle frente.

La investigación realizada por el Centro de Estudios Estratégicos e Internacionales (CSIS) de los Estados Unidos muestra que las grandes organizaciones emplean aproximadamente cuarenta y siete herramientas de ciberseguridad diferentes en sus redes de un promedio de diez proveedores (Crumpler & Lewis, 2020). Gestionar este cúmulo de herramientas conlleva cuantiosos gastos monetarios y humanos, sobre todo para brindar respuesta a las amenazas de forma proactiva y veloz. Como solución a esta problemática, se ha planteado la integración de diferentes tecnologías como los sistemas de Gestión de Información y Eventos de Seguridad (SIEM, siglas del término en inglés, *Security Information and Event Management*), de Detección y Respuesta Extendidas (XDR, *eXtended Detection and Response*), IA y la computación en la nube (European Union Agency for Cybersecurity., 2023; Uzoma et al., 2023).

La detección oportuna y la identificación precisa de cambios en los patrones de comportamiento en la infraestructura son cruciales para garantizar el funcionamiento efectivo de los mecanismos de protección de ciberseguridad. La naturaleza no estacionaria del ciberespacio plantea un desafío significativo para el desarrollo de soluciones de protección robustas y actualizadas. Por lo tanto, los métodos de IA a emplear necesitan ajustarse de forma dinámica a las condiciones cambiantes del entorno, con el fin de que los sistemas de defensa respondan con eficacia a las amenazas. Como solución se propone el empleo de algoritmos de Aprendizaje por Refuerzo, *Reinforcement Learning*), capaces de generar modelos que imitan el razonamiento humano. Estos son adaptativos, aprenden del entorno en que se encuentran y seleccionan las acciones que maximizan una recompensa acumulada, basándose en la experiencia previa. (Adawadkar & Kulkarni, 2022; Alonso et al., 2021; Hore et al., 2022; Huang et al., 2022; Mathew, 2021; Phan & Bauschert, 2022; Sewak et al., 2022).

Este trabajo presenta un sistema basado en herramientas de *software* libre y código abierto para gestionar información de ciberseguridad, empleando técnicas de RL para responder de forma autónoma a las amenazas existentes en la infraestructura gestionada. Permite a los

especialistas de ciberseguridad supervisar la toma de decisiones del sistema y obtener una respuesta rápida y efectiva a los posibles ciberataques. El sistema diseñado busca alcanzar los siguientes objetivos: (1) disminuir tiempos de detección, análisis y respuesta a amenazas y vulnerabilidades; (2) automatizar procesos de gestión de ciberseguridad en entidades de Cuba; y (3) mantener la infraestructura segura y estable con mínima intervención humana.

Materiales y métodos

En la presente investigación, se emplearon los métodos de investigación científica de análisis-síntesis para el estudio de los componentes de soluciones SIEM y XDR necesarios para implementar la gestión automatizada de la ciberseguridad en entornos heterogéneos. Se recurre al método del modelado y simulación computacional para representar los ataques cibernéticos que el sistema tiene que detectar y mitigar de forma automatizada. Por último, se emplean los métodos empíricos de observación, medición y experimentación durante la validación de la solución propuesta.

Para conformar el sistema se utilizaron las soluciones de *software* libre y código abierto Wazuh, OpenVAS, Suricata y Snort. Para el desarrollo de la solución de IA se usó el lenguaje de programación Python junto con la librería PyTorch como *framework* principal para la construcción de los modelos. La solución propuesta se encuentra en la fase de desarrollo y optimización, no obstante, existe una primera versión de la plataforma operativa.

Resultados y discusión

El sistema para la gestión autónoma de la ciberseguridad está integrado por seis componentes: (1) recolección de datos, (2) monitorización y análisis, (3) detección y clasificación, (4) autonomía decisional, (5) supervisión especializada y (6) almacenamiento. Para garantizar la estabilidad y disponibilidad, su diseño se sustenta en la arquitectura de microservicios. Esto permite la progresión de los componentes y posibilita la recolección de los datos por grupos de colectores de manera simultánea. La interrelación entre componentes se muestra en la Figura 1.

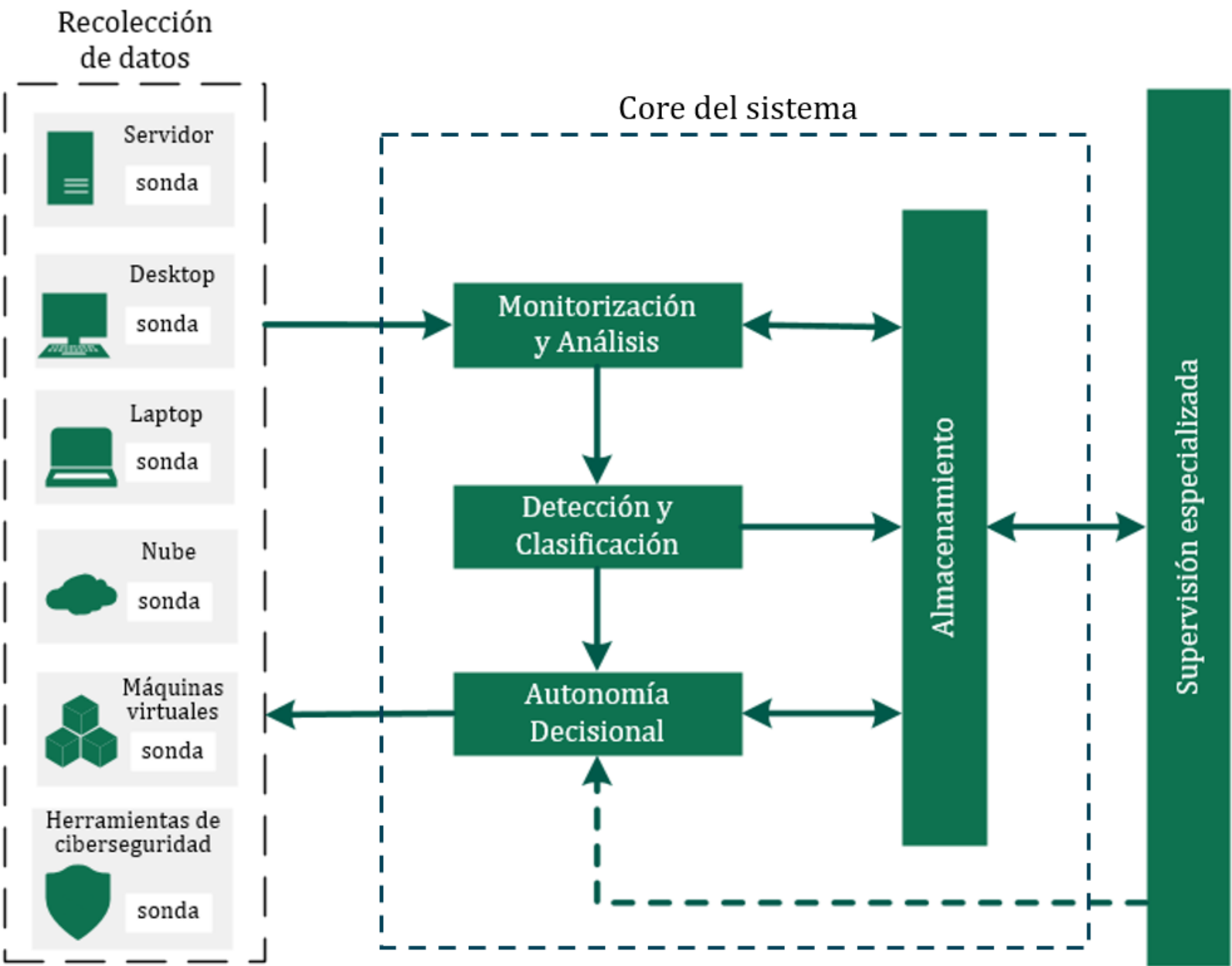


Figura 1. Componentes internos del sistema autónomo de gestión integral de ciberseguridad

Recolección de datos

El componente principal es el de recolección de datos, encargado de la obtención de la información de interés. Está compuesto por sondas distribuidas en la infraestructura de red y agentes de recolección que se encuentran desplegados en los puntos finales, denominados colectores. Su característica fundamental es la capacidad de recopilar datos heterogéneos que incluyen: registros del sistema operativo, tráfico de red, registros de eventos de instancias de virtualización y computación en la niebla/nube, alertas de sistemas de seguridad como los Sistemas de Detección de Intrusos (IDS, *Intrusion Detection System*) o cortafuegos.

Los datos recopilados se pueden clasificar en dos grupos: generales y específicos. Los datos generales están constituidos por registros de tráfico, registros del sistema operativo, registros de aplicaciones y registros de seguridad. Los datos específicos consisten en confirmaciones de existencia de vulnerabilidades, datos de amenazas y datos de inteligencia colaborativa. Es importante recalcar que el tipo específico de datos variará en función de las herramientas de ciberseguridad

que se utilicen y de las necesidades de la organización; sin embargo, los datos generales y específicos descritos anteriormente se consideran esenciales para prevenir y detectar incidentes de ciberseguridad.

En concreto, el objetivo del componente de Recolección de Datos es obtener datos de ciberseguridad en bruto de diversas fuentes, en tiempo real o con un retraso mínimo. Estos serán enviados al componente Monitorización y Análisis para su procesamiento y conservación.

Monitorización y Análisis

Para procesar y almacenar los datos recogidos por los colectores es necesario poseer un punto de intercambio común para el sistema. En este sentido, el componente de Monitorización y Análisis provee interfaces de comunicación estándares con el componente de Recolección de Datos. Estas son *Syslog* (*System Logging Protocol*) que permite el intercambio de mensajes de registro con un formato estándar; y API REST (API, Interfaz de Programación de Aplicaciones basada en la arquitectura de Transferencia de Estado Representacional, REST).

Varios colectores pueden recopilar la misma información y representarla de forma distinta. Por este motivo, una de sus funciones fundamentales es la normalización de la información recibida. Esta función se realiza sustentada en expresiones regulares (Zheng et al., 2021), y en el formato JSON (del término en inglés, *JavaScript Object Notation*); formato de intercambio de datos ligero fácil de leer y escribir para los humanos y las máquinas (Banhara et al., 2023). Con su implementación se puede generar una estructura de datos ordenados, de fácil entendimiento, además de generalizar la información relevante para los procesos de análisis; basado en este primer paso de procesamiento de los datos en bruto, se pueden enriquecer los registros con información asociada. Por ejemplo, la dirección IP de un paquete puede dar a conocer el país y el proveedor de servicios correspondiente.

Otra función importante del componente de Monitorización y Análisis es la correlación de los datos obtenidos de los colectores, para agrupar la información y evitar su duplicación. Dado que las infraestructuras de red pueden variar en tamaños, es necesario que este componente sea capaz de recibir y normalizar un volumen elevado de información de forma simultánea.

Detección y Clasificación

Es el responsable de detectar amenazas, anomalías e infracciones de cumplimiento normativo y de generar alertas al identificar actividad sospechosa. Dicho componente se apoya en múltiples fuentes de inteligencia de amenazas y enriquece las alertas con datos contextuales para mejorar la precisión de la detección. Esto incluye la asignación de eventos al marco MITRE ATT&CK, la detección de vulnerabilidades con el servicio Wazuh CTI y la alineación de los hallazgos con estándares regulatorios como PCI DSS, RGPD, HIPAA, los puntos de referencia CIS y NIST 800-53 (Figura 2). Estas capacidades proporcionan información práctica para la búsqueda de amenazas, la detección de vulnerabilidades y la monitorización del cumplimiento normativo.

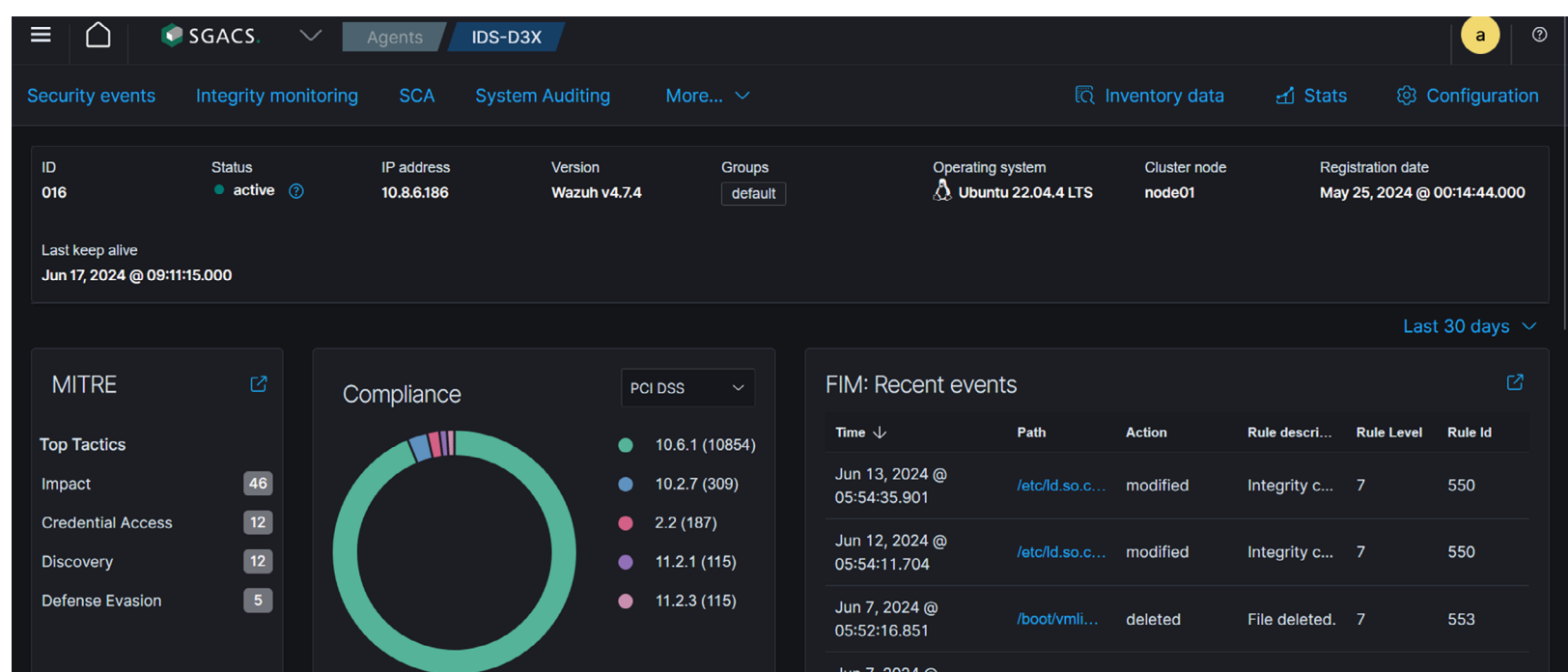


Figura 2. Ilustración del funcionamiento del componente de Detección y Clasificación

Mediante los complementos de la plataforma Wazuh, el componente de Detección y Clasificación se integra con plataformas externas para optimizar los flujos de trabajo. Algunos ejemplos incluyen sistemas de *tickets* como Jira o TheHive, así como herramientas de comunicación como Slack o Telegram.

Un valor muy importante lo tienen los modelos de Aprendizaje Automático (ML, siglas del término en inglés, *MachineLearning*). Estos se implementan como piezas clave en la detección y clasificación de amenazas. Para la implementación de este componente, se seleccionó el algoritmo kNN (*k-nearest neighbors*) que es un clasificador de aprendizaje supervisado que emplea la proximidad para realizar clasificaciones o predicciones sobre la agrupación de un punto de datos individual.

Autonomía decisional

El cerebro del sistema está representado por el componente Autonomía Decisional. Este se forma por dos tecnologías de IA: (1) Modelos Grandes de Lenguaje (LLM, *Large Language Models*) y (2) Aprendizaje por Refuerzo Profundo (DRL, *Deep Reinforcement Learning*). El LLM recibe información del componente Detección y Clasificación, la procesa y organiza en una matriz de observación que funciona como entrada del modelo DRL. Este último decide qué acción tomar para disminuir o mitigar los riesgos. La integración de estos modelos contribuye directamente a la autonomía del sistema.

La matriz de observación representa el estado de la infraestructura gestionada. En consecuencia, el LLM se entrena y prepara para conformar una imagen de la situación actual de la infraestructura. Para conseguirlo, se apoya en eventos pasados, datos del marco MITRE ATT&CK, datos sobre las vulnerabilidades presentes y estado del cumplimiento de las normativas internacionales y nacionales. Dado que el sistema está destinado principalmente para las instituciones cubanas, el LLM ha sido refinado con las regulaciones nacionales referidas a la ciberseguridad. Hasta el momento se ha experimentado con el modelo Llama 3.2:3b. En la actualidad, se ejecuta un estudio de contraste con otros modelos dentro del estado de la cuestión con el objetivo de seleccionar el que mejor se ajuste a los recursos disponibles.

Por otro lado, los estudios realizados por Adawadkar y Kulkarni (2022); Asmat et al. (2025); Hore et al. (2022), Huang et al. (2022) y; Phan y Bauschert (2022) señalan que un modelo de DRL posee la habilidad de tomar decisiones estratégicas y se basa tanto en el estado actual como en las condiciones anteriores del entorno. Para alcanzar este nivel de autonomía, se opta por utilizar una red neuronal de tipo DDQN (*Double Deep Q Network*), cuya principal ventaja radica en su capacidad de mejora continua y que opera como un programa que evoluciona de manera similar al aprendizaje humano. En cada iteración, el algoritmo analiza las acciones realizadas en función del entorno observado y ajusta su comportamiento con el objetivo de

maximizar la recompensa obtenida. En la Figura 3 se representa el diagrama de funcionamiento de este componente.

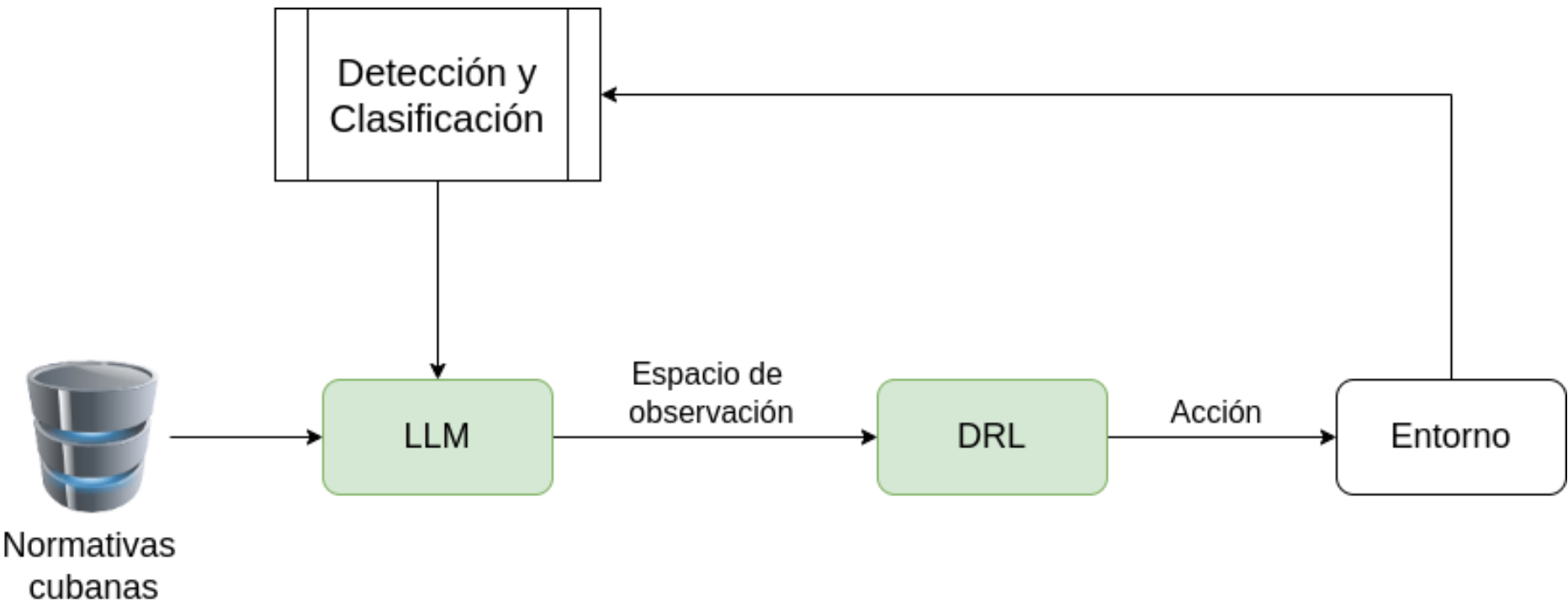


Figura 3. Diagrama de funcionamiento del componente de Autonomía Decisional

Almacenamiento

El componente de Almacenamiento gestiona toda la información generada por los componentes anteriores, utilizando formato JSON para facilitar su indexación. Esta función se implementa mediante OpenSearch (<https://opensearch.org/>), plataforma de búsqueda y observación de código abierto. La función principal del componente de Almacenamiento es la conservación de los datos, el cumplimiento normativo y el análisis forense. Al ser un sistema distribuido y desplegado sobre la tecnología Docker, el componente ofrece escalabilidad y persistencia. Gracias a la conservación de los datos durante largos períodos de tiempo permite, tanto a especialistas como investigadores, acceder a datos históricos y realizar un análisis forense. Además, posibilita analizar tendencias, identificar patrones recurrentes y evaluar la efectividad de las medidas implementadas. Al mismo tiempo, se garantiza uno de los requisitos de ciberseguridad primordial: la conservación de los datos de eventos para su análisis en situaciones que lo ameriten (*ISO/IEC 27001:2022*, 2022).

Supervisión especializada

Al igual que un docente supervisa el proceso de aprendizaje de sus estudiantes, es necesario implementar mecanismos de control que permitan evaluar las decisiones del sistema en tiempo real y

retrospectivamente. A pesar de los beneficios de la introducción de la IA en la ciberseguridad, la dificultad para entender cómo actúan estos modelos mientras operan de forma autónoma es un riesgo que no puede ser ignorado por los especialistas ni las organizaciones (Arreche et al., 2024; Atakishiyev et al., 2024; Biswas et al., 2024; Nwakanma et al., 2023; Rjoub et al., 2023; Tsakalakis et al., 2024). El componente de Supervisión Especializada es el que permite visualizar las acciones realizadas por el sistema, así como acceder a los datos normalizados y procesados representándolos de forma clara mediante diagramas, gráficos y tablas (Figura 4).

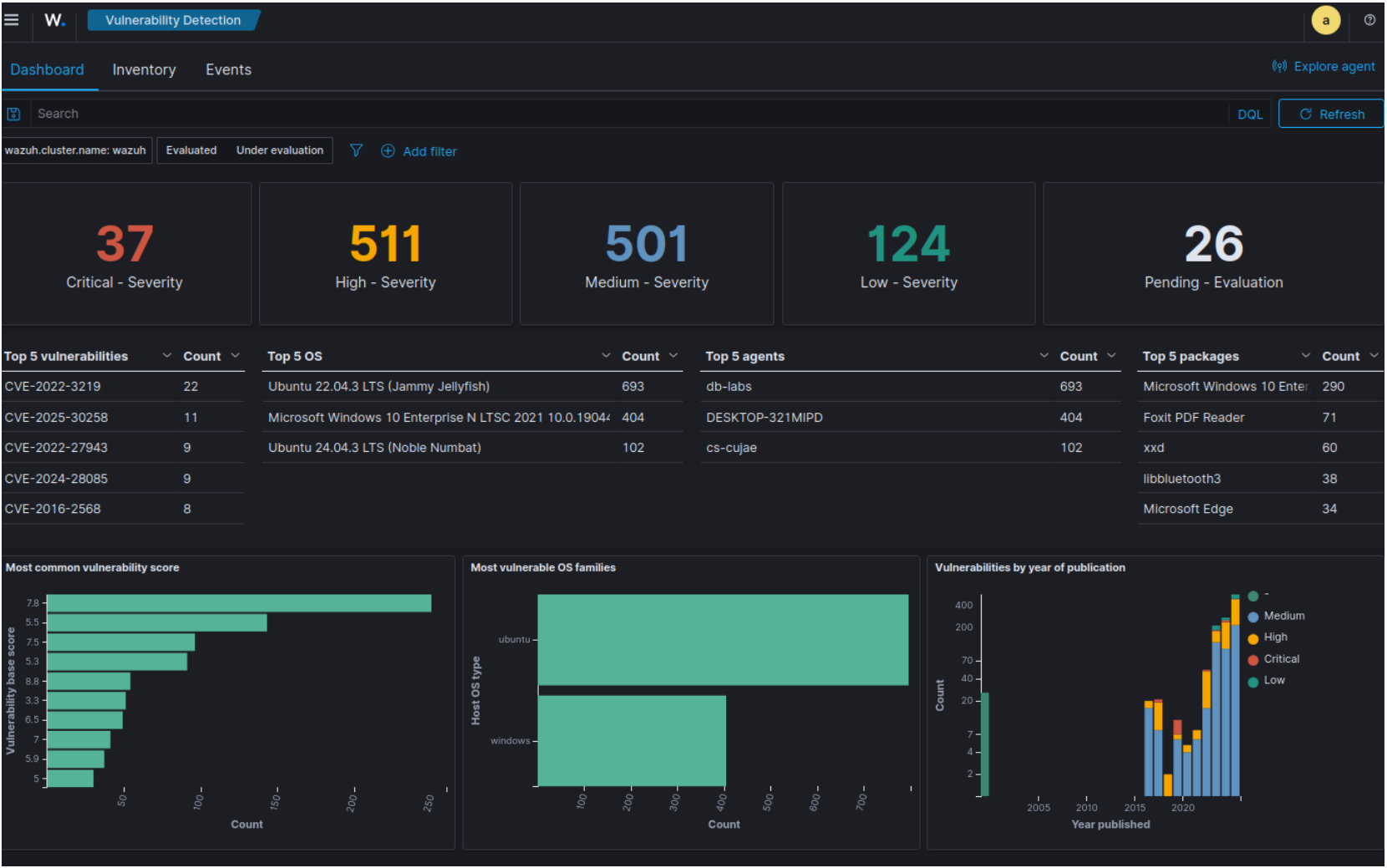


Figura 4. Visualización de las vulnerabilidades presentes en un dispositivo

Es el responsable de crear una representación gráfica intuitiva de la recopilación, el almacenamiento y el procesamiento de los datos de ciberseguridad. Esta representación puede ser aprovechada por las partes interesadas (como analistas de seguridad, gerentes de TI y ejecutivos), para comprender con mayor claridad el panorama de ciberseguridad de la organización y tomar decisiones informadas sobre la gestión de riesgos (Figura 5). Este componente facilita la interpretación de los elementos que desencadenan una acción específica en los modelos de IA, sobre todo el de RL. Su función es primordial para evitar el sobreajuste y corregir desviaciones del sistema de forma anticipada.

>	Jun 28, 2023 @ 21:40:25.861	PAM: User login failed.	5	5503
>	Jun 28, 2023 @ 21:40:25.847	Host Blocked by firewall-drop Active Response	3	651
>	Jun 28, 2023 @ 21:40:25.610	sshd: Attempt to login using a non-existent user	5	5710
>	Jun 28, 2023 @ 21:40:25.607	sshd: Attempt to login using a non-existent user	5	5710
>	Jun 28, 2023 @ 21:40:25.607	sshd: brute force trying to get access to the system. Non existent user.	10	5712
>	Jun 28, 2023 @ 21:40:24.156	sshd: Attempt to login using a non-existent user	5	5710
>	Jun 28, 2023 @ 21:40:24.156	sshd: Attempt to login using a non-existent user	5	5710
>	Jun 28, 2023 @ 21:40:21.847	PAM: User login failed.	5	5503
>	Jun 28, 2023 @ 21:40:21.843	PAM: User login failed.	5	5503
>	Jun 28, 2023 @ 21:40:21.618	sshd: Attempt to login using a non-existent user	5	5710
>	Jun 28, 2023 @ 21:40:21.618	sshd: Attempt to login using a non-existent user	5	5710
>	Jun 28, 2023 @ 21:40:21.618	sshd: Attempt to login using a non-existent user	5	5710

Figura 5. Registros de acciones y eventos del sistema

Conclusiones

El desarrollo de un sistema autónomo para la gestión integral de la ciberseguridad es un paso más hacia la soberanía tecnológica nacional. La integración de técnicas de IA permite reducir los tiempos de respuesta a incidentes y la corrección de vulnerabilidades antes de que sean explotadas. Los seis componentes que conforman el sistema cumplen funciones independientes, sin embargo al integrarse aportan una capacidad significativa a la solución desarrollada. El componente de Autonomía Decisional representa una de las principales contribuciones para lograr la ciberseguridad autónoma, pero implica un riesgo considerable si no se gestiona de forma adecuada. Esta función de control se delega en el componente de Supervisión Especializada, el cual permite la visualización de todas las acciones que realiza el sistema y su supervisión controlada. Es necesario mantener el constante perfeccionamiento de los modelos de lenguaje empleados, con el fin de optimizar el uso de recursos. Como línea de trabajo futura resulta esencial profundizar en el estudio y diseño de modelos de explicabilidad orientados a interpretar las decisiones del sistema y sus implicaciones.

Referencias bibliográficas

2024 SonicWall Cyber Threat Report. (2024). <https://www.sonicwall.com/resources/white-papers/2024-sonicwall-cyber-threat-report>

- Adawadkar, A. M. K., & Kulkarni, N. (2022). Cyber-security and reinforcement learning—A brief survey. *Engineering Applications of Artificial Intelligence*, 114, 105116. <https://doi.org/10.1016/j.engappai.2022.105116>
- Alonso, R. S., Prieto, J., La Prieta, F. D., Rodriguez-Gonzalez, S., & Corchado, J. M. (2021). A Review on Deep Reinforcement Learning for the management of SDN and NFV in Edge-IoT. 2021 IEEE Globecom Workshops (GC Wkshps), 1-6. <https://doi.org/10.1109/GCWkshps52748.2021.9682179>
- Arreche, O., Guntur, T., & Abdallah, M. (2024). XAI-IDS: Toward Proposing an Explainable Artificial Intelligence Framework for Enhancing Network Intrusion Detection Systems. *Applied Sciences*, 14(10), 4170. <https://doi.org/10.3390/app14104170>
- Asmat, H., Din, I. U., Almogren, A., & Khan, M. Y. (2025). Digital Twin with Soft Actor-Critic Reinforcement Learning for Transitioning from Industry 4.0 to 5.0. *IEEE Access*, 1-1. IEEE Access. <https://doi.org/10.1109/ACCESS.2025.3546085>
- Atakishiyev, S., Salameh, M., Yao, H., & Goebel, R. (2024). Explainable Artificial Intelligence for Autonomous Driving: A Comprehensive Overview and Field Guide for Future Research Directions. *IEEE Access*, 12, 101603-101625. <https://doi.org/10.1109/ACCESS.2024.3431437>
- Banhara, N., Duarte, D., & Schreiner, G. (2023). Extração de Esquemas de Documentos JSON: O que há de Novo? *Escola Regional de Banco de Dados (ERBD)*, 11-20. <https://doi.org/10.5753/erbd.2023.229421>
- Biswas, B., Mukhopadhyay, A., Kumar, A., & Delen, D. (2024). A hybrid framework using explainable AI (XAI) in cyber-risk management for defence and recovery against phishing attacks. *Decision Support Systems*, 177, 114102. <https://doi.org/10.1016/j.dss.2023.114102>
- CrowdStrike. (2024). *CrowdStrike 2024 Global Threat Report*. <https://www.crowdstrike.com/en-us/global-threat-report/>
- Crumpler, W. D., & Lewis, J. A. (2020). *Cybersecurity and the Problem of Interoperability*.
- European Union Agency for Cybersecurity. (2023). Artificial intelligence and cybersecurity research: ENISA research and innovation Brief. *Publications Office*. <https://data.europa.eu/doi/10.2824/808362>
- European Union Agency for Cybersecurity. (2024). ENISA threat landscape 2024: July 2023 to June 2024. *Publications Office*. <https://data.europa.eu/doi/10.2824/0710888>

- Fabio Assolini. (2024, marzo 19). *Panorama de Amenazas Cibernéticas en Cuba y región* [Conferencia magistral]. <https://www.informaticahabana.cu/actividad/panorama-de-amenazas-ciberneticas-en-cuba-y-region/>
- Hore, S., Shah, A., & Bastian, N. D. (2022). *Deep VULMAN: A Deep Reinforcement Learning-Enabled Cyber Vulnerability Management Framework* (No. arXiv:2208.02369). arXiv. <https://doi.org/10.48550/arXiv.2208.02369>
- Huang, Y., Huang, L., & Zhu, Q. (2022). Reinforcement Learning for feedback-enabled cyber resilience. *Annual Reviews in Control*, 53, 273-295. <https://doi.org/10.1016/j.arcontrol.2022.01.001>
- ISO/IEC 27001:2022—Information security management systems. (2022). [International Standard published]. ISO/IEC JTC 1/SC 27. https://www-iso-org.translate.google.com/standard/27001?x_tr_sl=en&x_tr_tl=es&x_tr_hl=es&x_tr_pto=tc
- Mathew, A. (2021). Deep Reinforcement Learning for Cybersecurity Applications. *International Journal of Computer Science and Mobile Computing*, 10(12), 32-38. <https://doi.org/10.47760/ijcsmc.2021.v10i12.005>
- Nwakanma, C. I., Ahakonye, L. A. C., Njoku, J. N., Odirichukwu, J. C., Okolie, S. A., Uzundu, C., Ndubuisi Nweke, C. C., & Kim, D.-S. (2023). Explainable Artificial Intelligence (XAI) for Intrusion Detection and Mitigation in Intelligent Connected Vehicles: A Review. *Applied Sciences*, 13(3), 1252. <https://doi.org/10.3390/app13031252>
- Phan, T. V., & Bauschert, T. (2022). DeepAir: Deep Reinforcement Learning for Adaptive Intrusion Response in Software-Defined Networks. *IEEE Transactions on Network and Service Management*, 19(3), 2207-2218. <https://doi.org/10.1109/TNSM.2022.3158468>
- Rjoub, G., Bentahar, J., Wahab, O. A., Mizouni, R., Song, A., Cohen, R., Otrok, H., & Mourad, A. (2023). A Survey on Explainable Artificial Intelligence for Cybersecurity. *IEEE Transactions on Network and Service Management*, 20(4), 5115-5140. <https://doi.org/10.1109/TNSM.2023.3282740>
- Sewak, M., Sahay, S. K., & Rathore, H. (2022). Deep Reinforcement Learning for Cybersecurity Threat Detection and Protection: A Review. En R. Krishnan, H. R. Rao, S. K. Sahay, S. Samtani, & Z. Zhao (Eds.), *Secure Knowledge Management In The Artificial Intelligence Era* (Vol. 1549, pp. 51-72). Springer International Publishing. https://doi.org/10.1007/978-3-030-97532-6_4
- Tsakalakis, N., Stalla-Bourdillon, S., Huynh, T. D., & Moreau, L. (2024). *A taxonomy of explanations to support Explainability-by-Design* (No. arXiv:2206.04438). arXiv. <https://doi.org/10.48550/arXiv.2206.04438>

Uzoma, J., Falana, O., Obunadike, C., Oloyede, K., & Obunadike, E. (2023). *Using artificial intelligence for automated incidence response in cybersecurity*.

Zheng, L.-X., Ma, S., Chen, Z.-X., & Luo, X.-Y. (2021). Ensuring the Correctness of Regular Expressions: A Review. *International Journal of Automation and Computing*, 18(4), 521-535. <https://doi.org/10.1007/s11633-021-1301-4>

