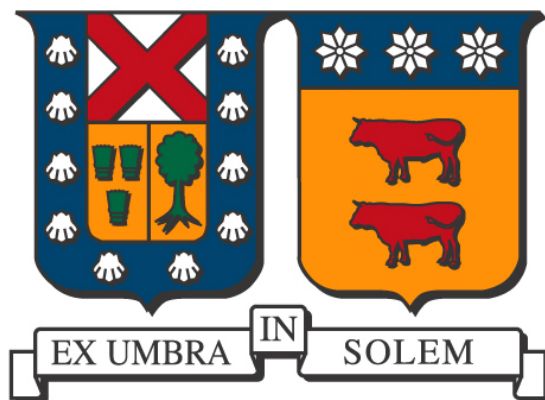


UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE ELECTRÓNICA
SANTIAGO - CHILE



**“ Plataforma de Gestión de Superficies de Ataque Cibernético para
Operadores de Telecomunicaciones”**

SEBASTIÁN ANDRÉS JARA PALOMINO

**MEMORIA DE TITULACIÓN PARA OPTAR AL TÍTULO DE INGENIERO CIVIL
TELEMÁTICO**

PROFESOR GUÍA:

NICOLÁS TORRES RUDLOFF

PROFESOR CORREFERENTE:

BERIOSKA CONTRERAS VARGAS

Marzo - 2025

Agradecimientos

Este camino ha estado lleno de desafíos, aprendizajes y momentos inolvidables, y no lo habría logrado sin el apoyo incondicional de quienes han estado a mi lado.

En primer lugar, agradezco profundamente a mis padres, Claudia Palomino y Esteban Jara, por su amor, paciencia y confianza inquebrantable, incluso cuando el camino se alargó más de lo esperado. Gracias por ser mi pilar y darme la fuerza para seguir adelante.

A mi hermano, Brian Jara, por estar siempre ahí, apoyándome y animándome en cada obstáculo. Su compañía ha sido un refugio en este proceso.

A mis amigos, mis panas, Constanza Corday, Gonzalo Soto, Juan de Dios Larrea y Agustín Ovando, por hacer más llevadera la universidad con su apoyo y compañía. Y a quienes iniciaron este viaje conmigo, mis reales: Fernanda Reed, Gabriela González, Catalina Salinas, Emilio Venegas y Diego Medel, gracias por las risas, los carretes y por hacer este camino más significativo.

A Jimena Opazo, mi compañera de vida, quien ha estado conmigo desde el principio, apoyándome en todas mis decisiones. Su constancia, inteligencia y pasión por lo que hace han sido un modelo a seguir para mí. Admiro su fortaleza y su capacidad de nunca rendirse, y le estaré eternamente agradecido por haber estado a mi lado en este camino.

Finalmente, a mi profesor guía, Nicolás Torres, siempre estuvo dispuesto a ayudarme y guiarme, y su orientación ha sido clave para el desarrollo de este trabajo.

A todos ustedes, gracias. Este logro también es suyo.

Dedico este trabajo a mi familia, quienes siempre me han apoyado incondicionalmente en cada paso de este camino, brindándome su amor, paciencia y confianza. Y a mi pareja, quien ha estado a mi lado en todo momento, compartiendo mis logros, alentándome en los desafíos y recordándome siempre el valor de la perseverancia. Sin ustedes, nada de esto habría sido posible.

Resumen

Las organizaciones enfrentan el reto de proteger sus activos ante amenazas cibernéticas en constante evolución. Este proyecto desarrolla una aplicación web escalable para gestionar y analizar la superficie de ataque de un operador de telecomunicaciones, centralizando información obtenida mediante una herramienta de reconocimiento personalizada. A través de modelos de aprendizaje de máquinas (*machine learning*), como Isolation Forest, se identifican patrones, anomalías y vulnerabilidades críticas para priorizar la respuesta ante riesgos.

El sistema evalúa el nivel de riesgo de cada activo en función de la presencia de 8 vectores de ataque específicos. Si se detecta al menos uno, se calcula un nivel de riesgo, permitiendo priorizar la mitigación de amenazas. La solución integra tecnologías como Flask, SQLite y Docker, garantizando un manejo eficiente de grandes volúmenes de datos con tiempos de respuesta óptimos. Además, la interfaz ha sido optimizada para mejorar la experiencia del usuario mediante una visualización clara de los riesgos, facilitando la interpretación de datos. Pruebas con usuarios revelaron mejoras en la interacción con etiquetas (*tags*), reducción en la complejidad de carga de datos y una navegación más intuitiva, favoreciendo la toma de decisiones y priorización de activos.

En términos de rendimiento, el sistema almacena hasta 26,000 registros y procesa datos en un tiempo promedio de 8.93 segundos. Puede soportar hasta 50 usuarios simultáneos sin degradación del rendimiento, con posibilidad de escalar a más de 200 con mejoras en la infraestructura. Su diseño está orientado a analistas de ciberseguridad, ingenieros de TI y auditores, quienes requieren herramientas avanzadas de visualización y análisis de riesgos.

El proyecto satisface los objetivos planteados al integrar tecnología avanzada con un enfoque práctico en la gestión de ciberseguridad. Futuras mejoras, como la incorporación de modelos más sofisticados, bases de datos robustas y un frontend dinámico basado en React o Vue.js, optimizarán su rendimiento y experiencia de usuario, posicionándolo como un estándar en la gestión proactiva de riesgos digitales.

Palabras Clave: Gestión de vulnerabilidades, ciberseguridad, aprendizaje automático, detección de anomalías, visualización de datos, marcos de referencia de aplicaciones web, análisis y gestión de riesgos

Abstract

Organizations face the challenge of protecting their assets from constantly evolving cyber threats. This project develops a scalable web application to manage and analyze the attack surface of a telecommunications operator, centralizing information obtained through a customized reconnaissance tool. Through machine learning models, such as Isolation Forest, patterns, anomalies and critical vulnerabilities are identified to prioritize risk response.

The system evaluates the risk level of each asset based on the presence of 8 specific attack vectors. If at least one is detected, a risk level is calculated, allowing prioritization of threat mitigation. The solution integrates technologies such as Flask, SQLite and Docker, ensuring efficient handling of large volumes of data with optimal response times. In addition, the interface has been optimized to improve the user experience through a clear visualization of risks, facilitating data interpretation. User testing revealed improved interaction with tags, reduced data loading complexity, and more intuitive navigation, facilitating decision making and asset prioritization.

In terms of performance, the system stores up to 26,000 records and processes data in an average time of 8.93 seconds. It can support up to 50 simultaneous users without performance degradation, with the possibility of scaling up to more than 200 with infrastructure improvements. It is designed for cybersecurity analysts, IT engineers and auditors who require advanced visualization and risk analysis tools.

The project meets the stated objectives by integrating advanced technology with a practical approach to cybersecurity management. Future enhancements, such as the incorporation of more sophisticated models, robust databases and a dynamic frontend based on React or Vue.js, will optimize its performance and user experience, positioning it as a standard in proactive digital risk management.

Keywords: Vulnerability management, cybersecurity, machine learning, anomaly detection, data visualization, web application frameworks, analytics and risk management

Tabla de Contenido

1	Introducción	1
1.1	Objetivos	3
1.1.1	Objetivo Principal	3
1.1.2	Objetivos Específicos	4
2	Situación Actual	5
2.1	Superficie de Ataque y su Importancia en el Reconocimiento	5
2.2	Etiquetado (<i>Tagging</i>).	6
2.3	Contextualización de tecnologías	7
2.3.1	Marco de Referencia Flask	7
2.3.2	Base de Datos SQLite	7
2.3.3	Bases de Datos Relacionales	8
2.3.4	Algoritmo no Supervisado Isolation Forest	8
2.4	Contenedores Docker y Reconocimiento con Eyewitness	9
2.5	Riesgo Tecnológico	10
2.6	Marcos de Referencia en Ciberseguridad	11
2.6.1	MITRE ATT&CK®	11
2.6.2	Proyecto de Seguridad de Aplicaciones OWASP	11
2.7	Principios de UX/UI y Diseño Universal en Ciberseguridad	11
2.8	Herramienta “DatosRecon”	12
2.9	Tecnologías y Metodologías de Reconocimiento	13
2.9.1	Reconocimiento Personalizado “DatosRecon”	13
3	Metodología Aplicada	15
3.1	Arquitectura de la Solución	15
3.1.1	Módulo de Reconocimiento Personalizado “DatosRecon”.	17
3.1.2	Base de Datos Centralizada (<i>Backend</i>)	17
3.1.3	Módulo Normalización de Fuentes de Datos (<i>DataSources</i>)	18

TABLA DE CONTENIDO

3.1.4	Módulo de Aprendizaje de Máquinas (<i>Machine Learning</i>)	19
3.1.5	Desarrollo Full Stack	19
3.2	Criterios de Severidad y Evaluación de Riesgos	22
3.3	Flujo de Trabajo	23
3.4	Implementación y Despliegue de la Arquitectura	26
3.5	Pruebas de Rendimiento	30
4	Resultados del Uso de la Plataforma	31
4.1	Resultados de la Interacción con la Superficie de Ataque	31
4.1.1	Visualización de la Superficie de Ataque	31
4.1.2	Resultados de la Interacción con Carga de Fuentes de Datos.	33
4.1.3	Resultados de la Interacción con Etiquetas	34
4.1.4	Resultados de la Interacción con Categoría de Riesgo	36
4.1.5	Detalles Específicos de Activos	38
4.2	Resultados de Rendimiento del Prototipo	43
5	Análisis de Resultados	44
5.1	Percepción de los usuarios	44
5.2	Logros Destacados	45
5.3	Limitaciones del Sistema	46
6	Conclusiones y Trabajo Futuro	47
	Referencias	48

Índice de figuras

3.1	Arquitectura del sistema.	16
3.2	Modelo de base de datos centralizada del sistema.	18
3.3	Diagrama del flujo de trabajo del sistema.	26
4.1	Tabla principal de visualización de activos en Inventarium.	32
4.2	Página de carga y validación de datos en Inventarium.	33
4.3	Zoom en la sección “Tags Seleccionado” con el tag “Cloud” activado.	34
4.4	Tabla principal filtrada por el tag “Cloud”.	35
4.5	Gestión de etiquetas en la sección de detalles de un activo	35
4.6	Zoom a activos con niveles de riesgo bajo, medio y crítico.	37
4.7	Vista general de los detalles de un activo en Inventarium.	39
4.8	Información específica del activo.	40
4.9	Gestión de comentarios para el activo.	41
4.10	Sección de captura de imágenes de la dirección IP vulnerada.	42
4.11	Lista cronológica de incidentes y vectores de ataque asociados al activo.	42

Índice de tablas

1	Comparativa de rendimiento del sistema para diferentes volúmenes de datos.	43
---	--	----

Glosario de Términos

- **Machine Learning:** Rama de la inteligencia artificial que permite a los sistemas identificar patrones y calcular riesgos a partir de datos.
- **Backend:** Parte del sistema que gestiona la lógica del negocio y la comunicación con la base de datos. Implementado con Flask.

- **Frontend:** Componente del sistema que interactúa con el usuario, permitiendo la visualización de datos y gestión de etiquetas.
- **MITRE ATT&CK:** Marco de referencia que clasifica tácticas y técnicas de ciberataques para mejorar la seguridad.
- **Nmap:** Herramienta para escaneo de redes y detección de vulnerabilidades.
- **API RESTful:** Interfaz que permite la comunicación entre el frontend y backend mediante solicitudes HTTP (GET, POST, etc.).
- **Jinja2:** Motor de plantillas utilizado en Flask para generar HTML dinámico basado en datos preprocesados del backend.
- **JSON:** Formato ligero para intercambio de datos estructurados entre sistemas.
- **XLSX:** Formato de archivo utilizado para hojas de cálculo en Microsoft Excel.
- **Vector de Ataque:** Medio por el cual un atacante explota una vulnerabilidad para acceder a un activo.
- **CVE (Common Vulnerabilities and Exposures):** Sistema de identificación para vulnerabilidades conocidas, utilizado para facilitar su referencia y manejo.
- **Fingerprint:** Huella digital única de un activo que incluye información como sistema operativo, servicios activos y configuraciones relevantes.

1. Introducción

En el contexto actual, donde la digitalización avanza rápidamente y las organizaciones dependen cada vez más de sus infraestructuras digitales, las amenazas de ciberseguridad se han vuelto una preocupación central para la continuidad y estabilidad de las operaciones empresariales. Cada vez es más común que las empresas enfrenten ataques y brechas de seguridad que comprometen la confidencialidad de sus activos, lo que destaca la importancia de una gestión efectiva de vulnerabilidades y riesgos. En este escenario, la empresa del sector de telecomunicaciones, con la que se desarrolla el presente proyecto, emplea una herramienta de reconocimiento personalizado de superficies de ataque para identificar los puntos de acceso susceptibles a compromisos de seguridad. Sin embargo, la falta de una interfaz intuitiva, una visualización clara del riesgo y una integración eficiente de etiquetas para la categorización limitan la toma de decisiones basada en riesgo, afectando la efectividad en la gestión de ciberseguridad.

Uno de los principales desafíos identificados en el uso de herramientas de seguridad es la manera en que los analistas interactúan con los datos. En pruebas preliminares realizadas con usuarios, se observó que la visualización de información de ciberseguridad debe ser clara y accesible, evitando sobrecarga de datos que dificulte la priorización de riesgos. Se identificó que los usuarios tienden a preferir interfaces que estructuren la información en paneles organizados, con opciones para filtrar y explorar activos de manera intuitiva. Adicionalmente, la carga y procesamiento de datos debe ser rápida y eficiente, permitiendo importar grandes volúmenes de información sin tiempos de espera prolongados. La interacción con etiquetas (*tags*) ha sido otro aspecto clave en la experiencia del usuario, ya que facilita la organización y priorización de vulnerabilidades, lo que optimiza la gestión de incidentes y la toma de decisiones estratégicas.

Este proyecto tiene como objetivo desarrollar una aplicación web robusta y escalable que permita a un Operador de Telecomunicaciones visualizar de forma centralizada la superficie de ataque obtenida a partir de su herramienta de reconocimiento, junto con la correlación de estos datos con los incidentes de seguridad registrados por la empresa. Este enfoque permitirá

a los responsables de ciberseguridad evaluar, no solo las vulnerabilidades detectadas en los activos, sino también el historial de incidentes asociados, proporcionando una visión integral del estado de ciberseguridad de la organización. Para mejorar la eficiencia del análisis, se ha optimizado la carga y gestión de datos, permitiendo una identificación más ágil de patrones y anomalías mediante algoritmos de aprendizaje de máquina (*machine learning*).

Para el desarrollo de esta aplicación, se emplearán tecnologías web modernas con un enfoque en la escalabilidad y el rendimiento, adaptadas a las necesidades de un Operador de Telecomunicaciones. La arquitectura se diseñará para soportar un gran volumen de datos y consultas complejas, garantizando tiempos de respuesta rápidos y una experiencia de usuario optimizada. Las herramientas utilizadas incluirán bibliotecas especializadas en la aplicación del usuario y de la lógica del negocio, así como bases de datos (*frontend y backend*) para el desarrollo de interfaces gráficas de usuario eficientes. Adicionalmente, la implementación de etiquetas (*tags*) facilitará la organización de los activos y la priorización de vulnerabilidades.

El sistema ha sido diseñado para ofrecer una experiencia de usuario optimizada, integrando mecanismos de retroalimentación visual y categorización de activos que permiten a los analistas tomar decisiones fundamentadas en riesgo de manera eficiente. Los usuarios podrán explorar activos mediante tableros interactivos, filtrar información según niveles de criticidad y visualizar alertas generadas por el modelo de aprendizaje automático. Además, se han implementado mecanismos para agilizar la carga de datos, asegurando que la importación de archivos voluminosos no afecte el rendimiento del sistema.

El alcance del proyecto incluye la implementación de un sistema que no solo permita visualizar las vulnerabilidades detectadas y correlacionarlas con los incidentes previos, sino también realizar un seguimiento de los cambios en el nivel de riesgo de cada activo a lo largo del tiempo, ofreciendo una perspectiva dinámica y actualizada de la seguridad de la organización. Además, la aplicación podrá escalar a medida que aumente la cantidad de activos, incidentes y datos procesados, adaptándose a las nuevas necesidades de expansión que surjan.

En términos de proyección, esta herramienta busca convertirse en un componente esencial de la operación táctica de ciberseguridad en Operadores de Telecomunicaciones nacionales

e internacionales, proporcionando una plataforma para el monitoreo constante y efectivo de sus activos. La integración de aprendizaje de máquina (*machine learning*) permitirá una mejora continua en la precisión de los modelos de riesgo y una adaptación proactiva ante nuevas amenazas. Con esta aplicación, se espera apoyar a la empresa en la toma de decisiones fundamentadas en el análisis de riesgos y la correlación de incidentes, mejorando así su estrategia de defensa frente a las crecientes amenazas en el ámbito digital.

El documento está estructurado en seis capítulos. En el Capítulo 1, se introduce el contexto del problema, junto con los objetivos del proyecto y su relevancia en el ámbito de la ciberseguridad. El Capítulo 2 presenta la situación actual, donde se fusionan el marco teórico y el estado del arte, proporcionando una base conceptual sobre la superficie de ataque, la gestión de vulnerabilidades y el uso de aprendizaje automático en la priorización de riesgos. En el Capítulo 3, se describe la metodología aplicada, detallando la arquitectura de la solución, los módulos que la componen y las herramientas utilizadas en su implementación. El Capítulo 4 expone el flujo de trabajo del sistema, abordando la interacción entre los módulos y el proceso de análisis de datos. En el Capítulo 5, se presentan los resultados obtenidos tras la implementación del sistema, incluyendo métricas de rendimiento, validaciones y percepciones de los usuarios sobre la usabilidad e impacto de la solución. Finalmente, el Capítulo 6 recoge las conclusiones del estudio y las oportunidades de mejora identificadas, estableciendo líneas de trabajo futuro para optimizar el sistema y su aplicabilidad en entornos de telecomunicaciones.

1.1. Objetivos

1.1.1. Objetivo Principal

Desarrollar una aplicación web robusta y escalable para una empresa de telecomunicaciones nacional, que permita visualizar y analizar la superficie de ataque de sus activos obtenida mediante una herramienta de reconocimiento personalizada, integrando una correlación con los incidentes históricos de seguridad de la empresa y proporcionando un análisis de riesgos

basado en aprendizaje de máquinas (*machine learning*) para apoyar la toma de decisión en la operación táctica de la ciberseguridad.

1.1.2. Objetivos Específicos

- **Diseñar interfaz web:** Crear una interfaz intuitiva y fácil de usar que centralice la visualización de la superficie de ataque de la organización, facilitando el acceso a la información de vulnerabilidades y detalles de activos de manera clara y eficiente.
- **Correlacionar datos:** Integrar los datos de incidentes históricos de la empresa con la superficie de ataque, proporcionando un análisis que identifique patrones y comportamientos de riesgo en la seguridad de la empresa.
- **Utilizar modelos de aprendizaje de máquinas para gestión de riesgos:** Diseñar e implementar modelos de aprendizaje de máquinas (*machine learning*) que permitan calcular el nivel de riesgo de cada activo, detectando patrones y anomalías mediante análisis predictivo.
- **Garantizar escalabilidad y rendimiento:** Diseñar la aplicación para manejar grandes volúmenes de datos y consultas, asegurando tiempos de respuesta rápidos y una experiencia de usuario optimizada.
- **Priorizar activos y vulnerabilidades:** Proporcionar herramientas que permitan a los analistas y operadores de ciberseguridad del Operador priorizar la atención a los activos y vulnerabilidades según su nivel de riesgo, facilitando decisiones fundamentadas.

2. Situación Actual

En el contexto de la ciberseguridad moderna, donde las amenazas evolucionan constantemente, las organizaciones enfrentan el desafío de proteger sus activos digitales de manera proactiva. Esto requiere la integración de tecnologías que permitan la identificación de vulnerabilidades, el análisis predictivo de riesgos y la gestión eficiente de grandes volúmenes de datos.

Para garantizar un almacenamiento y procesamiento eficiente, el sistema está diseñado para manejar cientos de miles de registros, manteniendo tiempos de respuesta óptimos. En particular, el uso de SQLite permite gestionar bases de datos de hasta varios gigabytes de tamaño con tiempos de consulta optimizados, asegurando la integridad de la información incluso en operaciones concurrentes.

En este proyecto, herramientas como Flask, SQLite e Isolation Forest trabajan en conjunto para proporcionar una solución que integre análisis en tiempo real, almacenamiento estructurado de datos históricos y detección de anomalías en activos críticos. Esta sección detalla los conceptos clave y tecnologías utilizadas para desarrollar una plataforma escalable y eficiente que apoye la toma de decisiones en ciberseguridad.

2.1. Superficie de Ataque y su Importancia en el Reconocimiento

La superficie de ataque se refiere al conjunto de puntos de entrada a través de los cuales un atacante podría comprometer la seguridad de los activos de una organización. Estos puntos pueden incluir puertos abiertos, servicios vulnerables, aplicaciones expuestas a internet y cualquier otro recurso accesible que pueda ser explotado. Reducir la superficie de ataque es un principio fundamental en la gestión de riesgos de ciberseguridad, ya que minimiza las oportunidades para los atacantes y fortalece la resiliencia de los sistemas ante amenazas emergentes.

La etapa de reconocimiento, definida en el marco MITRE ATT&CK® **mitre_att&ck**, es crucial dentro del proceso de seguridad ofensiva. Este paso inicial permite mapear la su-

perficie de ataque de una red mediante técnicas como escaneos de puertos, identificación de servicios y análisis de configuraciones. Sin embargo, la literatura destaca desafíos como la falta de una memoria organizacional compartida entre los especialistas, lo que conlleva a la duplicación de esfuerzos y resultados inconsistentes. Además, la heterogeneidad en los niveles de conocimiento técnico dentro de los equipos de ciberseguridad impacta significativamente la capacidad de identificar y mitigar riesgos de manera efectiva.

Existen diferentes enfoques para mapear y analizar la superficie de ataque. Según Everson y Cheng [1], la identificación de la superficie de ataque en redes modernas enfrenta retos debido a la evolución constante de las amenazas y la expansión del perímetro digital, impulsada por tecnologías como la computación en la nube y el Internet de las Cosas (*IoT*). Para abordar estos desafíos, es fundamental apoyarse en marcos de referencia que brinden una estructura clara y metodológica, como MITRE ATT&CK®, que proporciona técnicas estandarizadas para la identificación y priorización de amenazas.

En este proyecto, se utiliza una herramienta de reconocimiento automatizado diseñada para escanear los activos de la empresa y generar un listado detallado de posibles puntos vulnerables. La información obtenida es clave para las evaluaciones de riesgo, permitiendo priorizar la mitigación de amenazas con base en la criticidad de los activos y su nivel de exposición a ataques.

2.2. Etiquetado (*Tagging*).

El etiquetado es un método de categorización que permite asignar descriptores a activos o registros con el propósito de facilitar su organización, búsqueda y priorización dentro de un sistema. Este mecanismo es ampliamente utilizado en el ámbito de la gestión de información, ya que posibilita una estructura flexible y adaptable para la clasificación de datos.

En el contexto de este proyecto, el etiquetado permite clasificar activos y vulnerabilidades de acuerdo con criterios específicos, optimizando la capacidad de los usuarios para identificar con rapidez los elementos más relevantes o críticos en términos de seguridad. Al asociar etiquetas a los registros, se mejora la gestión de riesgos y se facilita la toma de decisiones

basada en la categorización de amenazas y activos.

La aplicación desarrollada incorpora un sistema de etiquetado dinámico, el cual permite la asignación, modificación y eliminación de etiquetas de manera directa desde la interfaz de usuario. Además, las etiquetas pueden emplearse como filtros dentro del sistema, permitiendo a los usuarios visualizar únicamente aquellos activos o vulnerabilidades que cumplan con ciertos criterios, lo que mejora la eficiencia en el análisis de datos y la priorización de acciones de mitigación.

2.3. Contextualización de tecnologías

2.3.1. Marco de Referencia Flask

Flask [2] es un framework ligero y flexible para el desarrollo web, ampliamente utilizado en proyectos que requieren escalabilidad y personalización. En este proyecto, Flask se utiliza tanto para el backend como para la integración del frontend, proporcionando las siguientes capacidades:

- Integración en tiempo real: Flask permite exponer APIs RESTful que conectan el frontend con el backend, garantizando una interacción fluida entre las capas.
- Procesamiento de riesgos: Las APIs desarrolladas con Flask manejan los cálculos de riesgo basados en datos de la superficie de ataque y modelos de machine learning.

2.3.2. Base de Datos SQLite

SQLite [3], [4] es un sistema de gestión de bases de datos relacional ligero y eficiente, ideal para proyectos que no requieren un servidor dedicado. En este proyecto, SQLite almacena datos históricos de vulnerabilidades, incidentes y riesgos, proporcionando una estructura optimizada para el acceso y análisis de información. Entre sus ventajas destacan:

- Consultas rápidas sobre grandes volúmenes de datos.

- Almacenamiento estructurado de los resultados de los escaneos realizados por la herramienta de reconocimiento.
- Persistencia de los registros de etiquetas asignadas a los activos.

Sin embargo, SQLite tiene limitaciones de tamaño que deben considerarse al diseñar sistemas escalables. El tamaño máximo de una base de datos en SQLite es de 281 terabytes (TB), pero su rendimiento puede degradarse significativamente cuando el archivo de base de datos supera los 10 GB [5], especialmente en sistemas con concurrencia alta [6]. Además, admite un máximo de 2,000,000 de filas aproximadamente por tabla y hasta 32,000 columnas en versiones recientes, aunque el manejo eficiente de estos valores depende del hardware y la configuración del sistema.

2.3.3. Bases de Datos Relacionales

Las bases de datos relacionales [7] son un modelo de gestión de datos ampliamente utilizado en aplicaciones empresariales y de ciberseguridad. Este modelo organiza la información en tablas estructuradas que están relacionadas entre sí mediante claves primarias y foráneas. Una de sus principales ventajas es la capacidad de manejar grandes volúmenes de datos de forma eficiente, garantizando la integridad y consistencia de la información mediante el uso de transacciones y restricciones.

En el contexto de ciberseguridad, las bases de datos relacionales se utilizan para almacenar información clave como registros de escaneos, inventarios de activos, vulnerabilidades detectadas y resultados de análisis de riesgos. Estas bases de datos permiten realizar consultas avanzadas para extraer información relevante en tiempo real, lo cual es fundamental para la toma de decisiones en entornos dinámicos y de alta exigencia.

2.3.4. Algoritmo no Supervisado Isolation Forest

El uso de modelos de aprendizaje de máquina ha revolucionado la gestión de riesgos en ciberseguridad, permitiendo la detección automatizada de anomalías y la priorización eficien-

te de activos vulnerables. En este proyecto, se implementa Isolation Forest [8], un algoritmo de Machine Learning no supervisado diseñado para la detección de anomalías en grandes volúmenes de datos.

Isolation Forest se basa en la premisa de que las anomalías son observaciones poco frecuentes y, por lo tanto, más fáciles de aislar que los datos normales. Su funcionamiento consiste en la construcción de múltiples árboles de aislamiento (Isolation Trees), donde cada nodo divide aleatoriamente los datos seleccionando una característica y un valor de partición. Las observaciones que requieren menos divisiones para ser aisladas son clasificadas como anomalías, ya que presentan un comportamiento atípico en el conjunto de datos.

Esta capacidad de identificar vectores de ataque críticos permite reducir la dependencia en la experiencia individual de los analistas y estandarizar los procesos de toma de decisiones. En el contexto de este proyecto, Isolation Forest se emplea para evaluar la superficie de ataque, detectar comportamientos anómalos y asignar niveles de criticidad a los activos en función de su exposición a amenazas. Su eficiencia computacional y capacidad para manejar grandes volúmenes de datos lo convierten en una solución ideal para la detección de amenazas en tiempo real dentro de entornos empresariales.

2.4. Contenedores Docker y Reconocimiento con Eyewitness

Docker [9] es una plataforma de contenedores que permite empaquetar aplicaciones y sus dependencias en un entorno aislado, facilitando su despliegue y portabilidad. En este proyecto, Docker es fundamental para ejecutar Eyewitness, una herramienta de ciberseguridad utilizada para capturar imágenes de pantalla de servicios expuestos y generar informes sobre posibles puntos vulnerables en la superficie de ataque [10].

El uso de Eyewitness permite obtener una referencia visual de los activos identificados, proporcionando documentación detallada para el análisis de seguridad. Docker garantiza que Eyewitness funcione de manera consistente y segura, sin afectar el rendimiento de la aplicación principal.

2.5. Riesgo Tecnológico

El análisis de riesgo en ciberseguridad es el proceso de evaluar la probabilidad e impacto de amenazas sobre los activos de una organización. El riesgo se determina como una función de la vulnerabilidad del activo, la probabilidad de explotación y el valor del activo para la empresa.

Para clasificar la criticidad de las vulnerabilidades, este proyecto adopta la metodología de evaluación del Common Vulnerability Scoring System (CVSS), desarrollado por el Forum of Incident Response and Security Teams (FIRST) [11]. CVSS asigna una puntuación de 0 a 10 a las vulnerabilidades, las cuales se agrupan en las siguientes categorías de severidad:

- Baja (0.1 - 3.9): Impacto y probabilidad de explotación reducidos.
- Media (4.0 - 6.9): Riesgo moderado, requiere mitigaciones oportunas.
- Alta (7.0 - 8.9): Vulnerabilidades con alto impacto o facilidad de explotación.
- Crítica (9.0 - 10.0): Fallos graves que pueden ser explotados de forma remota con consecuencias severas.

Si bien esta clasificación proporciona una medida estándar de criticidad, el riesgo final de un activo en este proyecto no se determina únicamente a partir de CVSS. En su lugar, el cálculo de riesgo considera también antecedentes históricos de incidentes y la superficie de ataque detectada en cada activo. De esta manera, dos activos con la misma vulnerabilidad pueden tener riesgos diferentes según su exposición real, su historial de ataques previos o la cantidad y tipo de servicios expuestos.

Para ello, la aplicación desarrollada emplea un modelo de machine learning, que integra datos de CVSS, superficie de ataque e incidentes históricos, generando una lista priorizada de activos en riesgo. Este enfoque permite una gestión de ciberseguridad más precisa y proactiva, enfocada en las amenazas que representan un peligro real para la organización.

2.6. Marcos de Referencia en Ciberseguridad

2.6.1. MITRE ATT&CK®

El marco MITRE ATT&CK® [12] proporciona un modelo estructurado de tácticas y técnicas utilizadas por adversarios en ciberataques. Este marco se organiza en 14 etapas consecutivas, desde el reconocimiento hasta el impacto, lo que permite a los especialistas sistematizar auditorías y desarrollar defensas más robustas. Además, facilita un lenguaje común para los equipos de seguridad, mejorando la comunicación y el entendimiento técnico dentro de organizaciones diversas.

2.6.2. Proyecto de Seguridad de Aplicaciones OWASP

El estándar de ejecución de pruebas de penetración (PenTest) de OWASP [13] complementa el marco MITRE ATT&CK® al definir siete fases claras que abarcan desde el reconocimiento pasivo hasta la post-explotación. Este enfoque proporciona guías prácticas para la ejecución de auditorías de seguridad y herramientas específicas para cada etapa. La integración de PenTest con ATT&CK® permite diseñar auditorías más completas, adaptadas a los contextos específicos de cada organización.

2.7. Principios de UX/UI y Diseño Universal en Ciberseguridad

Un aspecto clave en el desarrollo de herramientas de ciberseguridad es la usabilidad y accesibilidad de la interfaz de usuario. Un diseño centrado en el usuario (*User-Centered Design, UCD*) facilita la identificación y gestión de amenazas al reducir la complejidad en la presentación de datos críticos. La aplicación de principios de UX/UI, como la simplicidad visual, la jerarquización de la información y la retroalimentación en tiempo real, permite mejorar la interacción entre los especialistas en seguridad y los sistemas de monitoreo.

El diseño de interfaces de usuario en sistemas de ciberseguridad debe considerar principios de interacción que optimicen la toma de decisiones en entornos de alto riesgo. Estudios

han demostrado que la visualización de datos estructurada y la integración de elementos interactivos en tableros de visualización (*dashboards*) de seguridad mejoran la velocidad de respuesta ante incidentes críticos [14], [15].

Además, el Diseño Universal busca garantizar que estas soluciones sean accesibles para un amplio rango de usuarios, incluyendo personas con discapacidad o con diferentes niveles de conocimiento en ciberseguridad. Esto implica la adopción de estándares como el Web Content Accessibility Guidelines (*WCAG*), que aseguran una experiencia equitativa en la manipulación de datos de seguridad [16]. La accesibilidad digital es un factor clave en la adopción de herramientas de seguridad en organizaciones con equipos diversos [17].

En ciberseguridad, la combinación de UX/UI y principios de diseño universal permite que herramientas como tableros de visualización (*dashboards*) y paneles de análisis de amenazas sean más intuitivos y eficientes, optimizando el tiempo de respuesta ante incidentes [18]. La integración de estas prácticas con marcos de referencia como MITRE ATT&CK® y OWASP WSTG fortalece la seguridad organizacional sin comprometer la usabilidad.

2.8. Herramienta “DatosRecon”

La detección y análisis de vulnerabilidades en redes corporativas requieren herramientas especializadas que permitan mapear la superficie de ataque y detectar posibles puntos de explotación. En este proyecto, se emplea una herramienta de reconocimiento personalizado basada en el escáner de red Nmap [19], que permite analizar activos y servicios expuestos en entornos complejos.

Para optimizar su rendimiento, la herramienta implementa distintos métodos de escaneo, basados en los enfoques documentados en la literatura sobre reconocimiento de red y detección de vulnerabilidades [20]. Entre estos, destacan:

1. **Escaneo de red:** Detección de servicios expuestos y puertos abiertos en activos.
2. **Exploración de recursos web:** Identificación de aplicaciones y configuraciones en servicios HTTP/S.

3. **Enumeración de protocolos internos:** Análisis de recursos compartidos en red, como SMB y otros servicios críticos.
4. **Pruebas de autenticación básica:** Evaluación de credenciales débiles mediante técnicas de acceso controlado.

Estos escaneos permiten a los equipos de ciberseguridad priorizar riesgos, identificando activos con alta exposición y posibles configuraciones vulnerables.

2.9. Tecnologías y Metodologías de Reconocimiento

2.9.1. Reconocimiento Personalizado “DatosRecon”

El reconocimiento de activos en entornos corporativos requiere una estrategia personalizada que no solo automatice la detección de superficies de ataque, sino que también contextualice la información recopilada en función de tácticas y técnicas documentadas en el marco MITRE ATT&CK `mitre_att&ck`.

En este proyecto, el sistema DatosRecon permite ejecutar escaneos adaptativos, identificando servicios expuestos, puertos abiertos y configuraciones vulnerables, alineándose con técnicas de reconocimiento como:

- T1595 - Active Scanning: Exploración de infraestructura expuesta mediante barridos de puertos y detección de servicios accesibles.
- T1592 - Gather Victim Host Information: Identificación de sistemas, direcciones IP y configuraciones del objetivo para establecer un perfil de ataque.
- T1590 - Gather Victim Network Information: Mapeo de la topología de red y enumeración de activos conectados.

Además de permitir la recopilación estructurada de datos históricos, DatosRecon facilita la correlación de vulnerabilidades con eventos previos, reduciendo la duplicación de esfuerzos y optimizando el análisis de riesgo. Al integrar inteligencia artificial en este proceso,

CAPÍTULO 2 : SITUACIÓN ACTUAL

se mejora la capacidad de priorizar amenazas y tomar decisiones basadas en información consolidada y contexto real.

3. Metodología Aplicada

3.1. Arquitectura de la Solución

La solución propuesta está diseñada con una arquitectura modular, en la que cada componente cumple una función específica dentro del flujo de procesamiento de datos y análisis de ciberseguridad. La integración de estos módulos permite la detección, almacenamiento, normalización y evaluación del riesgo en la superficie de ataque de la organización. La interacción entre los módulos ocurre de la siguiente manera:

1. La herramienta DatosRecon realiza el escaneo de red de los activos, detectando puertos abiertos, servicios activos y configuraciones inseguras.
2. Los datos generados son almacenados en la Base de Datos Centralizada, asegurando su integridad y accesibilidad para su posterior análisis.
3. El Normalizador de Datos procesa y estructura la información, eliminando duplicados o inconsistencias y asegurando un formato unificado.
4. El Módulo de Machine Learning analiza la información estandarizada, identificando anomalías y calculando niveles de riesgo en los activos mediante algoritmos como Isolation Forest.
5. El Backend de la API actúa como intermediario entre los módulos anteriores y el Frontend de la Aplicación Web, procesando las solicitudes y gestionando las consultas de los usuarios.
6. Finalmente, el Frontend de la Aplicación Web proporciona una interfaz gráfica intuitiva, permitiendo a los usuarios visualizar la superficie de ataque, gestionar activos y priorizar riesgos de manera eficiente.

Este enfoque estructurado garantiza una interacción fluida entre todos los módulos, asegurando que los datos se procesen y analicen de manera óptima para la toma de decisiones en ciberseguridad

En la Figura 3.1 se presenta la arquitectura del sistema, mostrando la interacción entre sus principales módulos. El flujo de datos inicia con la entrada de información a través de DatosRecon, que es almacenada en la Base de Datos Centralizada. Posteriormente, los datos son normalizados y procesados por el Backend/API REST, que se encarga de gestionar la comunicación con el Frontend y el módulo de Machine Learning, el cual realiza el análisis de riesgos y retroalimentación del sistema.

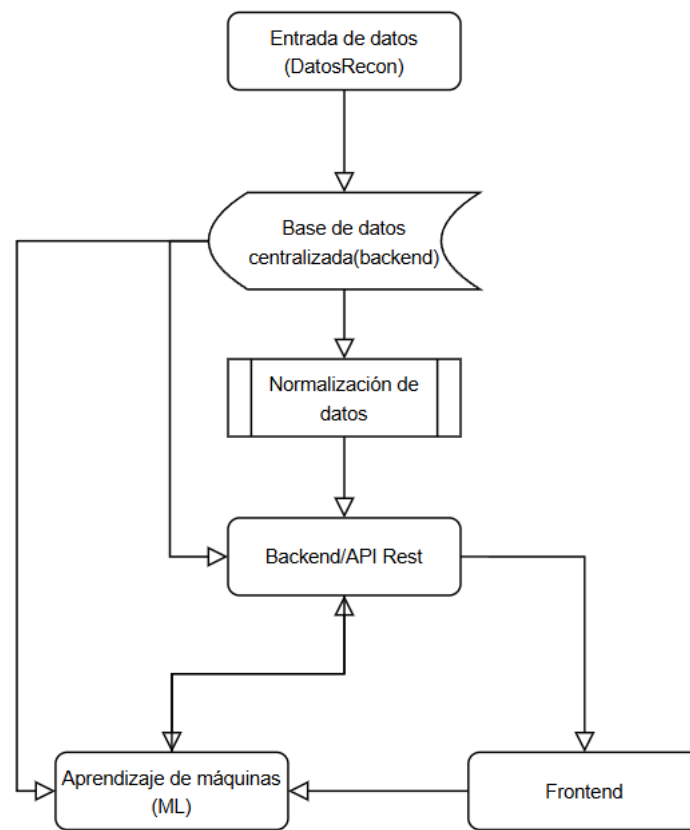


Figura 3.1: Arquitectura del sistema.

Componentes principales

La arquitectura se compone de los siguientes módulos principales:

3.1.1. Módulo de Reconocimiento Personalizado “DatosRecon”.

DatosRecon es el módulo encargado del escaneo de red, basado en Nmap, permitiendo la identificación de la superficie de ataque del Operador de Telecomunicaciones. Su objetivo principal es detectar puertos abiertos, servicios activos y configuraciones inseguras en los activos monitoreados.

Principales funcionalidades de DatosRecon:

- Escaneo de red: Identificación de puertos abiertos mediante técnicas como TCP SYN Scan.
- Enumeración de servicios: Detección de configuraciones expuestas, incluyendo SMB y credenciales débiles.
- Análisis de vulnerabilidades: Generación de reportes con activos y configuraciones inseguras detectadas.
- Integración con el sistema: Los datos recolectados se almacenan en la Base de Datos Centralizada para su posterior análisis.

Este módulo juega un papel clave en la detección temprana de amenazas, proporcionando la **información inicial** necesaria para el análisis de riesgos y la correlación con incidentes históricos en el sistema.

3.1.2. Base de Datos Centralizada (*Backend*)

La Base de Datos Centralizada es el componente encargado del almacenamiento estructurado de la información recolectada por DatosRecon, asegurando la integridad y disponibilidad de los datos críticos del sistema.

Principales funcionalidades:

- Gestión de datos de escaneos: Almacena los resultados obtenidos por DatosRecon para su análisis posterior.

- Registro de incidentes históricos: Permite la correlación de vulnerabilidades detectadas con eventos de seguridad previos.
- Análisis de riesgos: Organiza la información utilizada por el Módulo de Machine Learning para evaluar la criticidad de los activos.
- Integridad y consistencia: Implementa controles para garantizar la validez y seguridad de los datos almacenados.

En la Figura 3.2 se muestra el modelo de base de datos centralizada implementado en el sistema. Esta base de datos almacena la información recolectada durante los escaneos de seguridad, incluyendo activos, direcciones IP, incidentes, detecciones y vectores de ataque. Su diseño relacional permite mantener la integridad y coherencia de los datos, facilitando consultas eficientes y la correlación de eventos de seguridad.

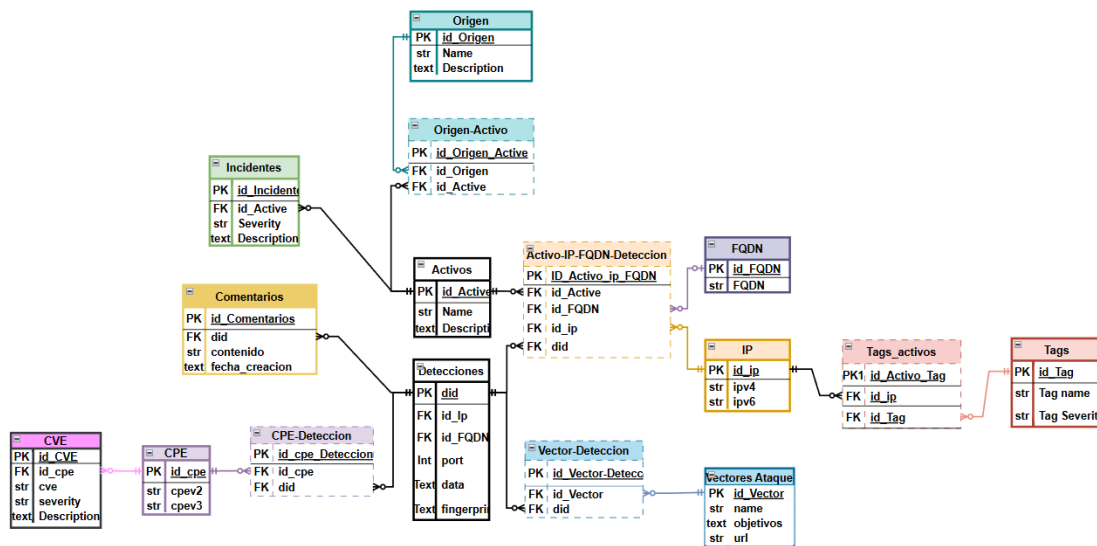


Figura 3.2: Modelo de base de datos centralizada del sistema.

3.1.3. Módulo Normalización de Fuentes de Datos (DataSources)

El Módulo de Normalización de Fuentes de Datos es responsable de procesar y estandarizar la información obtenida de diferentes fuentes, garantizando su calidad y coherencia antes

de ser utilizada en los análisis de riesgos.

Principales funcionalidades:

- Estandarización de datos: Convierte los datos a un formato estructurado y unificado para su procesamiento.
- Detección y eliminación de duplicados: Identifica registros repetidos y corrige inconsistencias en la información.
- Manejo de grandes volúmenes de datos: Procesa entre 10,000 y 200,000 registros de escaneo, asegurando eficiencia y precisión en el análisis.

3.1.4. Módulo de Aprendizaje de Máquinas (*Machine Learning*)

El Módulo de Aprendizaje de Máquinas es el encargado de analizar la superficie de ataque y calcular el nivel de riesgo de los activos mediante técnicas de detección de anomalías.

Principales funcionalidades:

- Identificación de anomalías: Utiliza Isolation Forest para detectar activos con comportamientos fuera de lo normal.
- Cálculo del nivel de riesgo: Estima la probabilidad de explotación de vulnerabilidades en cada activo monitoreado.
- Generación de modelos predictivos: Evalúa patrones en los datos para mejorar la toma de decisiones en ciberseguridad.

3.1.5. Desarrollo Full Stack

El sistema ha sido desarrollado bajo un enfoque Full Stack, integrando tanto el Frontend como el Backend para garantizar una comunicación fluida y eficiente entre la interfaz de usuario y la lógica de negocio.

1. Módulo de Interfaz Programática de la Lógica del Negocio (*Backend*)

El backend fue desarrollado utilizando Flask, un microframework ligero para Python, que permite la creación de APIs RESTful y la integración con modelos de Machine Learning para la evaluación de riesgos. Principales funcionalidades del backend:

- Procesamiento de datos: Validación y almacenamiento de información importada desde archivos Excel (*.xlsx) y JSON (*.json).
- Cálculo de riesgos: Ejecución del modelo Isolation Forest para detectar activos críticos.
- Interacción con el frontend: A través de APIs y renderización dinámica de datos.

Interacción con el Frontend

Para garantizar una comunicación eficiente entre ambas capas, se implementan dos enfoques complementarios:

a) APIs RESTful (Interfaz Programática de Aplicaciones para la Transferencia de Estados Representacionales) [21]: Las APIs RESTful permiten la comunicación estandarizada entre el cliente y el servidor mediante el protocolo HTTP. En este proyecto, se utilizan para el acceso y manipulación de datos en tiempo real, permitiendo operaciones como:

- Consulta de activos y vulnerabilidades.
- Asignación de etiquetas.
- Recuperación de información mediante solicitudes HTTP (GET, POST, etc.).

Este enfoque responde a la necesidad de interoperabilidad y escalabilidad, ya que REST es ampliamente adoptado en sistemas distribuidos, proporcionando una arquitectura ligera y flexible que facilita la integración con otros servicios en el futuro.

b) Proceso de transformación de datos con Jinja2 [22]: Jinja2 es un motor de plantillas para Python que permite la generación dinámica de contenido HTML mediante la inyección de datos preprocesados en el backend. En este proyecto, se emplea para:

- Renderizar plantillas HTML con información estructurada.
- Facilitar la visualización inicial de registros como activos, vulnerabilidades y etiquetas.

La elección de Jinja2 se debe a su eficiencia en el procesamiento de plantillas, su compatibilidad con Flask y su capacidad para reducir la carga en el frontend. Al procesar los datos en el servidor antes de enviarlos al cliente, mejora el rendimiento y reduce la latencia en la carga de páginas.

2. Módulo de Aplicación de Usuario e Interfaz Gráfica (*Frontend*)

El frontend fue desarrollado utilizando HTML, CSS y JavaScript, aprovechando Jinja2 para la renderización dinámica y consumo de las APIs RESTful proporcionadas por el backend. Principales funcionalidades del frontend:

- Visualización dinámica de datos: Tablas interactivas con funciones de búsqueda, filtrado y paginación de registros.
- Gestión de etiquetas: Asignación y edición de etiquetas directamente desde la interfaz gráfica.
- Validación de archivos: Subida de archivos con vistas previas para garantizar la calidad de los datos antes de su importación.
- Gestión de activos y riesgos: Visualización de la superficie de ataque con información detallada de cada activo, incluyendo incidentes históricos y niveles de riesgo.

Propósito del Módulo Frontend:

- Proporcionar una interfaz gráfica intuitiva para los usuarios.
- Facilitar la exploración y administración de los activos y vulnerabilidades detectadas.
- Permitir el filtrado dinámico y categorización de los registros.

3.2. Criterios de Severidad y Evaluación de Riesgos

La estimación y priorización del riesgo es un elemento central en la gestión de ciberseguridad. Para ello, se han desarrollado diferentes metodologías que permiten evaluar el impacto y la probabilidad de explotación de vulnerabilidades, facilitando la toma de decisiones para mitigar amenazas de manera efectiva.

Modelos de Estimación de Riesgo

Diversos enfoques han sido propuestos para estandarizar la evaluación del riesgo en sistemas de seguridad:

- Modelo FAIR (Factor Analysis of Information Risk): Este modelo permite cuantificar el riesgo cibernético mediante la evaluación de la frecuencia y magnitud de eventos de pérdida, proporcionando un enfoque estructurado para priorizar incidentes críticos.
- CVSS (Common Vulnerability Scoring System) [23]: Amplia y utilizada herramienta de clasificación que asigna un puntaje a vulnerabilidades basándose en factores como facilidad de explotación, impacto en la confidencialidad, integridad y disponibilidad.
- Herramientas de Visualización y Priorización: Sistemas que integran modelos predictivos y datos del reconocimiento de superficies de ataque para mapear los riesgos de forma visual e intuitiva, optimizando la toma de decisiones.

Implementación en el Sistema Desarrollado

En el presente trabajo, la evaluación de riesgos se realiza a partir de un enfoque basado en aprendizaje automático. Se emplea Isolation Forest para detectar anomalías en los activos escaneados, asignando un nivel de riesgo en función de la presencia de ciertos vectores de ataque. Además, los datos obtenidos del escaneo inicial son normalizados y procesados para garantizar su coherencia antes de ser analizados.

El sistema desarrollado implementa estos principios de priorización de riesgos, integrando modelos de análisis con visualización dinámica en el frontend. Esto permite que los analistas puedan explorar los resultados, aplicar filtros y tomar decisiones fundamentadas sobre la criticidad de cada activo monitoreado.

3.3. Flujo de Trabajo

El flujo de trabajo describe las etapas del proceso desde la generación de datos iniciales por DatosRecon hasta su análisis y presentación en la interfaz del sistema.

Pasos del Flujo de Trabajo

1. Entrada de Datos: Generación del Archivo de Escaneo

- **Descripción:** El proceso inicia con DatosRecon, que ejecuta escaneos sobre la red de activos de un Operador de Telecomunicaciones. Los resultados son exportados en un archivo Excel estructurado, que contiene información detallada de la superficie de ataque y los activos evaluados.
- **Salida:** Archivo Excel generado por DatosRecon.

2. Carga de Datos al Sistema

- **Descripción:** Los usuarios cargan el archivo Excel a través de la interfaz web del sistema. Este archivo es procesado para poblar la base de datos centraliza-

da, donde se almacenan los registros relacionados con los activos y su estado de seguridad.

- **Salida:** Datos integrados en la base de datos del sistema.

3. Almacenamiento en la Base de Datos

- **Descripción:** La información cargada es organizada y almacenada en la base de datos centralizada. Este paso asegura la disponibilidad de los datos para los procesos de análisis posteriores.
- **Salida:** Base de datos actualizada con los nuevos registros.

4. Normalización de los Datos

- **Descripción:** Un módulo de normalización procesa los datos cargados, ajustándolos a un formato estandarizado y eliminando posibles inconsistencias. Esto garantiza la integridad y calidad de los datos utilizados en los análisis posteriores.
- **Salida:** Datos normalizados y listos para su análisis.

5. Análisis de Riesgos

- **Descripción:** Los datos normalizados son procesados por un modelo de aprendizaje automático (Isolation Forest), que identifica anomalías y calcula un nivel de riesgo asociado a cada activo. Este análisis proporciona métricas objetivas para priorizar la atención a vulnerabilidades críticas.
- **Salida:** Informes de riesgos con priorización de activos según su criticidad.

6. Retroalimentación y Actualización

- **Descripción:** El sistema permite la retroalimentación mediante la carga de nuevos archivos Excel o la actualización manual de datos existentes. Este ciclo iterativo garantiza que la información del sistema esté actualizada y en constante mejora.

- **Salida:** Flujo continuo de datos actualizados para análisis iterativos.

7. Presentación de Resultados

- **Descripción:** Los resultados del análisis son accesibles a través del módulo frontend. Los usuarios pueden consultar dashboards interactivos para visualizar la superficie de ataque, filtrar resultados por criterios específicos y generar reportes personalizados.
- **Salida:** Visualización estructurada de los datos procesados en la interfaz del sistema.

La Figura 3.3 representa el flujo de procesamiento del sistema. DatosRecon realiza el escaneo de la red y almacena la información en la base de datos centralizada. Luego, el módulo de normalización estandariza los datos para su análisis de riesgo basado en aprendizaje de máquinas (ML), donde se identifican anomalías y se calculan niveles de riesgo. Posteriormente, los usuarios pueden retroalimentar y actualizar la información para mejorar la precisión del análisis. Finalmente, la API Backend conecta estos módulos con el Frontend, permitiendo la visualización y gestión de riesgos en una interfaz gráfica.

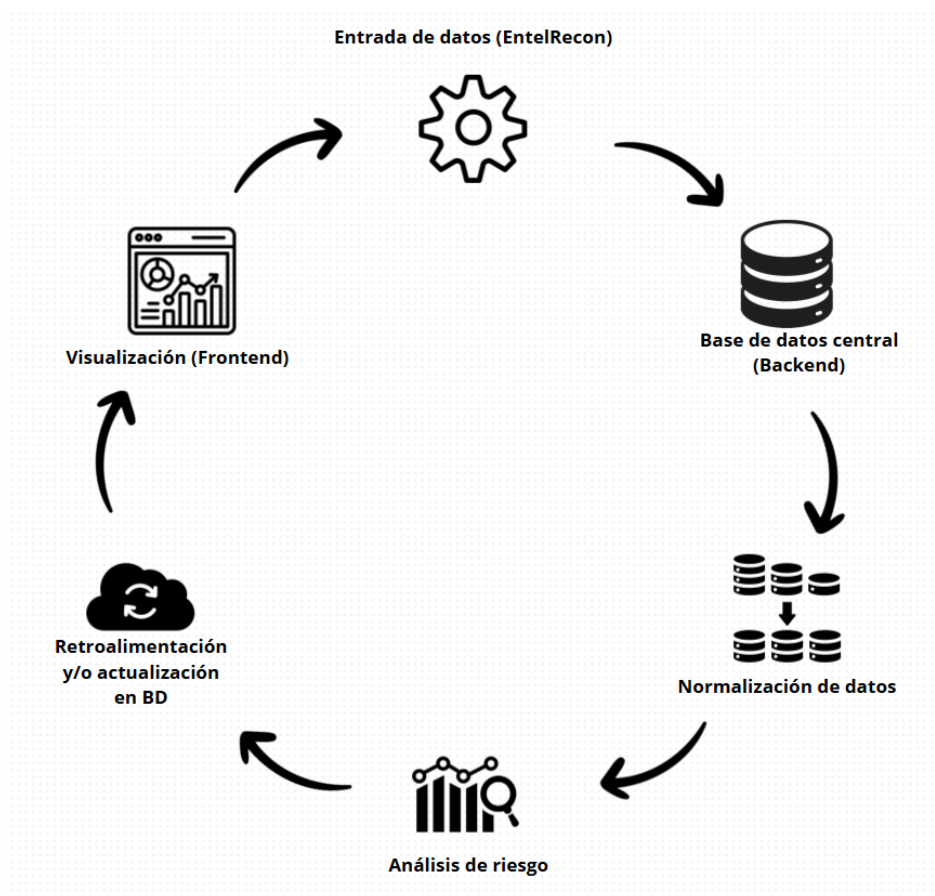


Figura 3.3: Diagrama del flujo de trabajo del sistema.

3.4. Implementación y Despliegue de la Arquitectura

En esta sección se describe la implementación técnica del sistema desarrollado, incluyendo los principales componentes del código, la estructura del proyecto y los enfoques utilizados para cumplir con los requisitos establecidos.

Estructura del Proyecto

El proyecto se implementó utilizando Flask como framework principal para el desarrollo del backend. La base de datos está gestionada con SQLAlchemy, lo que permite la abstracción de las consultas SQL mediante un ORM (Object Relational Mapping). La estructura del

proyecto incluye los siguientes elementos principales:

1. Carpeta principal del proyecto:

- *run.py*: Archivo principal para ejecutar la aplicación.
- *__init__.py*: Archivo para inicializar el proyecto y configurar los módulos necesarios, incluyendo Flask y SQLAlchemy.

2. Módulos y carpetas:

- *models.py*: Contiene las definiciones de las tablas de la base de datos y sus relaciones.
- *routes.py*: Define las rutas de la aplicación y las vistas para interactuar con los datos.
- *funciones.py*: Módulo que agrupa funciones utilitarias para procesamiento de datos, manejo de archivos y ejecución de EyeWitness.
- *constant.py*: Archivo que almacena constantes y configuraciones globales como vectores de ataque y pesos.
- *operaciones_db.py*: Contiene funciones específicas para operaciones con la base de datos.
- *templates/*: Carpeta que almacena los archivos HTML para las vistas del frontend.
- *static/*: Carpeta para recursos estáticos como hojas de estilo y archivos JavaScript.

Principales Funcionalidades

1. Importación de Datos

Se desarrolló una funcionalidad para importar datos desde archivos .xlsx y .json a la base de datos. Esta funcionalidad permite cargar información de detecciones, direcciones IP, y otros activos. La lógica principal incluye:

- Lectura de archivos en diferentes formatos.
- Validación de los datos antes de insertarlos en la base de datos.
- Creación de registros únicos en caso de duplicados (como IPs repetidas).
- Inserción eficiente mediante SQLAlchemy.

2. Visualización de Datos

Los datos importados se muestran en una tabla principal del frontend, con funcionalidades de búsqueda y filtrado. Se implementó:

- Renderizado Dinámico: Uso de Jinja2 para inyectar datos del backend en las plantillas HTML.
- Interactividad: Incorporación de JavaScript para manejo de eventos y actualizaciones en tiempo real.
- Paginación: División de los datos en páginas para optimizar la navegación y facilitar el acceso a los datos.
- Vistas específicas: Páginas como *detail.html* permiten editar y visualizar detalles de activos, mientras que *uploader.html* facilita la subida de archivos con validación previa.

3. Gestor de Tags

Se incluyó un sistema de gestión de etiquetas (tags) para categorizar y organizar los registros. La implementación permite:

- Asociar etiquetas a elementos seleccionados mediante un formulario POST.
- Mantener la integridad de las etiquetas en la base de datos.
- Mover el botón de *Crear Tag* a una sección específica para mejorar la experiencia del usuario.

4. Extracción de Datos

Una funcionalidad clave incluye la extracción de información específica de cadenas,

como el hostname, a partir de fingerprints. Esta información se inserta en la columna *Description* de la tabla *Activos* en la base de datos.

- **Algoritmo de Extracción:** Uso de expresiones regulares para identificar patrones y extraer los datos relevantes.
- **Actualización Automática:** Los registros existentes se actualizan automáticamente con la nueva información.

5. Detección de Vectores de Ataque

El sistema identifica vectores de ataque relevantes mediante la evaluación de columnas específicas en los datos importados. Los vectores con valores de riesgo (“1”) se destacan para su análisis posterior.

6. Cálculo de Riesgos y Machine Learning

Se emplea Isolation Forest como modelo principal para detectar anomalías y calcular el riesgo asociado a los activos. Este cálculo se integra con los datos almacenados para priorizar activos críticos y mejorar la toma de decisiones en ciberseguridad.

7. Visualización interactiva

El frontend incluye tablas dinámicas que permiten a los usuarios explorar y analizar los datos procesados. Estas tablas se integran con el backend mediante el uso de Jinja2, lo que permite:

- Filtrado por atributos específicos como nivel de riesgo o etiquetas asignadas.
- Paginación y búsqueda para facilitar la navegación en grandes volúmenes de datos.
- Actualizaciones dinámicas de los registros mostrados en respuesta a eventos del usuario.
- Gestión de Etiquetas (Tags): Los usuarios pueden interactuar con etiquetas asignadas a los registros, permitiendo:

- a) Agregar nuevas etiquetas a los activos desde la interfaz gráfica.
- b) Filtrar registros por etiquetas específicas para identificar categorías relevantes.
- c) Editar y eliminar etiquetas existentes de forma dinámica.

Tecnologías Utilizadas

- Lenguajes: Python, HTML, CSS, JavaScript.
- Frameworks: Flask, SQLAlchemy, Jinja2.
- Herramientas: SQLite para la base de datos, Bootstrap para diseño responsivo del frontend.

3.5. Pruebas de Rendimiento

Para evaluar la eficiencia del sistema en el procesamiento de grandes volúmenes de datos, se diseñaron pruebas de rendimiento considerando diferentes tamaños de archivos de entrada. Se establecieron dos escenarios de prueba con volúmenes de datos de 2,000 y 26,000 registros, midiendo:

- El tiempo de carga de los archivos en la base de datos.
- La cantidad de activos procesados correctamente.
- El porcentaje de activos clasificados como críticos.

Las pruebas fueron realizadas en un entorno controlado, garantizando condiciones uniformes para obtener resultados comparables. Los detalles de estas pruebas se presentan en la sección de resultados.

4. Resultados del Uso de la Plataforma

La presente sección detalla los resultados obtenidos durante la implementación y validación del sistema desarrollado para la gestión de riesgos en ciberseguridad de la organización. Los resultados se presentan en función de las funcionalidades principales implementadas, ilustradas mediante capturas de pantalla y ejemplos de uso real del sistema. Además, se analiza cómo estas funcionalidades cumplen con los objetivos planteados y su contribución a la mejora en la gestión de ciberseguridad en el contexto organizacional.

4.1. Resultados de la Interacción con la Superficie de Ataque

4.1.1. Visualización de la Superficie de Ataque

En la pantalla principal, los usuarios pueden observar una tabla interactiva que centraliza la información de los activos escaneados. Los elementos destacados incluyen:

- Tabla de activos: Presenta columnas para nivel de riesgo, IP, hostname, sistema operativo, puertos abiertos, vectores vulnerados y tags asignados.
- Resumen general: Muestra el número total de máquinas registradas.
- Niveles de riesgo: Un indicador visual que clasifica los activos según su nivel de criticidad (bajo, medio, alto, crítico).

En la Figura 4.1 se observa la tabla principal, donde se muestran 341 activos registrados. Los colores asociados al nivel de riesgo permiten identificar rápidamente los activos críticos que requieren atención prioritaria.

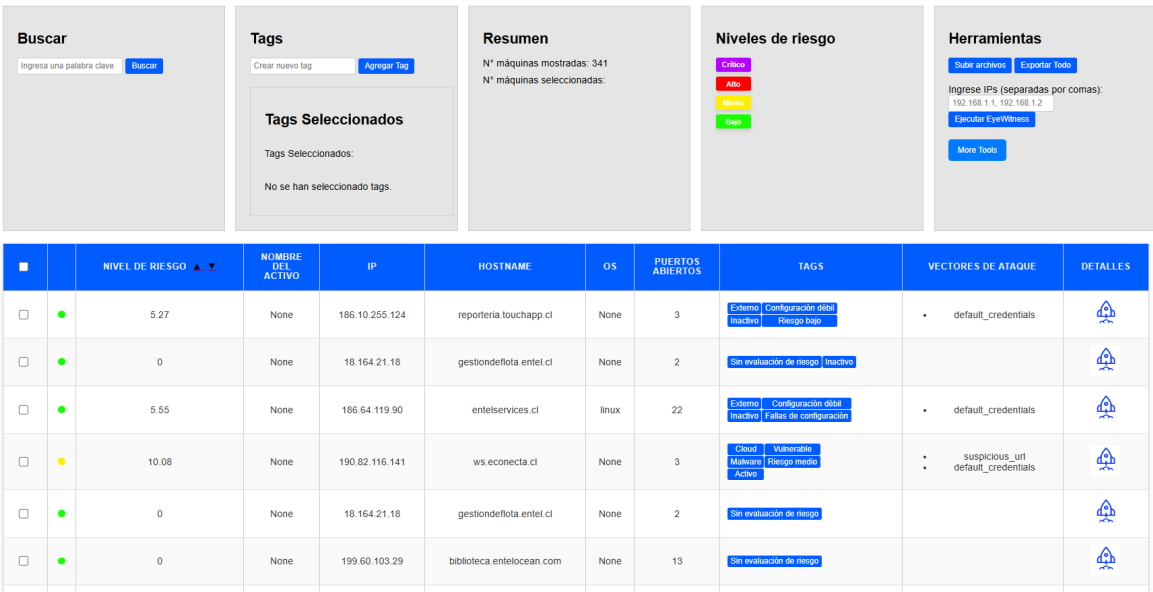


Figura 4.1: Tabla principal de visualización de activos en Inventarium.

Percepción del Usuario

Durante las pruebas de la aplicación, los usuarios destacaron varios aspectos clave de la visualización de la superficie de ataque:

- Usabilidad: La interfaz intuitiva y el uso de filtros dinámicos facilitan la búsqueda y organización de los activos escaneados.
- Rapidez: Gracias a la optimización de consultas en la base de datos y la paginación en la tabla, la carga de información es rápida y fluida, incluso con grandes volúmenes de datos.
- Precisión: La estructuración clara de la información y la integración de niveles de criticidad garantizan que los usuarios puedan identificar rápidamente los activos con mayor riesgo, permitiendo una toma de decisiones eficiente.

Estos factores mejoran la experiencia del usuario al interactuar con la aplicación, optimizando la gestión de riesgos en la red monitoreada.

CAPÍTULO 4 : RESULTADOS DEL USO DE LA PLATAFORMA

4.1.2. Resultados de la Interacción con Carga de Fuentes de Datos.

El sistema incluye una funcionalidad para cargar datos mediante archivos en la forma de planillas Excel con extensión *.xlsx y archivos de notación de objeto JavaScript con extensión *.json, permitiendo:

- La validación previa de los datos cargados.
- La integración automática de nuevos registros en la base de datos.

La Figura 4.2 muestra la pantalla de carga de datos, donde el usuario puede previsualizar el contenido del archivo antes de importarlo al sistema.

Sube tu archivo

Selecciona el tipo de archivo para subir:

[Subir Detecciones](#) [Subir Incidentes](#) [Subir Detecciones JSON](#)

[Subir Archivo](#)

Vista previa del archivo:

id	did	port	address	hostname	service
c57a9940-7c7c-43e9-8e93-7885e7ad892d	0660a5aa82e21a5e1f118030f2f6f107d	22	34.176.11.2	webmail.en.tel	'name': 'ssh', 'product': 'OpenSSH', 'version': '7.4', 'extraInfo': 'protocol 2.0', 'method': 'probed', 'conf': '10', 'cpe': 'cpe:/a:openssh:openssh:7.4'
ecc26e93-ac84-4a4b-891e-64c48c22b23f	0f383154715359b64ca5dbe48998b208	21	34.176.11.2	webmail.en.tel	'name': 'ftp', 'product': 'Pure-FTPd', 'method': 'probed', 'conf': '10', 'cpe': 'cpe:/a:pureftpd:pure-ftpd'
4661c037-a5d0-4ecd-83b3-c7080bcb665b	274225cb177464df651e17638f4725f	587	34.176.11.2	webmail.en.tel	'name': 'smtp', 'product': 'Exim smtpd', 'version': '4.96.2', 'hostname': 'whmi', 'method': 'probed', 'conf': '10', 'cpe': 'cpe:/a:exim:exim:4.96.2'
dc104949-ac5f-4515-9aa5-a090b6c4821c	793088bf2ec0a6111771529f99485590	443	34.176.11.2	webmail.en.tel	'name': 'http', 'product': 'Apache httpd', 'tunnel': 'ssl', 'method': 'probed', 'conf': '10', 'cpe': 'cpe:/a:apache:http_server'
ef72f88d-928e-4f8d-a00a-a5d2479b1d91	aeel16f0943f8705720b4823384d94c	465	34.176.11.2	webmail.en.tel	'name': 'smtp', 'product': 'Exim smtpd', 'version': '4.96.2', 'hostname': 'whmi', 'tunnel': 'ssl', 'method': 'probed', 'conf': '10', 'cpe': 'cpe:/a:exim:exim:4.96.2'
e2b2fc71-1171-474e-b893-9fa45af389d	b8ed1a8a4c1565976b9b96d310b531f2	3306	34.176.11.2	webmail.en.tel	'name': 'mysq', 'product': 'MySQL', 'extraInfo': 'unauthorized', 'method': 'probed', 'conf': '10', 'cpe': 'cpe:/a:mysql:mysql'
051e9dc8-6faa-4fa4-a7d2-276820f79ce6	c05450bab7ac9cbbad342824be2cc9b	80	34.176.11.2	webmail.en.tel	'name': 'http', 'product': 'Apache httpd', 'method': 'probed', 'conf': '10', 'cpe': 'cpe:/a:apache:http_server'
0293cf89-e19d-44d7-9b3b-97e2017479d0	ce13affe1776564dc713ea60ae2232a	53	34.176.11.2	webmail.en.tel	'name': 'domain', 'product': 'PowerDNS Authoritative Server', 'version': '4.7.3', 'method': 'probed', 'conf': '10', 'cpe': 'cpe:/a:powerdns:authorita
a82f8ec9-c754-4529-bedf-fa2f8b8e23be	fe4c0b02b1e46d0d21693e2f32bf17c	25	34.176.11.2	webmail.en.tel	'name': 'smtp', 'method': 'table', 'conf': '3'
6ec98eb5-0694-4abb-be3f-010e86458455	98deda5f6038de5d70c0ceace03ed5f1	443	34.160.23.96	cdn-eds.en.tel	'name': 'https', 'product': 'UploadServer', 'servicefp': 'SF-Port443-TCP:V=7.94SVN%T=SSL%I=7%D=5/13/Time=66427292P=x86_64-pc-lir
74789d08-fb04-4c1b-b1ca-55b669413c5	c5fc4f2167b0f6db8b8318fb064da216f	443	34.111.190.110	cdn-eds-dev.en.tel	'name': 'https', 'product': 'UploadServer', 'servicefp': 'SF-Port443-TCP:V=7.94SVN%T=SSL%I=7%D=5/13/Time=66427284P=x86_64-pc-lir
c757545e-1843-4b28-8110-5745ac2fb402	19e7f1d567e20eb22edc53dc9da5f39	443	34.128.139.195	cdn-eds-cms.en.tel	'name': 'https', 'product': 'UploadServer', 'servicefp': 'SF-Port443-TCP:V=7.94SVN%T=SSL%I=7%D=5/13/Time=66427279P=x86_64-pc-lir
52b25804-05ce-4457-48bc-06410c186429	ef06fba2024ee61f96bc67498e47495d	443	34.111.50.14	cdn-eds-uat.en.tel	'name': 'https', 'product': 'UploadServer', 'servicefp': 'SF-Port443-TCP:V=7.94SVN%T=SSL%I=7%D=5/13/Time=66427258P=x86_64-pc-lir
904ab09a-c846-4538-8470-7dde62ce26f7	ee5dde2354a811b87b57cbb75a4d9783	443	34.36.250.97	balancer.en.tel	'name': 'https', 'tunnel': 'ssl', 'method': 'probed', 'conf': '10'

Figura 4.2: Página de carga y validación de datos en Inventarium.

Percepción del Usuario

Durante la interacción con la funcionalidad de carga de datos, los usuarios destacaron los siguientes aspectos:

- **Fluidez y rapidez:** La carga de archivos se ejecuta de manera rápida y eficiente, permitiendo importar grandes volúmenes de datos sin afectar el rendimiento del sistema.
- **Flexibilidad:** La compatibilidad con archivos .xlsx y .json facilita la integración de diferentes fuentes de datos, adaptándose a diversas necesidades de los usuarios.

- Trazabilidad y control de versiones: El sistema no mantiene un control de versiones explícito de los datos cargados, pero sí registra la fecha y origen de cada importación, lo que permite identificar cuándo y cómo se integraron nuevos registros.

Estos aspectos contribuyen a mejorar la experiencia del usuario, asegurando que la carga de datos sea rápida, accesible y confiable.

4.1.3. Resultados de la Interacción con Etiquetas

El sistema permite filtrar los activos según las etiquetas (*tags*) asignadas. Los usuarios pueden seleccionar uno o varios tags desde la sección de *Tags Seleccionados*, lo que actualiza automáticamente la tabla para mostrar solo los activos que corresponden al filtro aplicado.

En la Figura 4.3, se muestra un acercamiento a la sección *Tags Seleccionados*, donde el usuario ha seleccionado el tag *Cloud*. Este tag se aplica como criterio de filtrado en la tabla principal.

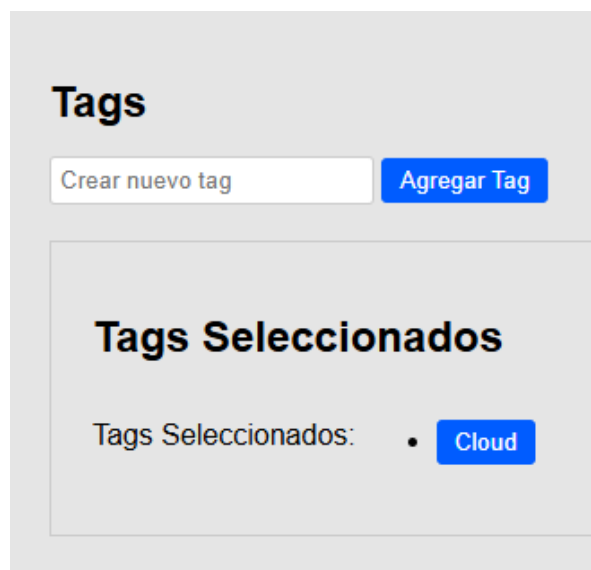


Figura 4.3: Zoom en la sección “Tags Seleccionado” con el tag “Cloud” activado.

Posteriormente, en la Figura 4.4, se observa la tabla principal actualizada, mostrando únicamente los activos que cumplen con el criterio del tag *Cloud*. Este proceso permite a los usuarios organizar y analizar los datos de manera eficiente.

CAPÍTULO 4 : RESULTADOS DEL USO DE LA PLATAFORMA

		NIVEL DE RIESGO ▲ ▼	NOMBRE DEL ACTIVO	IP	HOSTNAME	OS	PUERTOS ABIERTOS	TAGS	VECTORES DE ATAQUE	DETALLES
☐	●	9.67	None	190.82.116.141	ws.econecta.cl	None	3	Cloud Vulnerable Malware Riesgo medio Activo	- suspicious_url - default_credentials	
☐	●	17.36	None	186.10.255.130	corp.facturanet.cl	None	2	Cloud Activo	- web_login - suspicious_url - default_credentials	
☐	●	18.37	None	186.10.255.184	kintun.touchapp.cl	None	4	Cloud Activo	- default_credentials - dbms	
☐	●	0	None	186.10.95.11	securecloud.entel.cl	None	4	Cloud Activo		
☐	●	24.03	None	34.176.11.2	webmail.en.tel	None	9	Vulnerable Cloud Activo No esencial	- default_credentials - dbms	
☐	●	5.09	None	190.120.229.65	ofertas.entel.pe	None	2	Externo Cloud	- cve_analysis - default_credentials	

Figura 4.4: Tabla principal filtrada por el tag “Cloud”.

La Figura 4.5 muestra cómo, desde la sección de detalles de cada activo, se pueden crear y eliminar etiquetas para mejorar su clasificación. Esto permite a los usuarios personalizar la categorización de los activos según necesidades específicas

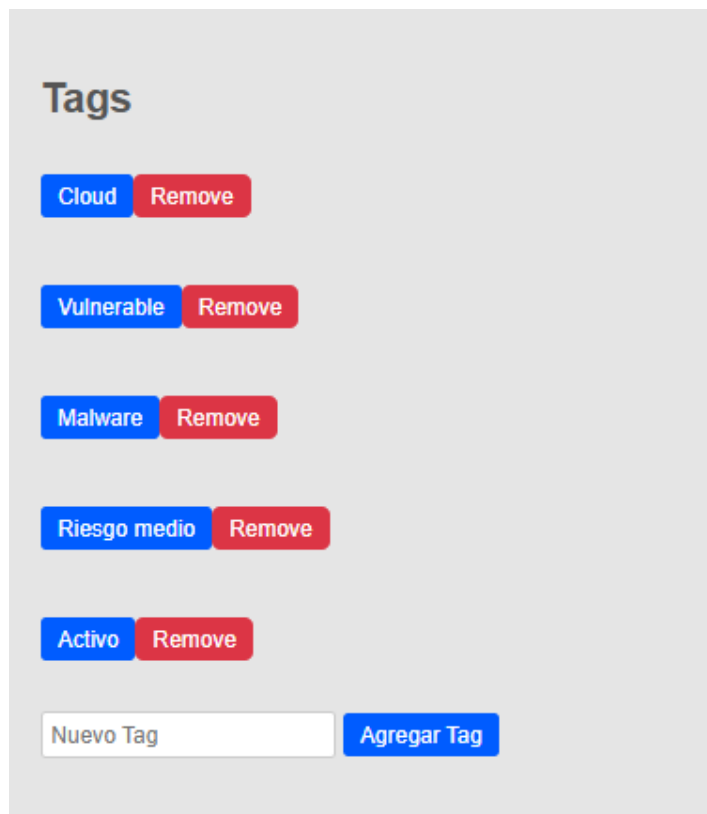


Figura 4.5: Gestión de etiquetas en la sección de detalles de un activo

Percepción del Usuario

Durante la interacción con el sistema de etiquetas, los usuarios destacaron los siguientes aspectos:

- Intuitividad en la asignación de etiquetas: La interfaz de gestión de tags es fácil de usar, ya que permite asignar, modificar y eliminar etiquetas con pocos clics, sin requerir configuraciones avanzadas.
- Facilidad de edición y eliminación de etiquetas: Desde la vista de detalles de cada activo, los usuarios pueden añadir y eliminar etiquetas de manera rápida, lo que mejora la flexibilidad en la clasificación de activos.
- Utilidad de las etiquetas en la productividad: El uso de filtros basados en tags facilita la búsqueda de activos específicos y optimiza el análisis de datos, permitiendo a los usuarios concentrarse en categorías relevantes sin revisar grandes volúmenes de información manualmente.

4.1.4. Resultados de la Interacción con Categoría de Riesgo

El modelo de machine learning Isolation Forest fue implementado para calcular el nivel de riesgo de cada activo. Este análisis ayuda a priorizar aquellos activos con anomalías detectadas.

En la Figura 4.1, se muestra la vista general de la tabla principal, donde se observa el nivel de riesgo asignado a cada activo. Para mejorar la claridad en la clasificación, se utilizó un sistema de colores que representa distintos niveles de criticidad:

- Bajo (*verde*): Activos sin anomalías detectadas, que requieren monitoreo estándar.
- Medio (*amarillo*): Activos con factores de riesgo moderados, donde se recomienda evaluación detallada.
- Alto (*rojo*): Activos con vulnerabilidades relevantes que requieren pronta mitigación.

- Crítico (*morado*): Activos con alto nivel de exposición y amenazas graves, demandando acción inmediata.

Este esquema de colores permite a los usuarios identificar rápidamente los activos prioritarios sin necesidad de analizar manualmente cada registro.

En la Figura 4.6, se presentan acercamientos a activos representativos con niveles de riesgo bajo, medio y crítico. Esto permite observar cómo se distribuyen los niveles de riesgo y destaca la importancia de priorizar acciones según la criticidad. Los activos con riesgo bajo pueden mantenerse en monitoreo continuo, mientras que los activos con riesgo crítico deben ser abordados de forma inmediata para mitigar vulnerabilidades graves.

■		NIVEL DE RIESGO ▲ ▼	NOMBRE DEL ACTIVO	IP	HOSTNAME	OS	PUERTOS ABIERTOS	TAGS	VECTORES DE ATAQUE	DETALLES
☐	●	5.03	None	186.160.127.244	miportal.entel.pe	None	1		default_credentials	
☐	●	0	None	192.168.226.128	None	None	29	Inactivo		
☐	●	24.11	None	135.181.141.38	acs.op.entel.cl	linux	11	Interno Bajo ataque Vulnerable	- default_credentials - dbms	
☐	●	5.03	None	186.10.94.98	amipassagentanyplace.touchapp.cl	None	12	Riesgo bajo Inactivo	- cve_analysis - default_credentials	
☐	●	9.21	None	135.181.141.38	acs.op.entel.cl	windows	11	Vulnerable Fallas de configuración Cloud Storage	default_credentials	

Figura 4.6: Zoom a activos con niveles de riesgo bajo, medio y crítico.

Percepción del Usuario y Uso de Colores en la Inclusividad

Durante la evaluación del sistema, los usuarios destacaron que el uso de colores como un mapa de calor facilita la comprensión del riesgo, especialmente para aquellos con perfiles no especializados en ciberseguridad.

- Claridad en la interpretación del riesgo: La diferenciación visual por colores evita ambigüedades y permite a cualquier usuario comprender rápidamente la severidad de las amenazas.
- Inclusividad para distintos roles: Profesionales de TI y administrativos pueden utilizar la plataforma sin necesidad de conocimientos avanzados en ciberseguridad, ya que la representación gráfica reduce la complejidad técnica.

- Filtrado eficiente: Los colores asociados a los niveles de riesgo permiten ordenar y segmentar activos de manera intuitiva, facilitando la toma de decisiones estratégicas según el perfil del usuario.

4.1.5. Detalles Específicos de Activos

La interfaz incluye una vista detallada para cada activo, donde se muestran las siguientes funcionalidades:

- Información específica, como fingerprint, puertos abiertos y vectores de ataque detectados.
- Opciones para agregar y eliminar comentarios asociados a cada activo, facilitando la documentación colaborativa.
- Una sección de captura de imágenes de la dirección IP vulnerada, permitiendo documentar visualmente los servicios expuestos.
- Un gestor de etiquetas específicas que permite agregar y eliminar tags para mejorar la organización de los activos.
- Una lista cronológica de incidentes, vectores de ataque, CVEs y fingerprints asociados, ordenados de los más recientes a los más antiguos. Cada elemento incluye una descripción detallada de su naturaleza y contexto.

En la Figura 4.7, se muestra una vista general de la página de detalles de un activo. A continuación, se presentan vistas ampliadas de las secciones más importantes:

CAPÍTULO 4 : RESULTADOS DEL USO DE LA PLATAFORMA

Detalles

IP: 34.176.11.2

Actualizar

Hostname: webmail.en.tel

Actualizar

OS: None

Actualizar

Nº puertos abiertos: 9

Puertos Abiertos: 21, 22, 25, 53, 80, 443, 465, 587, 3306

Nivel de riesgo: 22.83

Tags

Vulnerable Remove

Cloud Remove

Activo Remove

Nuevo Tag

Agregar Tag

Comentarios

Revisado por gabriel.
Creado el: 2025-02-27 00:21

Eliminar comentario

Escribe tu comentario aquí...

Agregar Comentario

Imágenes capturadas para la IP
34.176.11.2

No se encontraron imágenes para esta IP.

Capturar imágenes

Leyenda

Incidentes

Vectores de Ataque

CVE

FingerPrint

Limpiar filtros

2024.05.13 16:08:48

Fingerprint

"name": "mysql", "product": "MySQL", "extrinfo":
"unauthorized", "method": "probed", "conf": "10",
"cpe": "cpe:/a:mysql:mysql"

Detalles

2025.01.15 00:08:34

Incidente BUG-099 - Activo: None

Descripción del incidente: Esta vulnerabilidad
parte del supuesto en el que la variable
PHPSESSID almacena la sesión de un usuario
y este valor puede ser robado. De esta manera
un atacante podría acceder a la cuenta de otro
usuario y acceder a los roles o información que
este tenga.
Estado: Abierto
Días abiertos: 41

Detalles

2024.05.13 16:08:48

Vector de ataque default_credentials

None

Detalles

2024.05.13 16:08:48

Vector de ataque dbms

None

Detalles

Figura 4.7: Vista general de los detalles de un activo en Inventarium.

Sección de Información Específica La Figura 4.8 muestra los detalles del activo, como la IP, hostname, sistema operativo, puertos abiertos y nivel de riesgo.

39

Detalles

IP:

Hostname:

OS:

N° puertos abiertos: 9
Puertos Abiertos: 21, 22, 25, 53, 80, 443, 465, 587, 3306
Nivel de riesgo: 24.03

Figura 4.8: Información específica del activo.

Gestión de Comentarios La Figura 4.9 presenta la sección donde los usuarios pueden agregar y eliminar comentarios relacionados con el activo.



Comentarios

Prueba
Creado el: 2025-01-15 00:32
[Eliminar comentario](#)

Seguimiento
Creado el: 2025-01-15 00:53
[Eliminar comentario](#)

Escribe tu comentario aquí...

[Agregar Comentario](#)

Figura 4.9: Gestión de comentarios para el activo.

Captura de Imágenes En la Figura 4.10, se visualiza la sección para capturar imágenes de los servicios expuestos de la dirección IP vulnerada.

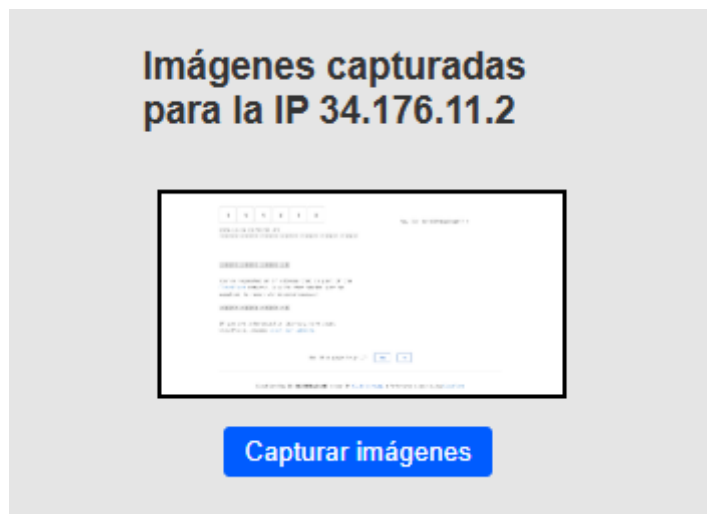


Figura 4.10: Sección de captura de imágenes de la dirección IP vulnerada.

Gestión de Etiquetas y Lista Cronológica

La gestión de etiquetas permite a los usuarios clasificar y filtrar activos de manera eficiente. Esta funcionalidad se evidencia en la Figura 4.5.

Además, en esta sección de detalles, los usuarios también pueden visualizar una lista cronológica de incidentes, vectores de ataque, CVEs y fingerprints asociados. La Figura 4.11 muestra esta lista organizada por orden temporal.



Figura 4.11: Lista cronológica de incidentes y vectores de ataque asociados al activo.

4.2. Resultados de Rendimiento del Prototipo

Comparativa de Rendimiento

La eficiencia del sistema fue evaluada bajo dos escenarios de carga, detallados previamente en la sección de Metodología Aplicada. En la Tabla 1 se presentan los tiempos promedio de carga, la cantidad de activos procesados y el porcentaje de activos clasificados como críticos en cada escenario.

Volumen de Datos (filas)	Tiempo Promedio de Carga (s)	Activos Procesados	Activos Críticos (%)
2,000	6.19	1,663	58 (3.48 %)
26,000	8.93	2,616	78 (2.98 %)

***Tabla 1:** Comparativa de rendimiento del sistema para diferentes volúmenes de datos.*

Análisis Comparativo

El tiempo de carga aumentó de manera proporcional con el volumen de datos, reflejando la eficiencia del sistema para manejar grandes archivos de entrada. A pesar del incremento en la cantidad de filas procesadas, el sistema mantuvo un rendimiento consistente.

La variación en el porcentaje de activos clasificados como críticos entre ambos escenarios (3.48 % frente a 2.98 %) sugiere que, a medida que el volumen de datos aumenta, la proporción relativa de activos críticos puede estabilizarse, lo cual es consistente con un análisis más representativo de una red extensa.

Estos resultados destacan la escalabilidad del sistema, evidenciando su capacidad para manejar grandes volúmenes de datos sin comprometer su funcionalidad.

5. Análisis de Resultados

Uno de los desafíos recurrentes en la ciberseguridad es la falta de criterios homogéneos para evaluar el riesgo de activos y vulnerabilidades. La ausencia de una metodología uniforme puede generar discrepancias en la clasificación de riesgos, afectando la precisión de las decisiones estratégicas.

En este sentido, el sistema desarrollado incorpora una estrategia de estandarización mediante la aplicación de modelos de riesgo basados en la severidad de vulnerabilidades como CVSS [23] y algoritmos de detección de anomalías como Isolation Forest. Esto permite un enfoque más estructurado y replicable en la priorización de amenazas, reduciendo la variabilidad en la toma de decisiones. La normalización de datos también juega un papel clave, asegurando que los registros de activos y vulnerabilidades sean consistentes en todas las evaluaciones.

No obstante, la implementación de este sistema en entornos empresariales requerirá una adaptación a normativas y estándares específicos del sector, como NIST, ISO 27001 y MITRE ATT&CK [12]. La alineación con estos marcos de referencia fortalecerá la confiabilidad del modelo y su aplicabilidad en escenarios reales.

5.1. Percepción de los usuarios

Para evaluar la efectividad del sistema en términos de usabilidad y precisión en el análisis de riesgos, se realizaron pruebas de usabilidad controladas durante reuniones con los usuarios. Durante estas presentaciones, los participantes interactuaron con la plataforma y proporcionaron retroalimentación en tiempo real sobre su experiencia.

Los especialistas en ciberseguridad valoraron la precisión del análisis de riesgo y la rapidez en la clasificación de amenazas, destacando la utilidad del modelo de aprendizaje automático para priorizar vulnerabilidades. Por otro lado, los usuarios con menor experiencia técnica resaltaron la facilidad de uso y la organización intuitiva de la información, aunque mencionaron la necesidad de documentación más detallada para interpretar ciertos elemen-

tos técnicos.

5.2. Logros Destacados

El desarrollo de este proyecto representa un avance significativo en la gestión de la ciberseguridad para Operadores de Telecomunicaciones, al implementar una solución escalable para la visualización y análisis de la superficie de ataque de sus activos. La integración de tecnologías modernas y metodologías avanzadas permitió alcanzar los objetivos planteados, proporcionando un sistema que optimiza la experiencia de usuario y la toma de decisiones en materia de seguridad.

Uno de los logros más relevantes es el diseño de una interfaz web basada en principios de usabilidad y accesibilidad, lo que ha permitido una gestión eficiente de los activos y riesgos. La implementación de visualizaciones interactivas, filtros personalizables y una organización clara de la información ha facilitado la identificación de activos críticos y la priorización de vulnerabilidades, reduciendo la complejidad del análisis.

Desde una perspectiva de análisis de datos, el sistema permitió correlacionar la información obtenida de la superficie de ataque con incidentes de seguridad previos, proporcionando un contexto más completo sobre la criticidad de los activos. Esto ha mejorado la capacidad de priorización de amenazas al evaluar no solo la exposición de cada activo, sino también su historial de incidentes, optimizando la asignación de recursos y fortaleciendo la capacidad de respuesta ante posibles ataques.

En términos de rendimiento, el sistema ha demostrado ser capaz de manejar grandes volúmenes de datos sin comprometer su estabilidad, permitiendo la importación y procesamiento de información de hasta 26,000 registros en tiempos óptimos. Además, se verificó su capacidad para soportar múltiples usuarios simultáneos sin degradación del servicio, lo que lo hace apto para su aplicación en infraestructuras en crecimiento.

5.3. Limitaciones del Sistema

A pesar de los avances logrados, el sistema presenta algunas limitaciones que podrían ser abordadas en futuros desarrollos:

- **Escalabilidad y Rendimiento:** El uso de SQLite ha sido adecuado para esta fase del proyecto, pero presenta restricciones al manejar grandes volúmenes de datos. Si bien el sistema logró procesar archivos de hasta 26,000 filas, el tiempo de carga aumentó proporcionalmente al tamaño del archivo. Se recomienda la implementación de bases de datos más robustas, como PostgreSQL o MongoDB.
- **Interactividad y Arquitectura del Frontend:** El sistema utiliza Flask y Jinja2 para la renderización de vistas dinámicas, lo que permite una integración eficiente con el backend. Sin embargo, la renderización en servidor limita la interactividad en comparación con tecnologías frontend modernas como React o Vue.js. La transición hacia una arquitectura desacoplada permitiría una mejor experiencia de usuario con actualizaciones dinámicas de datos.
- **Expansión del Perfil de Usuario:** Actualmente, el sistema está diseñado principalmente para especialistas en ciberseguridad. Sin embargo, otros perfiles dentro de la organización, como gestores de TI, auditores de seguridad o administradores de sistemas, podrían beneficiarse de la información generada sin requerir conocimientos avanzados en ciberseguridad. Se recomienda mejorar la accesibilidad del sistema mediante documentación detallada y paneles de análisis simplificados.

6. Conclusiones y Trabajo Futuro

Este trabajo ha desarrollado una plataforma efectiva para la gestión de la ciberseguridad en operadores de telecomunicaciones, proporcionando herramientas de análisis y priorización de riesgos basadas en aprendizaje automático. Los resultados obtenidos evidencian mejoras significativas en la identificación y mitigación de amenazas, facilitando la toma de decisiones en entornos de alta exigencia.

El sistema ha demostrado ser capaz de manejar grandes volúmenes de datos y múltiples usuarios simultáneos sin comprometer su estabilidad, lo que permite su aplicación en entornos empresariales reales. La interfaz web optimizada ha sido valorada positivamente por los usuarios, mejorando la experiencia en la gestión de activos y riesgos.

Sin embargo, aún existen áreas que pueden ser abordadas en futuras investigaciones. Se recomienda explorar modelos de machine learning más sofisticados para mejorar la precisión del análisis de riesgos, así como la implementación de bases de datos más robustas para optimizar el rendimiento en entornos con grandes volúmenes de información. Además, la transición a una arquitectura frontend desacoplada con tecnologías modernas como React o Vue.js permitiría una interactividad superior y una experiencia de usuario más fluida.

Asimismo, futuras investigaciones podrían enfocarse en mejorar la accesibilidad del sistema, diseñando interfaces más intuitivas y herramientas de apoyo para usuarios con menor experiencia técnica. Esto garantizaría que el sistema pueda ser utilizado por un público más amplio dentro de las organizaciones, maximizando su impacto en la gestión de la ciberseguridad.

Referencias

- [1] D. Everson y L. Cheng, “A Survey on Network Attack Surface Mapping,” *Digital Threats*, vol. 5, n.º 2, págs. 1-25, jun. de 2024. doi: 10.1145/3640019.
- [2] A. Ronacher. “Flask Documentation.” Accessed: March 6, 2025. (2010), dirección: <https://flask.palletsprojects.com/>.
- [3] F. Documentation. “Using SQLite 3 with Flask.” Accessed: March 7, 2025. (n.d), dirección: <https://flask.palletsprojects.com/en/stable/patterns/sqlite3/>.
- [4] D. R. Hipp. “SQLite Home Page.” Accessed: March 6, 2025. (2000), dirección: <https://www.sqlite.org/>.
- [5] H. Erz, “Why You Shouldn’t Use SQLite for Large-Scale Applications,” *Hendrik-Erz.de Blog*, jul. de 2023, Accessed: March 6, 2025. dirección: <https://www.hendrik-erz.de/post/why-you-shouldnt-use-sqlite>.
- [6] D. R. Hipp. “Limits In SQLite.” Accessed: March 12, 2025. (2000), dirección: <https://www.sqlite.org/limits.html>.
- [7] J. Sánchez, “Principios sobre bases de datos relacionales,” *Informe, Creative Commons*, vol. 11, pág. 20, 2004.
- [8] K. M. T. F. T. Liu y Z.-H. Zhou, “Isolation Forest,” *IEEE International Conference on Data Mining*, vol. 8, n.º 1, págs. 413-422, 2008. doi: 10.1109/ICDM.2008.17. dirección: <https://ieeexplore.ieee.org/document/4781136>.
- [9] S. Singh y N. Singh, “Containers and Docker: Emerging Roles and Future of Cloud Technology,” en *Proceedings of the 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)*, Bangalore, India: IEEE, 2016, págs. 804-807. doi: 10.1109/ICATCCT.2016.7912109.

REFERENCIAS

- [10] C. Truncer. “Eyewitness: A Tool for Capturing Screenshots of Services.” Accessed: March 6, 2025. (2024), dirección: <https://github.com/FortyNorthSecurity/EyeWitness>.
- [11] F. of Incident Response y S. T. (FIRST). “Common Vulnerability Scoring System (CVSS) v3.1 Specification Document.” Accessed: March 20, 2025. (2019), dirección: <https://www.first.org/cvss/v3-1/specification-document>.
- [12] M. Corporation, “MITRE ATT&CK: Design and Philosophy,” MITRE, inf. téc., mar. de 2020. dirección: https://attack.mitre.org/docs/ATTACK_Design_and_Philosophy_March_2020.pdf.
- [13] O. Foundation. “OWASP Web Security Testing Guide (WSTG).” Accessed: March 7, 2025. (2024), dirección: <https://github.com/OWASP/wstg>.
- [14] P. C. S. Kraemer y T. Sanquist, “Human factors and ergonomics in cybersecurity: Examining the potential of usability design,” *IEEE Transactions on Human-Machine Systems*, vol. 43, n.º 5, págs. 485-497, 2013. DOI: 10.1109/JSYST.2013.2260680. dirección: <https://ieeexplore.ieee.org/document/6636035>.
- [15] I. Hardy, V. J. Reyes, L. Phillips y M. Hamid, “Re-presentations and im-possibilities: the politics of dashboard data,” *Journal of Education Policy*, vol. 39, págs. 1-21, jul. de 2024. DOI: 10.1080/02680939.2024.2383655.
- [16] J. H. F. J. Lazar y H. Hochheiser, *Research Methods in Human-Computer Interaction*, 2nd. San Francisco, CA, USA: Morgan Kaufmann, 2017.
- [17] S. Harper e Y. Yesilada, *Web Accessibility: A Foundation for Research*. New York, NY, USA: Springer, 2018. DOI: 10.1007/978-1-4471-7440-0. dirección: <https://www.springer.com/gp/book/9781447174387>.
- [18] J. Nielsen y R. Budiu, *Mobile Usability*, 2nd. New York, NY, USA: New Riders, 2016.
- [19] A. Orebaugh y B. Pinkard, *Nmap in the Enterprise: Your Guide to Network Scanning*. Elsevier, 2011.

REFERENCIAS

- [20] G. Lyon, *Nmap Network Scanning - Official Guide*, Accessed: March 6, 2025, 2024. dirección: <https://nmap.org/book/man-examples.html>.
- [21] H. Subramanian y P. Raj, *Hands-On RESTful API Design Patterns and Best Practices*. Birmingham, UK: Packt Publishing Ltd, 2019.
- [22] J2Logo. “Tutorial Flask – Lección 2: Uso de plantillas para las páginas HTML.” Accessed: March 20, 2025. (n.d), dirección: <https://j2logo.com/tutorial-flask-leccion-2-plantillas/>.
- [23] K. S. P. Mell y S. Romanosky, “A Complete Guide to the Common Vulnerability Scoring System Version 2.0,” National Institute of Standards y Technology (NIST), inf. téc., jul. de 2007. dirección: <https://api.semanticscholar.org/CorpusID:13910056>.