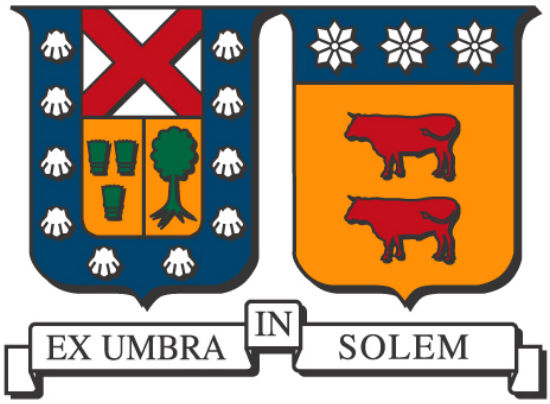
UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE ELECTRÓNICA
SANTIAGO - CHILE



"Detección y Mitigación de Tecnologías en la Sombra (Shadow IT) en Infraestructuras de Telecomunicaciones"

Gonzalo Soto Ulloa

MEMORIA DE TITULACIÓN PARA OPTAR AL TÍTULO DE INGENIERO CIVIL TELEMÁTICO

PROFESOR GUÍA: **Berioska Contreras Vargas**

PROFESOR CORREFERENTE: **Luis Lizama**

Abril - 2025

Resumen

El avance de la digitalización y la conectividad ha traído consigo nuevos desafíos en el ámbito de la ciberseguridad, siendo uno de los más relevantes el fenómeno conocido como "Tecnologías en la Sombra"(Shadow IT). Este concepto se refiere al uso de dispositivos, software y servicios tecnológicos no autorizados o no registrados dentro de la infraestructura oficial de una organización, lo cual amplía considerablemente la superficie de ataque y dificulta la gestión de riesgos.

Este trabajo tiene como objetivo desarrollar una plataforma integral denominada Inventarium, diseñada para identificar, mapear y gestionar activos tecnológicos en un entorno de telecomunicaciones mediante técnicas avanzadas de reconocimiento, escaneo de vulnerabilidades y cálculo de riesgos basados en Inteligencia Artificial.

La metodología propuesta integra un conjunto de herramientas de búsqueda y exploración de servicios y protocolos de red que operan fuera del control oficial de la organización. Además, se implementa el algoritmo de aprendizaje no supervisado Isolation Forest para evaluar el riesgo de cada activo.

En la fase de evaluación, se revisaron 795 activos tecnológicos, con una duración total de 6 horas y 30 minutos (aproximadamente 2 activos por minuto). Como resultado, se descubrieron 171 activos no registrados en el inventario oficial, representando un 20 % del total analizado. De estos, se identificaron 10 activos operando en entornos no productivos, asociados a dominios de desarrollo y control de calidad. Asimismo, se detectaron 40 activos con malas configuraciones relacionadas con servicios inseguros de administración remota y bases de datos, entre otro. También se identificaron 9 activos desconocidos con vulnerabilidades críticas conocidas (CVE) con alto riesgo de explotación.

Los resultados obtenidos demuestran que la solución planteada permite no solo identificar activos desconocidos y mal configurados, sino también priorizar su mitigación de acuerdo con su nivel de riesgo, mejorando así la postura de seguridad de la organización.

Abstract

The advancement of digitalization and connectivity has brought new challenges in the field of cybersecurity, one of the most significant being the phenomenon known as Shadow IT. This concept refers to the use of unauthorized or unregistered devices, software, and technological services within an organization's official infrastructure, which significantly expands the attack surface and complicates risk management.

This work aims to develop a comprehensive platform called Inventarium, designed to identify, map, and manage technological assets in a telecommunications environment through advanced techniques of recognition, vulnerability scanning, and AI-based risk calculation.

The proposed methodology integrates a set of tools for discovering and exploring network services and protocols operating outside the official control of the organization. In addition, the Isolation Forest unsupervised learning algorithm is implemented to assess the risk level of each asset.

During the evaluation phase, 795 technological assets were reviewed over a period of 6 hours and 30 minutes (approximately 2 assets per minute). As a result, 171 previously unregistered assets were discovered, accounting for 20 % of the total analyzed. Among them, 10 assets were found operating in non-production environments, identified by domain indicators such as qa, test, dev, and -pp. Furthermore, 40 assets exhibited poor configurations related to insecure services such as FTP, SSH, Telnet, MySQL, PostgreSQL, and Oracle. In addition, 9 assets were found with critical known vulnerabilities (CVEs) with a high risk of exploitation.

The results demonstrate that the proposed solution not only identifies unknown and misconfigured assets, but also enables the prioritization of their mitigation based on risk level, thereby improving the organization's security posture.

Keywords: Cibersecurity, analytics and risk management, Vulnerability scanners, Penetration testing, Domain-specific security and privacy architectures, software design engineering, Shadow iT.

1 Glosario

- **Activo:** Cualquier sistema informático conectado a la red de la empresa, que pueda ser comprometido, vulnerado o atacado de manera cibernética.
- **OWASP:** Organización sin fines de lucro dedicada a mejorar la calidad en seguridad del software, que contiene proyectos como OWASP TOP 10, documento que identifica los diez riesgos de seguridad más importantes en las aplicaciones web.
- **Superficie de Ataque:** Puntos de entrada y vulnerabilidades potenciales de un activo, sistema o red.
- **Ethical Hacking:** Práctica que consiste en simular ciberataques para identificar y corregir vulnerabilidades de seguridad en sistemas informáticos.
- **Shadow IT:** Sistemas, procesos y unidades organizativas desarrollados de manera autónoma por los departamentos empresariales, sin el conocimiento, aprobación o soporte del departamento oficial de TI. Esto incluye hardware, software y aplicaciones, que operan fuera del control de la infraestructura oficial de TI de una organización.
- **ICMP:** *Internet Control Message Protocol* es un protocolo de la capa de red utilizado para enviar mensajes de control y notificación sobre problemas en la transmisión de datos.
- **TCP:** *Transmission Control Protocol* es un protocolo de la capa de transporte que garantiza la entrega confiable de datos entre aplicaciones en una red.
- **ARP:** *Address Resolution Protocol* es un protocolo de la capa de enlace que traduce direcciones IP a direcciones MAC (físicas) dentro de una red local. Esto permite que los dispositivos en una red puedan comunicarse directamente en el nivel físico.

- **UDP:** *User Datagram Protocol* es un protocolo de la capa de transporte que permite el envío rápido de datos sin establecer conexiones. A diferencia de TCP, no garantiza la entrega ni el orden de los paquetes, lo que lo hace adecuado para aplicaciones en tiempo real como video y voz.

- **Banner:**Información que un servicio (por ejemplo, un servidor web o FTP) expone al establecer una conexión. Puede incluir detalles como el nombre y la versión del software, y es útil para la identificación de servicios y la detección de vulnerabilidades.

- **Firewall:** Sistema de seguridad, que puede ser hardware, software o una combinación de ambos, diseñado para monitorear, filtrar y controlar el tráfico de red entrante y saliente basado en políticas de seguridad predefinidas.

- **Fingerprint:** Conjunto único de características que permite identificar un sistema, dispositivo, software o protocolo. Estas características se derivan del análisis de respuestas específicas a estímulos enviados, como paquetes de red o solicitudes de conexión.

- **IP:** *Internet Protocol*, Protocolo estándar para la transmisión de datos desde la fuente a los destinos en redes de comunicaciones por conmutación de paquetes y sistemas interconectados de dichas redes.

- **CVE:** *Common Vulnerabilities and Exposures* es un sistema de nomenclatura estándar utilizado para identificar y categorizar vulnerabilidades de seguridad en sistemas y software. Cada CVE es un identificador único asignado a una vulnerabilidad o exposición conocida, que permite a organizaciones, investigadores y desarrolladores comunicar de manera clara y uniforme problemas de seguridad específicos.

- **JSON:** Formato de texto ligero y legible por humanos para el intercambio de datos estructurados. Está basado en la estructura de objetos de JavaScript, con

pares clave-valor y listas ordenadas.

- **XLSX:** Formato estándar basado en XML para archivos de hojas de cálculo creados por Microsoft Excel. Es un formato comprimido en ZIP que contiene múltiples archivos y carpetas internas que definen el contenido de la hoja de cálculo.
- **ftp:** Es un protocolo de red que permite transferir archivos entre equipos conectados a una red.

Tabla de Contenido

Índice de figuras

Índice de tablas

2 Introducción

Desde los primeros ciberataques realizados a sistemas como el telégrafo, la seguridad de la información ha evolucionado de una amenaza poco visible a una preocupación crítica para cualquier sistema informático (SI). Hoy en día, vivimos en un entorno hiperconectado donde la digitalización y la conectividad son omnipresentes en nuestra vida cotidiana. Esta realidad, aunque llena de beneficios, ha incrementado drásticamente la superficie de ataque para actores maliciosos.

La ciberseguridad se ha convertido en un pilar esencial para la protección de los datos y la continuidad operativa de organizaciones, especialmente en sectores sensibles como telecomunicaciones y banca. No obstante, la respuesta a las amenazas ha sido históricamente reactiva: se actúa después de que ocurren los incidentes, aplicando medidas correctivas en lugar de preventivas. Este enfoque ha comenzado a cambiar con el desarrollo de simulaciones de ciberseguridad, enfrentando equipos defensivos (Blue Team) y ofensivos (Red Team), en entornos controlados que permiten detectar y mitigar vulnerabilidades en infraestructuras conocidas.

Sin embargo, surge una pregunta crítica: ¿qué ocurre si algún activo de información que utilizan los trabajadores no está autorizado por la organización, pero igualmente es utilizado en el día a día? Esta práctica da origen a un fenómeno conocido como Tecnologías en la Sombra (Shadow IT), que implica el uso de software, hardware o servicios tecnológicos fuera del control oficial del área de TI. Tal como se plantea en Acken, Gadellaa, Jansen et al. [1], esto representa una amenaza creciente en diversas organizaciones. En un estudio aplicado a una universidad, se encontró una gran cantidad de herramientas tecnológicas utilizadas por estudiantes, docentes y funcionarios que no estaban bajo supervisión institucional, revelando que el 21 % de los participantes reportaron incidentes de ciberseguridad relacionados con Shadow IT.

Este problema no es exclusivo del ámbito académico. También afecta a los Operadores de Telecomunicaciones, los cuales, a pesar de tener mejores controles, manejan

una infraestructura crítica de gran escala con información sensible. En este contexto, la existencia de activos no autorizados o no inventariados representa un riesgo considerable, ya que impide tener una visión completa del panorama de seguridad, dejando potenciales brechas sin abordar.

Para enfrentar esta problemática, este trabajo propone una solución sistémica y experimental denominada Inventarium, una plataforma integral diseñada para identificar, mapear y gestionar activos tecnológicos *incluso aquellos que no han sido autorizado* dentro de un operador de telecomunicaciones nacional. Mediante técnicas avanzadas de descubrimiento, escaneo de vulnerabilidades y análisis de riesgo basado en inteligencia artificial, Inventarium permite obtener una visibilidad completa de la infraestructura tecnológica, facilitando la toma de decisiones y mejorando la postura de seguridad de la organización.

Este documento se estructura de la siguiente manera: el Capítulo 3 presenta el estado del arte y los antecedentes relacionados con Shadow IT y gestión de activos. El Capítulo 4 describe la metodología empleada, incluyendo las herramientas utilizadas, el enfoque de escaneo y la aplicación del algoritmo de aprendizaje automático. En el Capítulo 5, se presentan los resultados obtenidos tras la implementación de Inventarium en un entorno real. Finalmente, en el Capítulo 7, se discuten las conclusiones y se plantean líneas futuras de trabajo.

2.1. Objetivos

2.1.1. Objetivo General

Crear un sistema que, a partir de un identificador único de la red, mapee todos aquellos activos disponibles visualizando su superficie de ataque y sus vulnerabilidades históricas, y realizar una búsqueda comparativa personalizada para encontrar aquellos activos que no están registrados, buscando evitar la ocurrencia de tecnologías en la sombra en la organización.

2.1.2. Objetivos Específicos

- **Implementar un sistema de mapeo de activos a partir de un identificador único de red**, que permita visualizar la superficie de ataque y vulnerabilidades históricas de cada activo, y realizar búsquedas comparativas personalizadas orientadas a detectar activos no autorizados, contribuyendo a la prevención de tecnologías en la sombra en la organización.

- **Optimizar el proceso de búsqueda y exploración de activos de información conocidos y desconocidos** mediante la integración de la herramienta propietaria del operador de telecomunicaciones con otras soluciones complementarias, logrando una detección más eficiente y distribuida en la red para mejorar la cobertura y reducir los tiempos de identificación de dispositivos.

- **Detectar activos de información no autorizados (Shadow IT)** mediante la comparación automatizada entre los activos descubiertos y los registros oficiales de la organización, permitiendo identificar tecnologías utilizadas sin aprobación formal.

- **Evaluar los riesgos de seguridad asociados a activos no autorizados**, analizando el impacto potencial que estas tecnologías no gestionadas pueden generar sobre la infraestructura organizacional.

3 Marco Teórico

3.1. Impacto Negativo de tecnologías en las sombras (*Shadow IT*)

El uso de *Shadow IT* expone a las organizaciones a un mayor riesgo de ciberincidentes. Un informe de [2] revela que el 77 % de las empresas a nivel mundial han experimentado incidentes cibernéticos en los últimos dos años, de los cuales el 11 % fueron atribuibles directamente al uso de aplicaciones no autorizadas por parte de los empleados. Este dato subraya la magnitud de la amenaza que tecnologías en las sombras *Shadow IT* representa en términos de vulnerabilidades de seguridad no gestionadas.

Sin embargo, la definición y caracterización de la tecnologías en las sombras *Shadow IT* es compleja y en ocasiones contradictoria. Lejos de ser una práctica puramente negativa, también puede verse como una respuesta organizacional a la falta de flexibilidad, lentitud o burocracia en los departamentos centrales de TI [3]. De hecho, la proliferación de estas tecnologías "en la sombra" puede ser vista como una expresión de innovación local que responde directamente a necesidades específicas de los usuarios finales [4].

1. **Exposición a Vulnerabilidades de Seguridad:** El uso de aplicaciones y servicios no autorizados puede introducir vulnerabilidades que no han sido identificadas ni gestionadas por los equipos de TI. Estas plataformas no están sometidas a los estrictos controles de seguridad que las herramientas aprobadas por la empresa sí tienen, lo que aumenta significativamente el riesgo de que los datos confidenciales sean comprometidos. Según [5], las empresas que permiten el uso de tecnologías en las sombras *Shadow IT* se enfrentan a un aumento sustancial en las posibles brechas de seguridad y en el ataque de hackers externos. Silic y Back [6] sostienen que estas prácticas no son necesariamente maliciosas, sino que muchas veces surgen del deseo legítimo de los usuarios de ser más eficientes en sus tareas diarias mediante el uso de herramientas populares como CCleaner,

VLC o PDFCreator.

2. **Filtración de Datos Confidenciales:** El uso de plataformas no autorizadas puede resultar en la filtración de datos personales, como información financiera combinada con datos individualizados de clientes. Un caso reciente subraya la importancia de monitorear el uso de servicios en la nube: el 77 % de los usuarios de IA generativa introducen esta tecnología en su trabajo sin la supervisión adecuada de la empresa, lo que incrementa el riesgo de filtraciones de datos corporativos [7]. Este dato refleja una tendencia creciente, ya que muchas organizaciones no han adaptado sus políticas de TI a la incorporación de nuevas tecnologías por parte de los empleados.

3. **Dificultades en la Gestión de Activos Tecnológicos:** La proliferación de herramientas no controladas hace que la gestión de activos tecnológicos sea mucho más compleja. Los departamentos de TI tienen dificultades para llevar un inventario preciso de las aplicaciones y dispositivos utilizados en la organización, lo que puede llevar a una gestión ineficiente y, por lo tanto, a la exposición de la empresa a riesgos adicionales. Como indica [5], esta falta de visibilidad sobre el uso de tecnologías no autorizadas puede aumentar la vulnerabilidad de las organizaciones a ciberataques. En un estudio realizado por [8], se evidenció que en entornos universitarios, más del 21 % de los participantes reportaron incidentes directamente relacionados con tecnologías no autorizadas utilizadas en el día a día académico.

El impacto de tecnología en las sombras *(Shadow IT)* se está intensificando debido a varias tendencias tecnológicas. Un informe reciente indica que el 77 % de los usuarios de IA generativa están introduciendo esta tecnología sin supervisión corporativa [7], lo que refleja cómo el uso de tecnologías emergentes sin control puede poner en riesgo a las empresas. Además, el cibercrimen ha crecido un 400 % entre 2022 y 2024, con el uso de tecnologías no aprobadas siendo un factor clave en este aumento. Otro indicador

importante es el creciente uso de tecnología en las sombras en el contexto de trabajo remoto, donde la necesidad de acceder a servicios de manera rápida y eficiente ha llevado a un aumento en el uso de software no aprobado [9].

A pesar de estos riesgos, algunos autores como Haag y Eckhardt [10] sugieren que Shadow IT puede representar una oportunidad de innovación. Plantean que en vez de ser eliminado, puede ser gobernado estratégicamente, reconociendo que las soluciones desarrolladas por los usuarios pueden ser más efectivas que las ofrecidas por TI central. En esta línea, Zimmermann y Rentrop [4] proponen una redistribución clara de responsabilidades que permita absorber y controlar estas iniciativas desde un enfoque formalizado.

Para mitigar los riesgos asociados con tecnología en las sombras *(Shadow IT)*, las organizaciones deben implementar políticas claras sobre el uso de tecnologías, mejorar la colaboración entre los departamentos de IT y los empleados, y proporcionar capacitación continua en seguridad de la información. La supervisión de las tecnologías emergentes y la adaptación a las necesidades de los trabajadores también son esenciales para evitar incidentes de seguridad y proteger los datos confidenciales de la empresa. La clave está en equilibrar el control con la flexibilidad, permitiendo aprovechar lo mejor del Shadow IT sin comprometer la seguridad organizacional.

En este contexto, comprender el fenómeno de las tecnologías en la sombra ya no es solo una necesidad técnica, sino una condición estratégica para proteger la infraestructura digital y fomentar la innovación segura en las organizaciones.

3.2. Tácticas, técnicas y procedimiento ofensivos ligados a las tecnologías en la sombra.

3.2.1. Movimiento Lateral

El *Movimiento Lateral* se refiere a un conjunto de técnicas que los atacantes utilizan para obtener acceso y control sobre sistemas remotos dentro de una red comprometida.

Este proceso implica explorar la red con el propósito de identificar y comprometer nuevos objetivos. Para lograrlo, los atacantes a menudo se desplazan entre múltiples sistemas y cuentas, ya sea utilizando herramientas de acceso remoto que instalan ellos mismos o aprovechando credenciales legítimas junto con herramientas nativas de la red y del sistema operativo, lo cual les permite operar de manera más discreta y difícil de detectar. Este método afecta especialmente a la tecnología en la sombra, ya que, los atacantes pueden alojarse en activos olvidados con vulnerabilidades explotables, que luego utilizan como punto de partida para pivotar al resto de la red.

Algunas de estas técnicas incluyen:

- **Explotación de Servicios Remotos:** Los atacantes pueden explotar servicios remotos para obtener acceso no autorizado a sistemas internos una vez dentro de la red. Esto ocurre cuando aprovechan errores de programación en aplicaciones, servicios, software del sistema operativo o incluso en el propio kernel para ejecutar código bajo su control. Uno de los objetivos principales de este tipo de explotación es facilitar el movimiento lateral, lo cual les permite acceder a otros sistemas remotos.

- **Transferencia de Herramientas Laterales:** Los atacantes pueden transferir herramientas u otros archivos entre sistemas comprometidos. Este movimiento de archivos ocurre después de comprometer un sistema inicial, permitiendo que los archivos maliciosos se copien de un sistema a otro sin ser detectados.

- **Secuestro de Sesión de Servicio Remoto:** Los adversarios pueden tomar el control de sesiones preexistentes establecidas con servicios remotos como Telnet, SSH o RDP. Cuando un usuario legítimo inicia sesión en un servicio, se crea una sesión que permite la interacción continua con dicho servicio. Los atacantes pueden secuestrar estas sesiones para moverse lateralmente dentro del entorno comprometido.

- **Servicios Remotos:** Los adversarios pueden utilizar cuentas válidas para auten-

ticarse en servicios que aceptan conexiones remotas, tales como Telnet, SSH o VNC. Este método permite a los atacantes conectarse a sistemas remotos usando credenciales legítimas y herramientas estándar del sistema operativo.

- **Replicación a través de Medios Extraíbles:** Los atacantes pueden comprometer sistemas, incluso aquellos aislados o en redes desconectadas, copiando malware en medios extraíbles. Este malware puede ser ejecutado automáticamente al insertar los medios en un sistema objetivo o ser renombrado para parecer un archivo legítimo, engañando a los usuarios para que lo ejecuten.

- **Contaminación de Contenido Compartido:** Los atacantes pueden introducir cargas maliciosas en ubicaciones de almacenamiento compartido, como unidades de red o repositorios de código internos. Esta técnica consiste en modificar archivos válidos o insertar nuevos archivos maliciosos que se ejecutan cuando los usuarios acceden al contenido compartido. Al abrir dicho contenido, el código del adversario puede ejecutarse, permitiendo así su movimiento lateral.

3.3. Superficie de Ataque

Superficie de ataque se define como el conjunto total de puntos de entrada o vectores que un atacante puede utilizar para intentar acceder a un sistema, red o aplicación.Delimita los posibles caminos que podrían ser explotados por amenazas externas o internas. Estos se pueden segregar en tres componentes:

- Superficie Digital: Incluye elementos como aplicaciones web, servicios en la nube, APIs y bases de datos accesibles en línea.

- Superficie Física: Comprende dispositivos de hardware suceptibles como lo son servidores, estaciones de trabajo, dispositivos móviles y access point.

- Superficie Humana: Relacionada a todo factor humano, como lo pueden ser configuraciones incorrecta que se apliquen, uso de contraseñas débiles o inseguras,

caer en ataques de ingeniería social, pishing, o un insider, que es un trabajador que desde dentro de la empresa ayuda a un ciber atacante a sus cometidos.

Dentro del mundo de las simulaciones de ciberseguridad, existen varias metodologías a seguir para el mismo fin, por lo que es importante definir que se seguirá lo contemplado por OWASP que es una organización sin fines de lucro dedicada a mejorar la calidad en seguridad del software. Dentro de esta organización se creo un estándar de ejecución para las pruebas de penetración, que consta de 5 etapas:

Figura 3.1: *Etapas del Pentesting.*

1. **Planificación y Reconocimiento**

   En esta etapa se busca la recopilación de información para identificar posibles superficies de ataque o vulnerabilidades, para ello se hace revisión de dominios públicos, OSINT (Open source Intelligence), Recolección de direcciones IP, DNS, y correos electrónicos.

2. **Reconocimiento Activo**

   En este proceso el rastreo es más intrusivo y detectable en los sistemas, ya que, se empiezan a analizar estos identificando diferentes puntos de entradas como lo

son el escaneo de puertos, identificación de servicios y versiones, enumeración de usuarios y recursos.

3. **Obtención de acceso**

   Al ya saber la superficie de ataque que nos encontramos, se explotan las vulnerabilidades identificadas para acceder al sistema, utilizando explotaciones conocidos, ataques de fuerza bruta, inyección de código, entre otros.

4. **Mantenimiento del acceso**

   Se establecen ciertos métodos para mantener el acceso no autorizado al sistema durante el tiempo que duren las pruebas, como lo pueden ser un puerta trasera (*backdoor*), escalamientos de privilegios y movimientos lateral dentro de la red.

5. **Análisis y Reporte**

   Dado que el periodo de pruebas termino, se hace un reporte al cliente de todos los hallazgos, vulnerabilidades explotadas y en algunos casos se proponen soluciones.

4 Metodología y Arquitectura del Sistema

En este capítulo se presenta la metodología e implementación realizada en el proyecto, describiendo cada una de las etapas y acciones que busca alcanzar cada una de ellas. Además, se detallan los resultados obtenidos en cada proceso.

Como se puede apreciar en 3.1, resumen lo anteriormente presentado para cada etapa, es por ello que este trabajo enfocará la etapa 2 y 3 del proceso, donde se trabajará con reconocimiento Activo de puertos y servicios de la red. Para ello se utiliza la herramienta NMAP.

4.1. Etapas del Experimento

Este capítulo presenta la metodología aplicada y la arquitectura desarrollada para abordar la problemática de tecnologías en la sombra (*Shadow IT*). Se estructura en tres etapas que dan respuesta directa a los objetivos específicos del proyecto: optimizar el descubrimiento de activos, detectar activos no autorizados y evaluar su riesgo de seguridad. Además, se detallan las herramientas utilizadas y la lógica del sistema implementado.

4.1.1. Etapa 1: Optimización del Descubrimiento de Activos (DatosRecon)

Para abordar el descubrimiento efectivo de activos de red, se desarrolló una herramienta personalizada denominada **DatosRecon**, basada en Nmap pero con mejoras específicas. Esta herramienta optimiza el proceso de reconocimiento al incorporar técnicas de escaneo personalizadas que permiten descubrir no solo puertos y servicios, sino también elementos ocultos o mal configurados que suelen pasar desapercibidos por escaneos tradicionales.

DatosRecon fue configurado para mejorar las cabeceras de búsqueda, integrar vectores de ataque adicionales y aumentar la sensibilidad en la identificación de servicios. Además, se le añadió un módulo complementario que permite analizar activamente URLs entregadas por los servicios web encontrados. Estas URLs son utilizadas por otra herramienta, EyeWitness, para realizar una verificación visual de los servicios en ejecución.

Esta etapa permitió la construcción de un inventario centralizado donde, a partir de una IP, se asocian múltiples atributos como sistema operativo, puertos abiertos, servicios, comentarios y vectores de ataque.

4.1.2. Etapa 2: Detección de Activos No Autorizados (DatosRecon + EyeWitness)

Una vez descubierto el universo de activos presentes en la red, se realiza una detección comparativa para identificar aquellos que no están registrados oficialmente. Esto se realiza mediante la comparación automatizada entre los datos obtenidos por **DatosRecon** y los registros oficiales de la organización.

NMAP Nmap (*Network Mapper*) es una herramienta de código abierto utilizada para auditorías de red y seguridad. En este proyecto, Nmap fue utilizado para:

- Identificar hosts activos mediante paquetes TCP, ICMP o ARP.
- Detectar puertos abiertos mediante escaneos SYN, CONNECT o UDP.
- Determinar versiones de servicios y sistemas operativos mediante fingerprinting del stack TCP/IP.
- Automatizar tareas con el uso del motor de scripting NSE.

EyeWitness Para complementar la información obtenida por Nmap, se utilizó la herramienta **EyeWitness**, que permite capturar imágenes de servicios web accesibles (HTTP, RDP, VNC, etc.). Esta herramienta está contenida en un contenedor Docker, lo que facilita su despliegue y posterior eliminación, dejando solo los datos recolectados. EyeWitness permite validar de forma visual los activos descubiertos y verificar si están en uso o presentan configuraciones inseguras [11].

El análisis cruzado entre los resultados de estas herramientas y el inventario existente permitió detectar tecnologías en la sombra (activos no autorizados) utilizadas por la organización pero que no estaban bajo control oficial.

4.1.3. Etapa 3: Evaluación del Riesgo de Seguridad (CVSS + IA)

Una vez identificados los activos no autorizados, se procedió a evaluar el riesgo que estos representan para la organización. Para ello, se empleó una combinación entre el estándar internacional **CVSS** y un modelo de inteligencia artificial basado en **Isolation Forest**.

CVSS (Common Vulnerability Scoring System) El **CVSS** es un sistema utilizado para medir la gravedad de las vulnerabilidades en activos informáticos. Se divide en tres métricas principales:

1. **Métricas Base:** Miden el impacto directo de una vulnerabilidad (confidencialidad, integridad, disponibilidad) junto con su facilidad de explotación (vector de ataque, complejidad, privilegios requeridos).
2. **Métricas Temporales:** Consideran la disponibilidad de exploits y madurez de soluciones.
3. **Métricas del Entorno:** Ajustan el puntaje según la criticidad del activo afectado y la tolerancia al riesgo de la organización [12].

El CVSS proporciona un puntaje cuantitativo entre 0.0 y 10.0 que fue utilizado como base para caracterizar las vulnerabilidades de cada activo donde dependerá el valor este para saber si es critico/alto/medio/bajo, estos valores se reparten de la siguiente manera:

Puntaje CVSS	Nivel de Severidad	Interpretación
0.0	Ninguna	Sin impacto o sin explotación posible
0.1 – 3.9	Bajo	Impacto leve, difícil de explotar
4.0 – 6.9	Medio	Impacto moderado, explotación posible
7.0 – 8.9	Alto	Riesgo alto, requiere atención
9.0 – 10.0	Crítico	Muy grave, explotación fácil y daño severo

Tabla 1: *Clasificación de severidad según puntaje CVSS v3.1*

Modelo IA: Isolation Forest Dado que cada organización tiene una estructura distinta y activos con diferentes grados de criticidad, se implementó un modelo de aprendizaje automático no supervisado llamado **Isolation Forest**, especialmente diseñado para detectar anomalías en los datos.

Este algoritmo crea múltiples árboles binarios que aíslan los puntos de datos. Los activos que son aislados con menos divisiones son considerados más anómalos, es decir, con mayor riesgo. Este enfoque permite calcular un **riesgo relativo** considerando el comportamiento del resto de los activos en la organización, ofreciendo así una métrica contextualizada y personalizada [13]. Cabe destacar que la selección de este algoritmo se basa en que como no se tienen los datos etiquetados, se debe utilizar este algoritmo a primeras instancias para que se pueda calcular un riesgo, de otra manera no sería posible. Además estas fueros las configuraciones utilizadas para el algoritmo.

Modelo	**Datos Utilizados**
Detection Model (por IP)	- IP convertida a número entero - Puerto de red - Vectores de ataque (8): *password_policy*, *cve_analysys*, *insecure_shared_folder*, *web_login*, *suspicious_url*, *default_credentials*, *possible_directory_traversal*, *dbms* - Pesos definidos por experto para cada vector
Active Model (por activo)	- IP convertida a número entero - Puerto de red - Mismos 8 vectores anteriores - Métricas adicionales (5): *num_incidents_active*, *num_incidents_mitigated*, *total_low_incidents*, *total_medium_incidents*, *total_critical_incidents*, *total_high_incidents* - Pesos asignados para vectores e incidentes
Preprocesamiento	- Valores nulos reemplazados por cero - Verificación de existencia de todas las columnas - Normalización posterior a 0–100 con comparación relativa de riesgo

Tabla 2: *Resumen de datos utilizados en el entrenamiento del algoritmo Isolation Forest*

Vector de Ataque	**Peso Asignado**
password_policy	14.8
cve_analysys	15.75
insecure_shared_folder	15.75
web_login	10.843
suspicious_url	8.714
default_credentials	19.687
possible_directory_traversal	10.984
dbms	23.625
num_incidents_active	2
num_incidents_mitigated	25
total_low_incidents	90
total_medium_incidents	60
total_critical_incidents	30
total_high_incidents	10

Tabla 3: *Vectores de ataque utilizados en el modelo y sus pesos asignados*

Para las métricas de validación, se implementaron validaciones comparativas, debido a que Isolation Forest es un método no supervisado, por lo que al no haber etiquetas como *Activo Riesgoso* o *Activo Normal*, estas validaciones se centro en la observación del comportamiento del modelo sobre los datos y comparando el nuevo riesgo con el

anterior, donde se puede observar de mejor manera en 6.1. Por otro lado, para obtener el valor final, separados en los criterios de Crítico/Alto/Medio/Bajo se tiene una salida dependiendo de los Umbrales percentiles 75, 85 y 95 calculados, donde dependiendo su distribución es donde entra cada activo.

Finalmente los posibles sesgos caen la selección de vectores y pesos porque son definidos manualmente por un experto, lo que introduce subjetividad. Además estos tienen una dependencia del histórico de detecciones,es decir, si la base de datos está incompleta, el modelo se entrena sobre una visión parcial del sistema que a medida que se va completando la base de datos, se tendrá una mayor precisión del riesgo de cada activo, y es por ello que con cada carga, estos se va re-entrenando para evitar lo mencionado.

4.1.4. Tecnologías utilizadas

- **Back-end:** Python + Flask + SQLAlchemy (ORM).
- **Base de datos:** PostgreSQL.
- **Front-end:** HTML, CSS y JavaScript.
- **Contenedores:** Docker para herramientas como EyeWitness.
- **Machine Learning:** Scikit-learn para Isolation Forest.

Cada componente del sistema fue diseñado para reducir la fricción técnica, mejorar el rendimiento y facilitar su integración en el entorno real de una organización de telecomunicaciones.

4.2. Arquitectura del Sistema

4.2.1. Flujo del Sistema

Dado que el usuario final es un Pentester, se ha diseñado una herramienta para automatizar su trabajo y facilitar la visualización de los resultados obtenidos de forma rápida y cómoda. Para ello, el sistema debe realizar un escaneo de red, procesar los datos recopilados, analizar y calcular los riesgos, y generar un informe detallado con los resultados de cada activo, todo centralizado en un aplicativo web.

El flujo de trabajo del sistema está diseñado para que el usuario solo necesite interactuar con los resultados, optimizando así su experiencia.

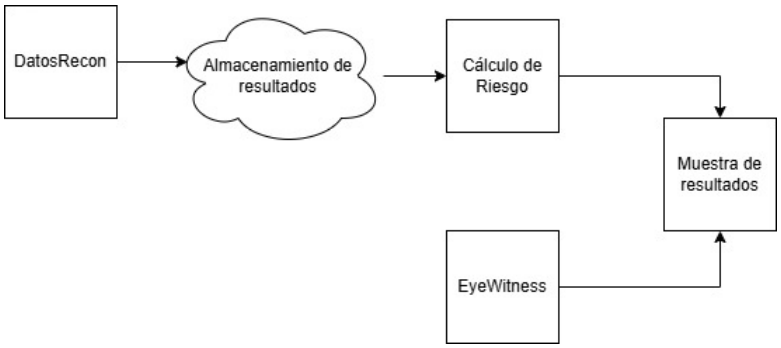


Figura 4.1: *Flujo del sistema.*

Como se muestra en la Figura 4.1, el sistema requiere únicamente la entrada de una IP o URL (también pueden ingresarse múltiples IPs, URLs o segmentos de red). A partir de esta información, se activa la herramienta *DatosRecon* y *EyeWitness* para recopilar toda la información de los activos. Esto permite escanear hasta 1000 activos sin necesidad de hacerlo manualmente uno por uno, delegando la tarea a una especie de sonda que automatiza el proceso y presenta únicamente los resultados.

Una vez generados los resultados mediante *DatosRecon*, el usuario continúa con un flujo predefinido para cargar, visualizar y analizar los datos a través del sistema.

Etapa de carga de datos: La herramienta *DatosRecon* genera un archivo en for-

mato Excel (.xlsx) que contiene la información recolectada de los activos detectados, incluyendo IPs, puertos abiertos, sistema operativo, fingerprint, entre otros atributos técnicos. Esta hoja de cálculo es posteriormente cargada al sistema mediante la interfaz web, como se muestra en la Figura 5.6.

Etapa de normalización y validación de datos: Una vez cargado el archivo, el sistema realiza un proceso de validación y normalización que incluye:

- Verificación de formato y estructura del archivo (columnas esperadas, tipos de datos, duplicados).
- Conversión y estandarización de valores (nombres de servicios o sistemas operativos, etc).
- Eliminación de registros incompletos o inconsistentes.
- Inserción en la base de datos relacional bajo los modelos definidos (por ejemplo, asociación de IPs con activos, o detecciones con vectores de ataque).

Este proceso garantiza que la información integrada en el sistema sea coherente, completa y comparable, lo que permite al usuario explorar los datos de forma efectiva y tomar decisiones basadas en una representación precisa del entorno tecnológico.

Etapa de detección y análisis de activos desconocidos: Luego de cargar y normalizar los datos, el sistema procede a realizar la detección de activos desconocidos. Esta detección se basa en una comparación automatizada entre los activos descubiertos por la herramienta *DatosRecon* y los activos ya registrados oficialmente en la base de datos de la organización. Aquellos que no tienen una correspondencia exacta en términos de IP, hostname o identificadores únicos se marcan como *no autorizados* o *desconocidos*.

Tasa de dispositivos desconocidos: A partir de esta comparación, se calcula una métrica porcentual denominada *tasa de activos desconocidos*, definida como:

$$\text{Tasa de activos desconocidos (\%)} = \frac{\text{Número de activos desconocidos}}{\text{Total de activos detectados}} \times 100$$

Este indicador permite observar la magnitud del problema de Shadow IT dentro de la organización, y su evolución en distintas campañas de escaneo.

Prioridad entre activos desconocidos: No todos los activos desconocidos representan el mismo nivel de riesgo. Aquellos que se encuentran expuestos a servicios inseguros, presentan vulnerabilidades críticas (según CVSS), o no están en ambientes controlados (por ejemplo, entornos de desarrollo sin protección), son considerados prioritarios.

Para establecer esta prioridad, se utiliza una combinación de:

- **Ponderación técnica:** nivel de criticidad del activo según puertos, fingerprint, sistemas operativos antiguos o detectores de CPE/CVE.
- **Ponderación de contexto:** ambiente donde fue detectado (producción, pruebas, DMZ), asociado a etiquetas o metadatos.
- **Modelo de riesgo relativo:** mediante Isolation Forest, el activo es comparado frente al resto, y se evalúa qué tan anómalo es su comportamiento o configuración.

Análisis por fuente de datos: Cada activo puede tener múltiples fuentes de evidencia (IP, FQDN, servicio web, puertos, fingerprint, etc). El sistema realiza una triangulación de estos elementos para verificar su existencia y clasificación. Por ejemplo, si un activo tiene IP válida pero hostname desconocido y expone servicios no autorizados, se incrementa su peso como activo no registrado.

Estas fuentes se analizan en conjunto para reforzar la certeza de que el activo efectivamente no pertenece al inventario oficial, y se agregan a un listado priorizado para ser analizado por el equipo de seguridad.

4.2.2. Flujo del usuario

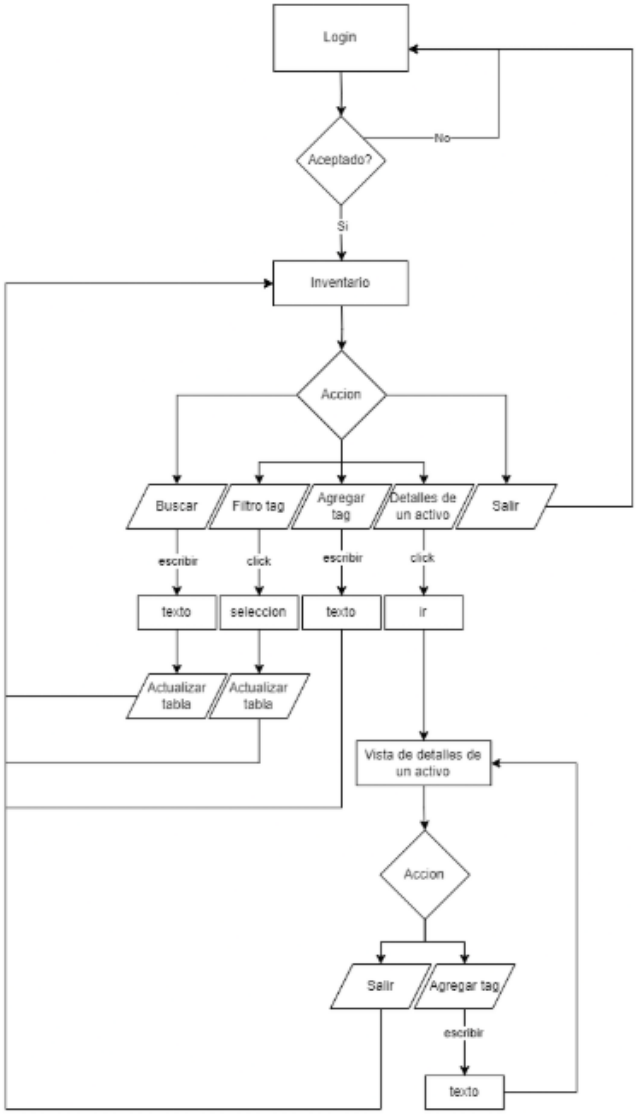


Figura 4.2: *Flujo del usuario.*

Como se muestra en la Figura 4.2, después de iniciar sesión, el usuario es redirigido al inventario de activos, donde se encuentran listados todos aquellos identificados por *DatosRecon*. Estos activos están organizados en una tabla ordenada según el riesgo previamente calculado. En esta vista, el usuario puede realizar búsquedas específicas utilizando criterios como IP, tags, URL o un nombre distintivo.

Los tags son agregados manualmente por el usuario con el propósito de facilitar la identificación y diferenciación de los activos, mejorando la experiencia de uso y la personalización del inventario. Además, cada activo dispone de una sección de detalles donde se presenta información recopilada tanto por *DatosRecon* como por *EyeWitness*. Entre los datos destacados se incluyen vectores de ataque, huellas digitales (*fingerprints*), puertos abiertos, sistema operativo y otra información relevante.

5 Resultados del Sistema

A partir de la propuesta planteada, se presenta la solución implementada, denominada *Inventarium*. Los resultados se describen inicialmente mediante capturas de pantalla del aplicativo para facilitar su análisis y comprensión. Se incluyen también ejemplos de casos reales que ilustran su funcionamiento, así como una evaluación de las funcionalidades desarrolladas y una comparativa sobre el cálculo de riesgos basado en el uso de inteligencia artificial.

Es importante destacar que, dado que este proyecto fue implementado para una compañía de Telecomunicaciones, algunos datos confidenciales, como las direcciones IP y los nombres de host, han sido anonimizados u ocultados en las visualizaciones, con el fin de preservar la privacidad organizacional y evitar la exposición de información crítica en el futuro.

Además, se ha considerado el cumplimiento de estándares y regulaciones internacionales en materia de seguridad y protección de datos, tales como el Reglamento General de Protección de Datos (GDPR) y la norma ISO/IEC 27001. En este contexto, la solución respeta los siguientes principios fundamentales:

- **Minimización de datos:** El sistema recolecta únicamente la información técnica necesaria para identificar los activos (como puertos abiertos, firmas de servicios o configuraciones), sin almacenar datos personales o sensibles innecesarios.
- **Anonimización y control de acceso:** Toda la información relacionada con acti-

vos detectados es almacenada bajo controles estrictos de acceso, y se aplican mecanismos de anonimización para campos sensibles como direcciones IP o nombres de host durante el proceso de visualización y análisis externo.

- **Consentimiento y alcance autorizado:** Las pruebas y escaneos realizados se enfocaron exclusivamente en rangos IP públicos previamente autorizados por la organización, asegurando que no se infringieran límites de auditoría ni se comprometiera infraestructura ajena.

- **Registro de auditoría:** Cada acción dentro del sistema (carga, modificación o eliminación de datos) queda registrada para fines de trazabilidad y cumplimiento de auditorías internas o externas, tal como lo recomienda ISO 27001.

Estos elementos aseguran que el uso de *Inventarium* no solo se alinea con los objetivos de detección de tecnología en las sombras, sino que lo hace dentro de un marco de cumplimiento normativo y ético, aspecto crucial para organizaciones que manejan infraestructura crítica o datos sensibles.

Durante el proceso se analizaron tres fuentes de datos distintas:

- Activos nuevos obtenidos mediante **DatosRecon**.

- Capturas visuales de servicios y aplicativos web **EyeWitness**.

- Información pasiva e histórica obtenida mediante **DatosRecon**.

Cada una de estas fuentes permitió complementar la detección, aumentar la certeza sobre el estado real de los activos y mejorar la cobertura total.

5.0.1. Visualización de la arquitectura del sistema

- **Vista General de Activos de Información.** Como se describió anteriormente, en la vista general de Inventarium (Ver 5.1), se han desarrollado varias funcionalidades clave, organizadas en dos secciones principales:

1. **Sección superior:** Diseñada para el filtrado y la búsqueda eficiente de activos. Incluye las siguientes características:
   - **Buscar y Tags:** Permite agregar etiquetas a varios activos de forma simultánea, facilitando la navegación y organización dentro del sistema.
   - **Filtrado por nivel de riesgo:** Los activos se clasifican según su nivel de riesgo en las categorías Crítico, Alto, Mediano y Bajo, lo que permite visualizar de manera precisa el estado de seguridad de cada activo.
   - **Herramientas especializadas:** Se incluye una sección adicional con herramientas diseñadas específicamente para el usuario. Esta funcionalidad se detalla más adelante en el documento.
2. **Tabla inferior:** Contiene la lista completa de activos detectados. En esta sección, los activos pueden ser ordenados según su nivel de riesgo o los tags asignados. Además, se presenta información detallada de cada activo, como vectores de ataque identificados, dirección IP, puertos abiertos, sistema operativo, dirección IP y nombre de host (si está disponible). Finalmente, en el extremo derecho de la tabla, se encuentra un enlace que conduce a la sección de detalles individuales de cada activo.

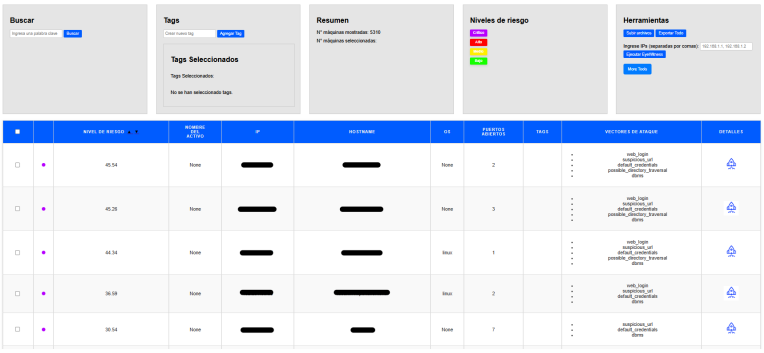


Figura 5.1: *Vista general de Inventarium.*

Continuando con el apartado de Herramientas, donde se incluyeron diferentes funcionalidades descritas a continuación.

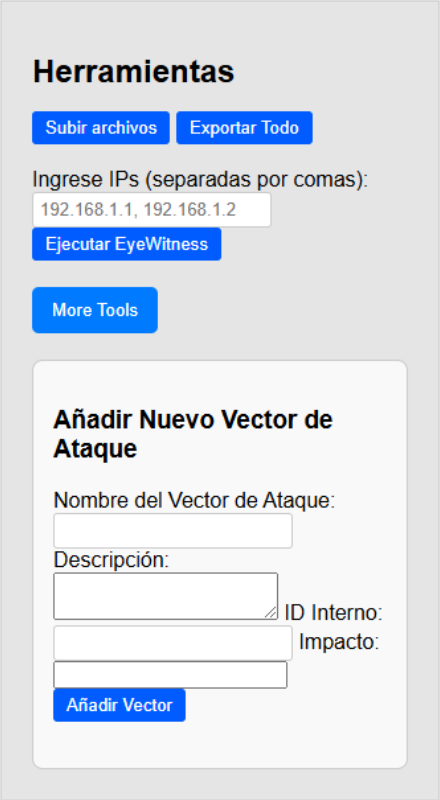


Figura 5.2: *Vista de herramientas en Inventarium.*

En la Figura 5.2 se puede observar tres acciones principales:

1. **Subir Archivos:** Esta funcionalidad permite cargar datos al sistema mediante archivos en formato XLSX o JSON. Los datos cargados son validados y el sistema calcula automáticamente el nivel de riesgo para cada una de las detecciones, como se ilustra en la Figura 5.6
2. **Ejecución de EyeWitness:** Esta herramienta se integra directamente en la aplicación y permite al usuario introducir un listado de direcciones IP separadas por comas. Al ejecutar el proceso, EyeWitness captura imágenes relacionadas con las IPs ingresadas y las guarda dentro de la misma aplicación. Posteriormente, estas imágenes se pueden visualizar en la sección de

detalles de cada activo.

3. **Añadir Nuevos Vectores de Ataque:** Dado que la ciberseguridad es un campo dinámico y los vectores de ataque evolucionan constantemente, se ha creado un apartado que permite añadir nuevos vectores de ataque al sistema. Para ello, se debe completar un formulario con la siguiente información:

   - **Nombre:** Identificación del vector de ataque.
   - **Descripción:** Breve explicación del vector de ataque.
   - **ID interno:** Un identificador único asociado a la organización.
   - **Impacto:** Valor numérico que será utilizado en el cálculo del riesgo asociado al vector.

De esta manera, se puede entrar a una vista detallada de cada uno de los activos para que el usuario pueda ver especificaciones técnicas del activo, esto se puede ver en la siguiente Figura 5.3:

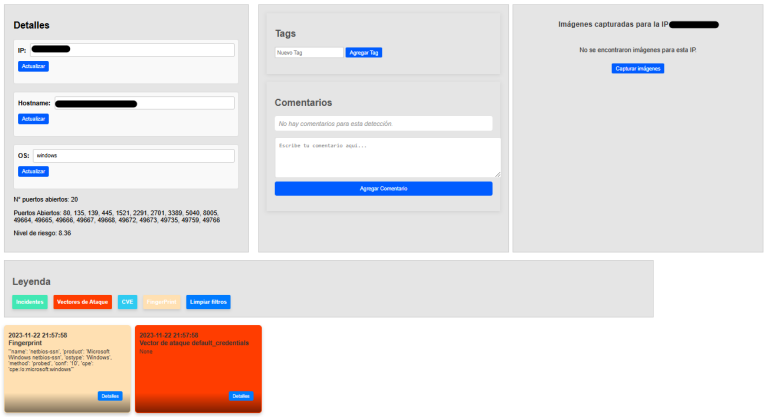


Figura 5.3: *Vista de detalles del activo.*

Como se puede observar, en esta vista se presenta información detallada del activo, incluyendo:

- Sistema Operativo (SO)
- Número y lista de puertos abiertos identificados

- Tags asociados: Etiquetas que permiten clasificar y organizar los activos.
- Sección de comentarios: Espacio donde los usuarios pueden agregar información adicional relevante sobre el activo.
- EyeWitness: Herramienta que permite visualizar imágenes capturadas directamente desde la aplicación.
- Leyendas: Información detallada de lo recopilado por DatosRecon, como incidentes, vectores de ataque, CVE y huellas digitales (Fingerprints).

Además, las imágenes capturadas por EyeWitness se visualizan como se muestra en la Figura 5.4

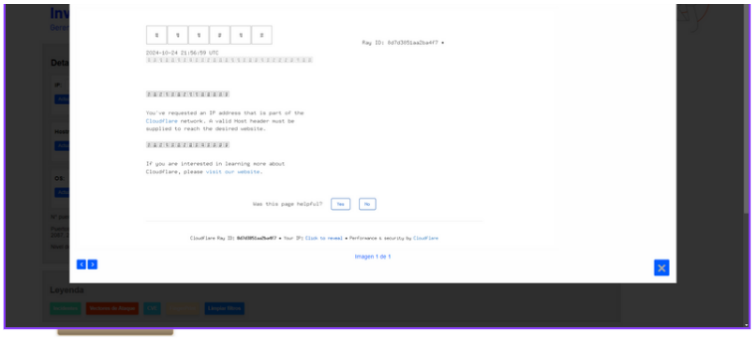


Figura 5.4: *Vista de imágenes capturadas mediante EyeWitness.*

Donde el usuario podrá ver de antemano, sin la necesidad de dirigirse al activo como tal, como este luce, lo cual le podrá entregar información valiosa la cual puede ser reportada como posibles vulnerabilidades.

Un aspecto crucial en la discusión es el *cálculo del nivel de riesgo*. Este valor, que oscila entre 0 y 100, es calculado considerando los datos obtenidos por DatosRecon y los vectores de ataque identificados. El nivel de riesgo es evaluado comparativamente entre los activos y se calcula siguiendo lo descrito en la Subsección 4.1.3. En la Figura 5.5 se presenta una visualización clara y detallada de este indicador, el cual permite al usuario identificar de manera cuantitativa qué tan vulnerable es un activo en relación con el resto.

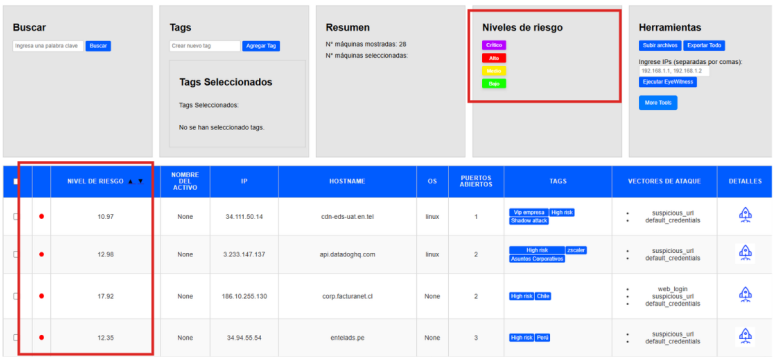


Figura 5.5: *Vista Riesgo calculado.*

- Subida de Archivos: Para cargar los datos al sistema como se describió en 4.2.1, estos se hacen a partir de archivos ya sean históricos de la compañía de Telecomunicaciones o nuevos generados por la herramienta DatosRecon, visualizandose de la siguiente manera:

Figura 5.6: *Vista Riesgo calculado.*

5.0.2. Búsqueda Comparativa de Tecnología en las Sombras

Para entender a qué se refiere el nivel de riesgo en las categorías, se presenta una tipificación basada en los datos reales recopilados. A continuación se detallan ejemplos representativos en cada categoría de riesgo:

1. **Riesgo Crítico:** Se identificaron **2 activos** con riesgo crítico. Uno de ellos corresponde a un servidor web en producción con una versión vulnerable de `Apache 2.2.15`, accesible desde Internet y sin autenticación. El otro es una base de datos

`PostgreSQL` expuesta en un entorno que no figura en los registros oficiales, con puerto por defecto y sin cifrado. Ambos presentan CVEs críticas activas.

2. **Riesgo Alto: 7 activos** fueron clasificados con riesgo alto. Entre ellos se incluyen aplicaciones mal configuradas expuestas en puertos no estándar y dispositivos IoT sin credenciales seguras. Uno de ellos corresponde a una interfaz de administración de red sin autenticación, expuesta públicamente.

3. **Riesgo Medio: 18 activos** fueron categorizados como riesgo medio, en su mayoría por configuraciones por defecto, versiones desactualizadas o servicios auxiliares sin protección (por ejemplo, servidores de pruebas expuestos con puertos estándar).

4. **Riesgo Bajo: 144 activos** fueron clasificados como de riesgo bajo. La mayoría corresponden a entornos de desarrollo olvidados o dispositivos sin acceso crítico, pero aún visibles desde el exterior.

Estas cifras reflejan que del total de activos desconocidos detectados, el 1.2 % se considera de riesgo crítico, el 4.1 % de riesgo alto, el 10.5 % de riesgo medio y el 84.2 % de riesgo bajo. La mayoría de los activos desconocidos, aunque presentan menor impacto individual, contribuyen significativamente a la superficie de ataque.

Atributos comunes de activos desconocidos: Los activos no autorizados suelen compartir ciertas características:

- No figuran en bases de datos oficiales ni tienen tags asignados.
- Exponen servicios inseguros (por ejemplo: `FTP`, `Telnet`, `MySQL` sin cifrado).
- En su mayoría pertenecen a ambientes de desarrollo (`qa`, `test`, `dev`).
- No cuentan con registros históricos en herramientas internas de monitoreo.

Comparación con estimaciones previas: Antes de contar con la solución *Inventarium*, se estimaba una cobertura del 95 % de los activos expuestos. Sin embargo, los

resultados demostraron que aproximadamente un **20 % de los activos** eran desconocidos, evidenciando una brecha significativa en el control de infraestructura.

Reacción del Operador de Telecomunicaciones: Los resultados fueron validados por el equipo de ciberseguridad de la organización. Se revisaron los hallazgos y se confirmó que varios activos detectados no se encontraban en los inventarios internos. Esta verificación externa fue clave para la aceptación de la herramienta como soporte para futuras campañas de control y limpieza de infraestructura.

Transición de activo desconocido a conocido: Una vez que el sistema identifica un activo como desconocido, se genera una alerta y se clasifica en el sistema como "no autorizado". Tras la validación del equipo, este activo puede ser marcado como reconocido y se integra al inventario oficial. Esta transición provoca una reevaluación del riesgo del activo, considerando ahora su entorno, historial y medidas de seguridad aplicadas, pudiendo así reducir su puntaje de riesgo en futuras evaluaciones.

5.0.3. Resultados de la Búsqueda Comparativa

Durante la búsqueda se trabajó con tres tamaños de fuentes de datos:

- Primer intento: conjunto pequeño de 100 activos, con un 27 % de detección de Shadow IT.
- Segundo intento: conjunto de 300 activos, con una reducción a un 21 %.
- Tercer intento (completo): conjunto de 795 activos, con una detección consolidada del 20 %.

Este comportamiento demuestra una tendencia decreciente en la aparición de activos desconocidos a medida que se expande la cobertura del inventario, lo que valida la efectividad de la herramienta en reducir brechas.

Aspecto	**Resultado**
Tiempo de ejecución	6 horas y 30 minutos, revisando 2 activos por minuto.
Tecnología de las sombras descubiertas	171 nuevos activos descubiertos en inventario (20 %).
Ambientes no productivos descubiertos	10 activos con dominios de desarrollo (*qa, test, dev, -pp*).
Malas configuraciones encontradas	40 activos con malas configuraciones (*ftp, ssh, telnet, mysql, postgres, oracle*).
Vulnerabilidades conocidas (CVE)	9 activos con vulnerabilidades críticas por fácil explotación.

Tabla 4: *Resumen de los resultados de la búsqueda comparativa de tecnología en las sombras.*

- Distribución de activos descubiertos por categoría de riesgo.

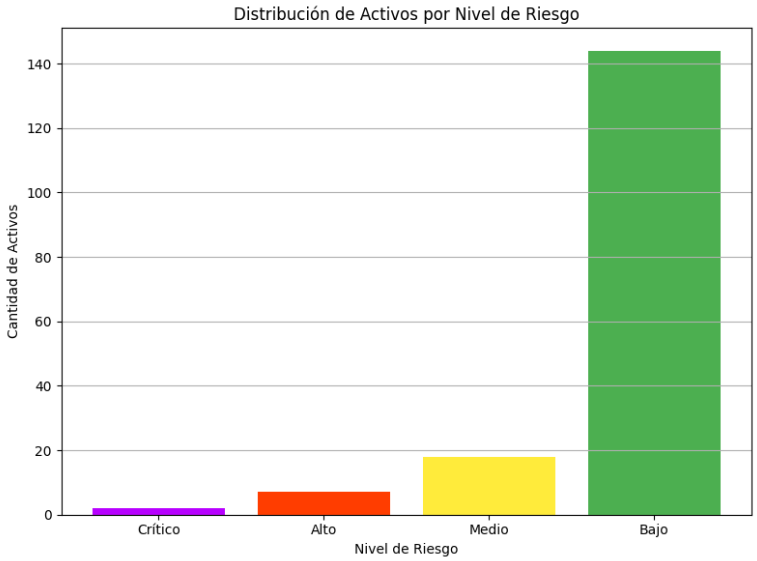


Figura 5.7: *Distribución de activos por nivel de riesgo.*

- Comparación entre activos conocidos y desconocidos.

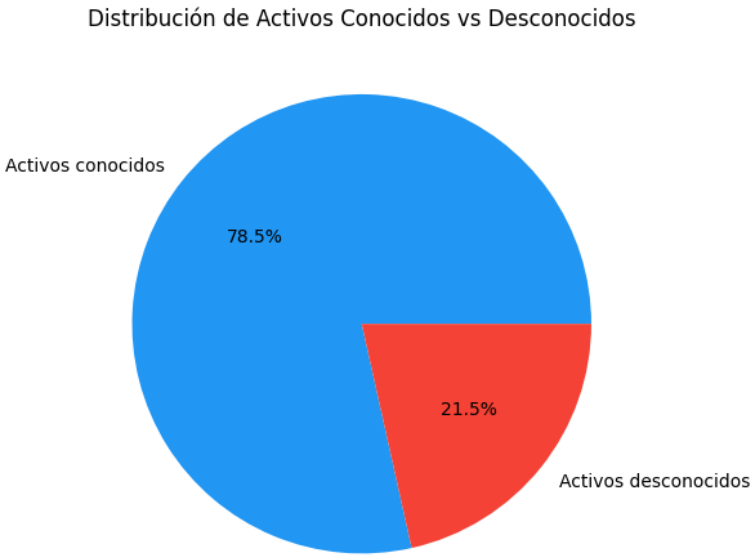


Figura 5.8: *Comparación de activos conocidos vs desconocidos.*

- Porcentaje de activos por tipo de error (ambiente no productivo, mala configuración, CVE).

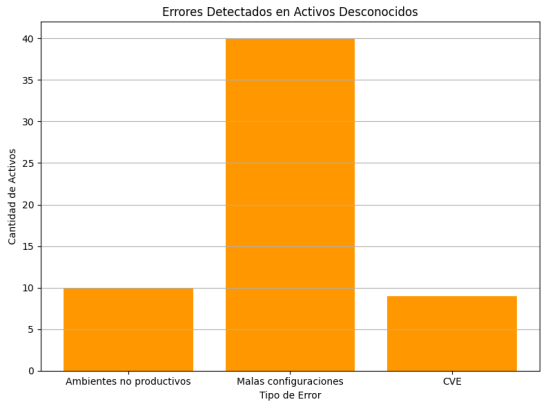


Figura 5.9: *Errores detectados en activos desconocidos.*

6 Análisis de Resultados

La búsqueda de tecnología en las sombras en una empresa de telecomunicaciones se realizó con el objetivo de identificar activos expuestos a internet que no estuvieran inventariados, es decir, aquellos activos que podrían representar un riesgo debido a su exposición no controlada. La investigación incluyó 795 activos y se centró en descubrir puertos filtrados, servicios activos, bases de datos expuestas y ambientes no productivos. A continuación, se presenta un análisis detallado de los resultados obtenidos:

1. **Tiempo de Ejecución** El tiempo total de ejecución fue de 6 horas y 30 minutos para revisar 795 activos, lo que equivale a una media de 2 activos revisados por minuto. Este tiempo de ejecución puede considerarse moderado, ya que, aunque no es excesivamente rápido, se ajusta a las necesidades de un proceso exhaustivo que involucra múltiples pruebas en cada activo. La moderación en la velocidad es necesaria para evitar bloqueos por parte de sistemas de protección como los firewalls o los WAFs (Web Application Firewalls), lo cual podría resultar en falsos negativos. Es importante destacar que, si un activo ya estaba registrado en el sistema, no se consideró como "tecnología en las sombras", sino que se actualizó su información. Esto ayuda a mantener un inventario más actualizado y evitar problemas relacionados con versiones desactualizadas o configuraciones incorrectas.

2. **Tecnología de las Sombras Descubiertas** Se descubrieron 171 nuevos activos que no estaban registrados en el inventario. Este hallazgo representa una parte significativa de los activos descubiertos, y es crucial para reducir la brecha en la cobertura de los activos de la organización. La meta de esta búsqueda es que, con el tiempo, este número disminuya hasta llegar a cero, lo que implicaría que todos los activos están adecuadamente registrados y gestionados.

   Además, dentro de los activos encontrados, se identificaron 10 que corresponden a dominios relacionados con ambientes no productivos, como los de pruebas o

desarrollo. Estos activos, pertenecientes a dominios como *qa, test, dev, y -pp*, no deberían estar expuestos a internet. La exposición de estos activos representa un riesgo potencial, ya que pueden no tener la misma seguridad que los ambientes productivos. Este hallazgo subraya la necesidad de revisar y ajustar las políticas de seguridad para estos entornos.

3. **Malas Configuraciones en los Activos** Uno de los hallazgos más significativos fue la identificación de 40 activos que presentaban malas configuraciones. Estas malas configuraciones, que incluyen protocolos inseguros como *ftp, ssh, telnet, mysql, postgres y oracle*, constituyen una categoría crítica de "tecnología en las sombras". Los protocolos mencionados, si no están configurados adecuadamente o si son accesibles desde internet, pueden ser explotados fácilmente por atacantes, lo que pone en riesgo la seguridad de los sistemas. Esta categoría de activos con malas configuraciones subraya la importancia de realizar auditorías de configuración y de aplicar buenas prácticas de seguridad en la configuración de servicios.

4. **Vulnerabilidades Conocidas (CVE)** Críticamente, se descubrieron 9 activos con vulnerabilidades conocidas (CVE) que representan un riesgo crítico debido a su fácil explotación. Las vulnerabilidades de este tipo son especialmente peligrosas porque ya han sido identificadas y documentadas por la comunidad de seguridad, lo que facilita su explotación si no se toman medidas correctivas. Estos activos deben ser priorizados para su corrección inmediata, ya que pueden ser utilizados como puntos de entrada para atacantes malintencionados. La existencia de estas vulnerabilidades resalta la importancia de mantener un proceso continuo de actualización y parcheo de los sistemas.

5. **Calculo de riesgo**

   Finalmente, al analizar el calculo de riesgo, se hace la comparación como antes la organización lo hacia, este valor antiguamente se calculaba mediante un experto

que otorgo valores dependiendo del vector de ataque conocido, cierto impacto y así con esas dos opiniones se obtenía el riesgo, por lo que con esta nueva metodología, se estandariza de manera tangible una mejor manera para obtener este valor, esto se puede apreciar en 6.1, donde muchos de estos, el riesgo presentado era mucho mayor del calculado, logrando una mejor apreciación del riesgo que contiene cada activo.

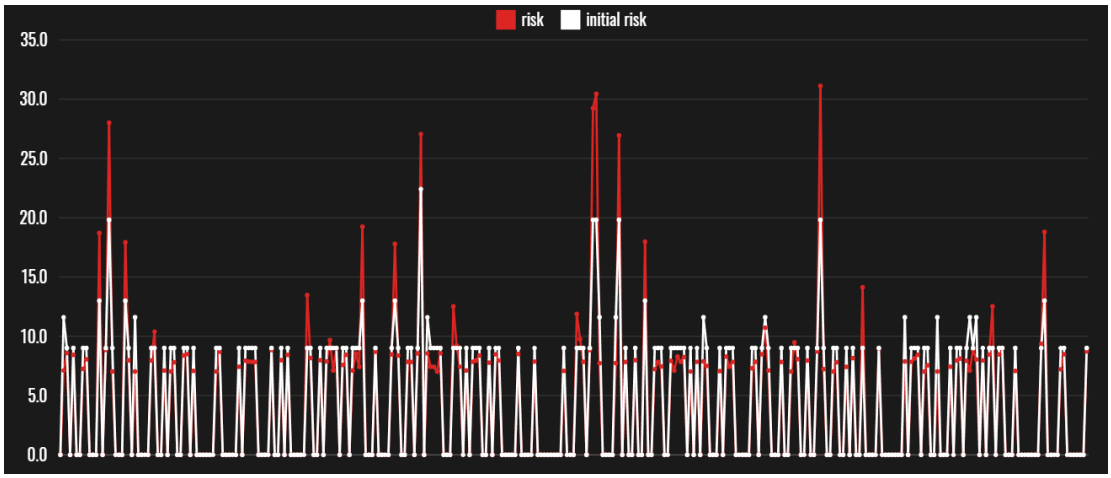


Figura 6.1: *Impacto del nuevo modelo de calculo de riesgo.*

7 Conclusiones

El análisis de resultados del sistema *Inventarium* evidencia la existencia de una brecha relevante en la visibilidad de activos tecnológicos dentro de una organización de telecomunicaciones. De los 795 activos revisados, se identificaron 171 como activos no registrados, lo que corresponde a un **21.5 % del total**. Este porcentaje es considerable, dado que a medida que se va utilizando esta herramienta, se irá completando la superficie de activos que utiliza la organización por lo que cada vez debería ir disminuyendo más este valor pero dado la importancia de datos que maneja este tipo de organización, este valor debería ser menor a 1 %.

Inventarium demostró ser una herramienta eficaz para detectar activos no autorizados, ya que logró identificar estos activos mediante múltiples fuentes de información (Nmap, EyeWitness y DatosRecon), integrando visualización, análisis y priorización del riesgo. Este enfoque sistemático permite no solo descubrir tecnologías en las sombras, sino también categorizar su nivel de riesgo y facilitar acciones correctivas.

El porcentaje detectado puede considerarse razonable dentro del contexto nacional e internacional para operadores de telecomunicaciones, especialmente considerando la complejidad de su infraestructura y la proliferación de servicios internos. Sin embargo, también representa un llamado a la acción para fortalecer las políticas de gobernanza de activos y control de inventarios.

En consecuencia, la evidencia obtenida a través de este proyecto podría llevar a reconsiderar el enfoque del trabajo, pasando de un diagnóstico a una propuesta más amplia de **plataforma de gobernanza activa de activos tecnológicos no autorizados**, haciendo visible una contribución que va más allá del descubrimiento, hacia la consolidación sistemática de la gestión de riesgos.

Principales hallazgos y acciones necesarias:

- **Reducción de activos no inventariados:** Con un 21.5 % de activos no autorizados detectados, se vuelve crucial mantener la ejecución periódica de *Inventarium*

y alimentar un inventario maestro continuamente actualizado.

- **Ambientes no productivos expuestos:** Un total de 10 activos operaban en ambientes de desarrollo (con dominios `qa`, `test`, `dev`, `-pp`) visibles desde internet. Esto requiere políticas más estrictas de segmentación de redes y control de accesos.

- **Malas configuraciones:** 40 activos presentaron protocolos o configuraciones inseguras. Este grupo debe priorizarse mediante auditorías automáticas recurrentes y buenas prácticas de hardening.

- **Vulnerabilidades críticas (CVE):** 9 activos mostraron vulnerabilidades con exploits conocidos, lo que requiere aplicar medidas inmediatas de parcheo y monitoreo reforzado.

- **Cálculo de riesgo estandarizado:** Se comprobó que el uso de inteligencia artificial con Isolation Forest mejora la asignación objetiva del riesgo, permitiendo priorizar de forma efectiva y detectar falsos negativos en métodos anteriores.

7.1. Comparación frente al mercado

Ahora bien, ¿Como se compara nuestra solución frente a otras existente en el mercado?. Para ello, se hace la comparación en la Tabla 5, donde podremos apreciar el enfoque principal, tecnologías claves y mayores virtudes de cada una de estas herramientas. Las soluciones comerciales analizadas destacan por su capacidad de descubrimiento en entornos específicos (por ejemplo, Censys y CybelAngel en entornos expuestos a internet, o Tenable en OT). Sin embargo, ninguna de ellas ofrece una solución integral personalizada que combine detección activa, cálculo de riesgo basado en inteligencia artificial, análisis de configuraciones inseguras, y gestión visual de activos en redes complejas como las de un operador de telecomunicaciones.

En contraste, Inventarium integra componentes técnicos adaptados al contexto real del entorno, utilizando un enfoque bottom-up que permite a las organizaciones detectar, clasificar y priorizar riesgos con flexibilidad. Su mayor virtud radica en la personalización técnica y la capacidad de adaptación al contexto organizacional sin depender de licencias comerciales ni infraestructura externa.

8 Discusión y Trabajo Futuro

La búsqueda comparativa realizada con *Inventarium* no solo aportó visibilidad sobre activos expuestos, sino también permitió inferir patrones que abren nuevas posibilidades de análisis. Al estudiar las características de los activos no autorizados (por ejemplo, puertos abiertos, encabezados de dominio, configuraciones por defecto), es posible aplicar técnicas de análisis de clústeres que agrupen los activos según su origen, configuración o comportamiento. Esto permitiría visualizar **zonas de riesgo homogéneas** dentro de la red corporativa.

Una línea prometedora para trabajo futuro consiste en triangular los datos provenientes de las distintas fuentes (escaneo activo, registros históricos, captura visual) para consolidar un modelo predictivo capaz de estimar con alta probabilidad si un activo nuevo descubierto corresponde a un riesgo alto o crítico. Esta triangulación permitiría pasar de una fase reactiva a una proactiva en la gestión del Shadow IT.

Otra reflexión importante es sobre la **localización física de los activos no autorizados**. Si bien estos activos se detectan a nivel lógico (a través de la red), en la práctica es posible mapear su ubicación física con suficiente precisión si se integra Inventarium con sistemas de gestión de red interna o con herramientas de geolocalización por IP, lo cual facilitaría su desmantelamiento o regularización.

Por el lado del modelo de entrenamiento, como se esta en una etapa temprana, y no tenemos los datos etiquetados para utilizar otro modelo, luego de esta solución no se tendrá esta problemática por lo que es buena manera para definir el riesgo de cada activo, utilizar otro modelo de entrenamiento con datos ya etiquetados.

Finalmente, desde una vista más comercial, se podría evaluar la opción de llevar esta solución a un mercado como lo puede ser la banca, salud o retail, ya que, como se presento en el escrito, esto brilla por su simplicidad y gran capacidad de adaptabilidad en cada detalle y siempre fue desarrollado para una mejoría futura.

9 Anexos

Solución	Enfoque principal	Tecnologías clave	Mayor virtud
Inventarium (propuesta)	Detección activa de activos no inventariados en redes expuestas	Nmap, EyeWitness, CVSS, Isolation Forest (IA)	Personalización completa y adaptabilidad al entorno de telecomunicaciones
CybelAngel ADM	Monitorización continua de Shadow IT externo	Escaneo contextual, analistas humanos, IA	Alta precisión en descubrimiento de activos expuestos sin falsos positivos
SunVizion Network Inventory	Gestión de infraestructura de red (física/lógica)	Mapeo gráfico, documentación jerárquica	Visualización completa y estructurada de toda la red física
ServiceNow CMDB + Vulnerability Response	Gestión de activos y remediación de vulnerabilidades	CMDB, correlación con escáneres, flujos automatizados	Integración robusta con flujos empresariales ITSM y visibilidad centralizada
OpenText Aviator IoT	Localización y gestión de activos IoT	Telemetría, IA, identidad de dispositivos	Gestión eficiente y remota de activos físicos dispersos

Continúa en la siguiente página

Tabla 5 – continuación de la página anterior

Solución	**Enfoque principal**	**Tecnologías clave**	**Mayor virtud**
Tenable OT	Gestión de exposición en entornos TI/OT/IoT	Inventario pasivo, análisis de vulnerabilidades	Visibilidad unificada entre tecnología operativa y de información
Censys EASM	Gestión externa de superficie de ataque	Descubrimiento continuo en Internet, análisis de riesgo	Detección rápida de exposición pública no autorizada

Tabla 5: *Comparación de soluciones para detección de activos no inventariados y Shadow IT*

Referencias

[1] J.-P. van Acken, J. F. Gadellaa, S. Jansen y K. Labunets, «Poster: The Unknown Unknown: Cybersecurity Threats of Shadow IT in Higher Education,» en *Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security (CCS '23)*, New York, NY, USA: Association for Computing Machinery, 2023, págs. 3633-3635. DOI: 10.1145/3576915.3624395. dirección: https://doi.org/10.1145/3576915.3624395.

[2] Kaspersky, «El uso de Shadow IT por parte de empleados expone a las empresas a ciberincidentes,» *Kaspersky*, 2025, Accedido: 2025-01-21. dirección: https://www.kaspersky.es/about/press-releases/el-uso-de-shadow-it-por-parte-de-empleados-expone-a-las-empresas-a-ciberincidentes?.com.

[3] D. Fürstenau y H. Rothe, «Shadow IT Systems: Discerning the Good and the Evil,» en *Proceedings of the 22nd European Conference on Information Systems (ECIS)*, Tel Aviv, Israel: AIS Electronic Library, 2014.

[4] S. Zimmermann, C. Rentrop y C. Felden, «Governing Identified Shadow IT by Allocating IT Task Responsibilities,» en *Multikonferenz Wirtschaftsinformatik (MKWI)*, Universität Ilmenau, Ilmenau, Germany, 2016.

[5] O. Security, «Riesgo latente: el impacto de Shadow IT en las empresas,» *OpenSecurity*, 2025, Accedido: 2025-01-21. dirección: https://www.opensecurity.es/riesgo-latente-shadow-it/.

[6] M. Silic y A. Back, «Shadow IT–A View from Behind the Curtain,» *Computers & Security*, vol. 45, págs. 274-283, 2014. DOI: 10.1016/j.cose.2014.06.007. dirección: https://doi.org/10.1016/j.cose.2014.06.007.

[7] C. Días, «El 77 % de los usuarios de IA generativa introduce esta tecnología en su trabajo sin supervisión de la empresa,» *Cinco Días*, 2025, Accedido: 2025-01-21. dirección: https://cincodias.elpais.com/companias/2025-01-

21/el-77-de-los-usuarios-de-ia-generativa-introduce-esta-tecnologia-en-su-trabajo-sin-supervision-de-la-empresa.html?.com.

[8] J. Chejfec, *Résultats de l'enquête sur le phénomène du « Shadow IT »*, Online survey, Available at: http://chejfec.com/2012/12/18/resultats-complets-delenquete-shadow-it/ [Accessed: March 2014], 2012.

[9] T. Australian, «Crackdown on remote workers' 'shadow IT',» *The Australian*, 2025, Accedido: 2025-01-21. dirección: https://n9.cl/pcr3x.

[10] S. Haag y A. Eckhardt, «Shadow IT,» *Business & Information Systems Engineering*, vol. 59, n.º 6, págs. 469-473, 2017. DOI: 10.1007/s12599-017-0497-x. dirección: https://doi.org/10.1007/s12599-017-0497-x.

[11] R. Siege. «EyeWitness.» Último acceso: 13 de enero de 2025. (2025), dirección: https://github.com/RedSiege/EyeWitness.

[12] National Institute of Standards and Technology (NIST), *National Vulnerability Database: CVSS Calculadora*, https://nvd.nist.gov/vuln-metrics/cvss, Accedido: 2025-01-22, 2024.

[13] F. T. Liu, K. M. Ting y Z.-H. Zhou, «Isolation Forest,» *Proceedings of the 2008 IEEE International Conference on Data Mining (ICDM)*, págs. 413-422, 2008. DOI: 10.1109/ICDM.2008.17. dirección: https://doi.org/10.1109/ICDM.2008.17.