Ciberseguridad aplicada

un enfoque integral de riesgos, gobernanza y análisis forense digital

Luis Eduardo Carrizo García
Diana Carolina Decimavilla Alarcón
José Ottón Pinela Tigua
Luis Arturo Caisaguano Caisaguano
Ivette Auxiliadora Mateo Washbrum
Zoila Amada Pineda Calle

Ciberseguridad aplicada

un enfoque integral de riesgos, gobernanza y análisis forense digital

Autores
Luis Eduardo Carrizo García
Diana Carolina Decimavilla Alarcón
José Ottón Pinela Tigua
Luis Arturo Caisaguano Caisaguano
Ivette Auxiliadora Mateo Washbrum
Zoila Amada Pineda Calle

Dirección Editorial: PhD. Jorge Luis León-González
Diseño de portada y edición: DI. Yunisley Bruno-Díaz

ISBN: 978-1-968794-45-3

DOI: https://doi.org/10.64092/MXCH1610

La evaluación científica y metodológica de la obra se realizó a partir del método de Revisión por Pares Abierta (Open Peer Review).

SOPHIA EDITIONS
8404 N Rome Ave, Tampa,
Florida, USA
Email: contact@sophiaeditions.com
Phone: +1 (813) 699-2557
https://sophiaeditions.com/

Ciberseguridad aplicada

un enfoque integral de riesgos, gobernanza y análisis forense digital

Autores
Luis Eduardo Carrizo García
Diana Carolina Decimavilla Alarcón
José Ottón Pinela Tigua
Luis Arturo Caisaguano Caisaguano
Ivette Auxiliadora Mateo Washbrum
Zoila Amada Pineda Calle

COMITÉ EDITORIAL

CONTENIDO

CONTENIDO

PREFACIO

En la actualidad, la información se ha consolidado como uno de los activos más valiosos en la sociedad contemporánea. Cada interacción digital, cada proceso organizacional y cada transacción electrónica generan datos que sustentan el funcionamiento de la economía, la gobernanza y la vida cotidiana. Sin embargo, este avance tecnológico ha traído consigo un incremento significativo en la exposición a riesgos y amenazas en el ciberespacio. En este escenario, la ciberseguridad emerge como un elemento fundamental para garantizar la continuidad operativa, la protección de los activos digitales y la confianza en los sistemas de información.

La ciberseguridad ha evolucionado desde un enfoque meramente técnico hacia una disciplina integral que abarca dimensiones estratégicas, organizacionales, legales y humanas. Los ciberataques han dejado de ser incidentes aislados para convertirse en fenómenos complejos, caracterizados por su sofisticación, persistencia y capacidad de adaptación. En consecuencia, se hace evidente la necesidad de comprender la ciberseguridad como un proceso dinámico, en constante transformación, que requiere una visión global e interdisciplinaria.

En este contexto, la presente obra titulada “Ciberseguridad aplicada: un enfoque integral de riesgos, gobernanza y análisis forense digital” se plantea como una contribución académica orientada a integrar los principales fundamentos, enfoques y aplicaciones de la ciberseguridad moderna. Su propósito es ofrecer una visión estructurada que permita comprender la interrelación entre los distintos componentes que conforman este campo, abordando desde los conceptos básicos hasta los aspectos más avanzados relacionados con la defensa, el ataque y la investigación forense.

A lo largo del desarrollo del libro, se adopta un enfoque sistemático que permite analizar la ciberseguridad

desde diferentes perspectivas complementarias. En primer lugar, se abordan los fundamentos, riesgos y amenazas que caracterizan el entorno digital, proporcionando al lector una base conceptual sólida para entender la naturaleza de los ciberataques y su evolución a lo largo del tiempo. Este análisis permite identificar las principales vulnerabilidades y comprender los mecanismos mediante los cuales los sistemas pueden ser comprometidos.

Posteriormente, se profundiza en el ámbito de la gobernanza y la gestión del riesgo en seguridad de la información, destacando la importancia de los marcos normativos, estándares internacionales y modelos de gestión que permiten a las organizaciones estructurar sus políticas de seguridad de manera eficiente. En este sentido, la obra enfatiza la necesidad de alinear la ciberseguridad con los objetivos estratégicos organizacionales, promoviendo una cultura de seguridad basada en la prevención, el control y la mejora continua.

El estudio de los enfoques ofensivos y defensivos constituye otro de los ejes centrales de la obra. En este apartado se analizan las estrategias utilizadas tanto por atacantes como por defensores, destacando el papel del hacking ético, las pruebas de penetración y las operaciones de equipos especializados como Red Team y Blue Team. Asimismo, se examina la ingeniería social como un vector de ataque crítico, evidenciando que el factor humano continúa siendo uno de los elementos más vulnerables dentro de los sistemas de información.

El análisis forense digital, abordado en el capítulo final, representa una de las áreas más relevantes dentro de la ciberseguridad contemporánea. En un entorno donde los incidentes dejan rastros digitales, la capacidad de identificar, preservar, analizar y presentar evidencia se convierte en un elemento esencial para la investigación y resolución de delitos informáticos. Este enfoque no solo permite reconstruir eventos y determinar responsabilidades, sino también fortalecer las estrategias de defensa mediante el aprendizaje derivado de incidentes previos.

Uno de los aspectos distintivos de esta obra radica en su carácter aplicado, el cual busca integrar la teoría con la práctica mediante un enfoque orientado a la resolución de problemas

reales. Esta perspectiva resulta especialmente relevante en un contexto donde la complejidad de las amenazas exige profesionales capaces de interpretar escenarios dinámicos y adoptar decisiones informadas. En este sentido, el contenido ha sido estructurado de manera que facilite la comprensión progresiva de los conceptos, permitiendo al lector desarrollar una visión integral de la ciberseguridad.

El público objetivo de este libro incluye estudiantes, profesionales e investigadores del área de tecnologías de la información, así como cualquier persona interesada en comprender los fundamentos y aplicaciones de la ciberseguridad. Si bien se presentan conceptos técnicos, la obra ha sido diseñada para ser accesible, promoviendo un aprendizaje gradual que no requiere necesariamente conocimientos especializados previos, pero sí una disposición al análisis crítico y a la reflexión.

En un entorno digital caracterizado por la constante evolución de las amenazas, la ciberseguridad se posiciona como un componente estratégico para el desarrollo sostenible de la sociedad. La capacidad de anticipar riesgos, implementar controles efectivos, responder a incidentes y aprender de ellos constituye un factor determinante para garantizar la resiliencia organizacional y la protección de la información.

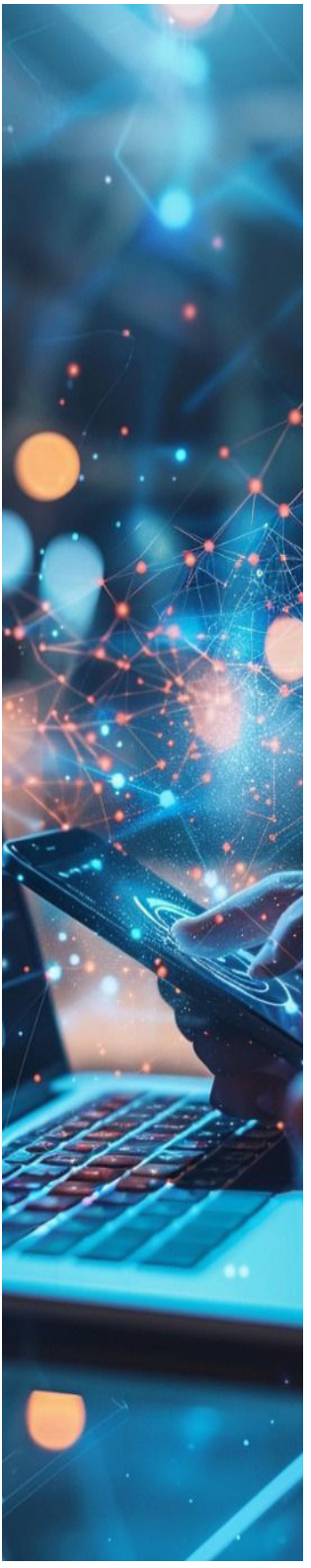
No obstante, es importante reconocer que la ciberseguridad no es un estado estático, sino un proceso continuo que requiere actualización permanente, adaptación a nuevos escenarios y fortalecimiento constante de capacidades. En este sentido, la presente obra no pretende ofrecer soluciones definitivas, sino proporcionar las herramientas conceptuales y metodológicas necesarias para comprender y enfrentar los desafíos del entorno digital.

Finalmente, se destaca que la ciberseguridad no debe ser concebida únicamente como una responsabilidad técnica, sino como un compromiso compartido que involucra a todos los actores que interactúan en el ecosistema digital. La protección de la información, la integridad de los sistemas y la confianza en las tecnologías dependen, en gran medida, de la responsabilidad, la ética y la conciencia de los usuarios.

En consecuencia, esta obra se presenta como una invitación a profundizar en el estudio de la ciberseguridad, promoviendo una visión integral que permita no solo comprender los riesgos, sino también desarrollar estrategias efectivas para su mitigación. En un mundo cada vez más interconectado, la seguridad digital se convierte en un pilar fundamental para la construcción de un futuro sostenible, seguro y confiable.

Los autores

INTRODUCCIÓN



El desarrollo vertiginoso de las tecnologías digitales ha transformado profundamente la forma en que las sociedades modernas operan, interactúan y evolucionan. Actualmente, la información se posiciona como un recurso estratégico que sustenta procesos económicos, sociales, políticos y organizacionales, lo que ha generado una dependencia creciente de los sistemas informáticos. En consecuencia, esta digitalización generalizada ha propiciado no solo oportunidades de innovación y eficiencia, sino también un incremento significativo en la exposición a amenazas que comprometen la seguridad de los entornos tecnológicos.

En este sentido, la ciberseguridad se consolida como un campo fundamental dentro del ámbito de las tecnologías de la información, orientado a la protección de los sistemas, redes y datos frente a accesos no autorizados, ataques maliciosos y fallas operativas. A diferencia de enfoques tradicionales centrados exclusivamente en soluciones técnicas, la ciberseguridad contemporánea requiere una visión integral que contemple

factores humanos, organizacionales, normativos y estratégicos. Por consiguiente, su estudio implica comprender no solo los mecanismos de protección, sino también las dinámicas de ataque que evolucionan de manera constante.

Cabe destacar que el ciberespacio se ha convertido en un entorno altamente dinámico, donde interactúan diversos actores con intereses heterogéneos. Entre ellos se encuentran usuarios legítimos, organizaciones, gobiernos y agentes maliciosos que buscan explotar vulnerabilidades con fines económicos, ideológicos o políticos. Bajo esta perspectiva, el análisis de la ciberseguridad trasciende el ámbito técnico, incorporando dimensiones sociales y jurídicas que permiten entender la complejidad de los incidentes digitales en un contexto globalizado.

Por otra parte, la evolución de los ciberataques ha evidenciado un cambio significativo en su naturaleza. Inicialmente, las amenazas se caracterizaban por su simplicidad y alcance limitado; sin embargo, en la actualidad, se observa la presencia de ataques altamente sofisticados, automatizados y persistentes, capaces de eludir mecanismos tradicionales de defensa. De este modo, la aparición de amenazas persistentes avanzadas, el uso de técnicas de evasión y la incorporación de inteligencia artificial en actividades maliciosas han redefinido el panorama de la seguridad digital, exigiendo respuestas más complejas y adaptativas.

En consecuencia, la gestión del riesgo adquiere un papel central en la ciberseguridad, ya que permite identificar, evaluar y mitigar posibles vulnerabilidades antes de que sean explotadas. Este enfoque proactivo facilita la toma de decisiones informadas y contribuye a la implementación de controles adecuados para proteger los activos de información. Asimismo, la adopción de estándares internacionales y marcos de referencia proporciona una base estructurada para el desarrollo de políticas de seguridad alineadas con los objetivos organizacionales.

De igual forma, la gobernanza en ciberseguridad se posiciona como un elemento clave para garantizar la correcta gestión de los recursos tecnológicos y la supervisión de las estrategias de seguridad. A través de modelos de gobernanza, las

organizaciones pueden establecer mecanismos de control que promuevan la transparencia, la responsabilidad y la mejora continua. En este contexto, la ciberseguridad deja de ser una función aislada del área técnica y se integra como un componente estratégico dentro de la estructura organizacional.

En contraste, el estudio de los enfoques ofensivos y defensivos permite comprender la dinámica de confrontación existente en el ciberespacio. Mientras que los atacantes desarrollan técnicas cada vez más sofisticadas para vulnerar sistemas, los defensores implementan mecanismos de protección que buscan anticipar y neutralizar dichas amenazas. En este marco, prácticas como el hacking ético y las pruebas de penetración resultan fundamentales para identificar debilidades y fortalecer la seguridad de los sistemas. Paralelamente, la implementación de estrategias como Red Team y Blue Team permite simular escenarios reales de ataque y defensa, contribuyendo a mejorar la capacidad de respuesta ante incidentes.

Por su parte, el factor humano adquiere una relevancia particular dentro de la ciberseguridad. La ingeniería social, entendida como el conjunto de técnicas orientadas a manipular el comportamiento de las personas, pone de manifiesto que muchas vulnerabilidades no se encuentran en la tecnología, sino en los usuarios. En este sentido, la formación y la concientización se convierten en herramientas esenciales para reducir el riesgo de incidentes, promoviendo una cultura de seguridad que involucre a todos los actores del entorno digital.

Desde otra perspectiva, el análisis forense digital emerge como una disciplina indispensable para la investigación de incidentes de seguridad. A través de la identificación, preservación y análisis de evidencia digital, es posible reconstruir eventos, determinar responsabilidades y generar información relevante para la toma de decisiones. Este proceso no solo contribuye al esclarecimiento de delitos informáticos, sino que también permite fortalecer las estrategias de prevención mediante el aprendizaje obtenido de incidentes previos.

Adicionalmente, la importancia del análisis forense se extiende al ámbito legal, donde la correcta gestión de la evidencia digital resulta determinante para su admisibilidad en procesos judiciales.

La aplicación de procedimientos rigurosos, el cumplimiento de la cadena de custodia y la adecuada documentación de los hallazgos son aspectos fundamentales para garantizar la validez de la evidencia. En este contexto, la intersección entre tecnología y derecho adquiere una relevancia creciente, evidenciando la necesidad de profesionales con formación interdisciplinaria.

A partir de estas consideraciones, la presente obra tiene como finalidad proporcionar una visión integral de la ciberseguridad, abordando sus principales dimensiones desde un enfoque estructurado y aplicado. A lo largo del contenido, se analizan los fundamentos teóricos, los modelos de gestión, las estrategias de defensa y los procesos de análisis forense, permitiendo al lector comprender la interrelación entre estos elementos.

En términos de organización, el libro se estructura en cuatro capítulos que desarrollan de manera progresiva los aspectos más relevantes de la ciberseguridad. En primer lugar, se presentan los fundamentos, riesgos y amenazas que caracterizan el entorno digital. Posteriormente, se examinan los elementos relacionados con la gobernanza y la gestión del riesgo. En tercer lugar, se abordan los enfoques ofensivos y defensivos, así como el impacto del cibercrimen y la ingeniería social. Finalmente, se profundiza en el análisis forense digital, destacando su importancia en la investigación de incidentes y la gestión de evidencia.

En síntesis, la ciberseguridad se configura como un pilar esencial para el desarrollo de la sociedad digital, ya que permite garantizar la protección de la información, la integridad de los sistemas y la confianza en las tecnologías. Su estudio requiere un enfoque integral que combine conocimientos técnicos, estratégicos y metodológicos, así como una actitud crítica frente a los desafíos emergentes.

Esta obra se presenta como una herramienta académica orientada a fortalecer el conocimiento en ciberseguridad, promoviendo una comprensión profunda que facilite la toma de decisiones y la implementación de estrategias efectivas. En un entorno caracterizado por la incertidumbre y la constante evolución de las amenazas, la capacidad de adaptación y aprendizaje continuo se convierte en un factor determinante para enfrentar los retos del futuro digital.

01.

Fundamentos, riesgos y amenazas en ciberseguridad

1.1. Evolución histórica y fundamentos de la ciberseguridad

La ciberseguridad, tal como se concibe en la actualidad, es el resultado de un proceso histórico marcado por la evolución tecnológica y la creciente digitalización de la sociedad. A medida que las tecnologías de la información y la comunicación han avanzado, también lo han hecho las amenazas que buscan explotar sus vulnerabilidades. En este sentido, la ciberseguridad ha pasado de ser un conjunto de medidas básicas de protección a convertirse en un campo estratégico fundamental para la protección de individuos, organizaciones y Estados. Su desarrollo ha estado estrechamente ligado al crecimiento de Internet, la interconectividad global y la transformación digital, factores que han incrementado tanto las oportunidades como los riesgos en el entorno digital.

a) Los inicios (años 60–80)

Durante las décadas de 1960 y 1970, los sistemas informáticos eran limitados, costosos y accesibles únicamente para instituciones académicas, gubernamentales y grandes corporaciones. En este contexto, la seguridad informática se centraba principalmente en el control del acceso físico a los equipos, ya que las amenazas externas eran prácticamente inexistentes. Sin embargo, con la creación de ARPANET en 1969, considerada la precursora de Internet, surgió un nuevo paradigma: la interconexión de sistemas a distancia. Este avance hizo evidente la necesidad de proteger no solo los equipos, sino también la información que circulaba entre ellos.

En este periodo también se registraron los primeros experimentos relacionados con software malicioso. Un ejemplo representativo es el programa Creeper (1971), considerado el primer virus informático, que se propagaba a través de la red mostrando un mensaje en los sistemas infectados. Aunque su propósito era experimental, marcó el inicio de una nueva categoría de riesgos en el ámbito digital.

b) Expansión del malware (años 80–90)

La década de 1980 estuvo marcada por la popularización de las computadoras personales, lo que amplió significativamente la superficie de ataque. Durante este periodo, los virus comenzaron a propagarse mediante disquetes, afectando sistemas de manera local. A diferencia de los experimentos previos, estos programas ya tenían la capacidad de replicarse y causar daños a los archivos y sistemas.

Uno de los eventos más relevantes de esta etapa fue el gusano Morris (1988), uno de los primeros ataques a gran escala en Internet, que logró infectar aproximadamente el 10% de los sistemas conectados en ese momento. Este incidente evidenció la vulnerabilidad de las redes y la necesidad de desarrollar mecanismos más sofisticados de defensa.

En la década de 1990, con la expansión de Internet a nivel global, surgieron nuevas amenazas como los troyanos, gusanos más avanzados y ataques de ingeniería social. Estos últimos introdujeron un componente psicológico en la seguridad, al explotar la confianza de los usuarios para obtener información sensible. Así, la ciberseguridad comenzó a evolucionar hacia un enfoque más integral que incluía tanto aspectos técnicos como humanos.

c) Profesionalización del cibercrimen (años 2000)

A partir de los años 2000, el crecimiento exponencial de Internet y la digitalización de servicios financieros, comerciales y gubernamentales propiciaron la aparición de un nuevo fenómeno: la profesionalización del cibercrimen. Los ataques dejaron de ser realizados únicamente por entusiastas o investigadores y pasaron a ser ejecutados por grupos organizados con fines económicos.

Durante esta etapa, se consolidaron prácticas como el phishing, el robo de identidad y los ataques de denegación de servicio (DoS y DDoS), los cuales permitían interrumpir el funcionamiento de servicios en línea. Asimismo, comenzaron a surgir mercados clandestinos en la red para la compra y venta de datos robados, herramientas de hacking y servicios ilícitos, lo que fortaleció la economía del cibercrimen.

d) La era de las amenazas avanzadas (2010 en adelante)

A partir de 2010, el panorama de la ciberseguridad experimentó un cambio significativo con la aparición de las Amenazas Persistentes Avanzadas (APT, por sus siglas en inglés). Estas amenazas se caracterizan por ser altamente sofisticadas, dirigidas a objetivos específicos y ejecutadas durante largos periodos de tiempo sin ser detectadas. A diferencia de los ataques tradicionales, las APT suelen estar vinculadas a intereses estratégicos, como el espionaje industrial o la obtención de información gubernamental.

Un caso emblemático de esta etapa es Stuxnet (2010), un malware diseñado para sabotear infraestructuras industriales en Irán. Este incidente demostró que los ciberataques podían tener consecuencias físicas en el mundo real, afectando infraestructuras críticas. A partir de entonces, la ciberseguridad dejó de ser considerada únicamente un problema técnico para convertirse en un asunto de seguridad nacional y geopolítica.

e) Ciberseguridad en la actualidad

En la actualidad, la ciberseguridad enfrenta desafíos cada vez más complejos debido a la constante evolución de la tecnología y la sofisticación de los atacantes. Entre las principales amenazas destacan el ransomware, que cifra la información de las víctimas a cambio de un rescate económico; el ciberespionaje, orientado al robo de información estratégica; y los ataques a infraestructuras críticas, que pueden afectar servicios esenciales como la electricidad, el agua o el transporte.

Asimismo, el auge de tecnologías como la computación en la nube, el Internet de las Cosas (IoT) y la inteligencia artificial ha ampliado considerablemente la superficie de ataque, generando nuevos riesgos que requieren enfoques innovadores de protección. En este contexto, la ciberseguridad se consolida como un pilar fundamental para el desarrollo seguro de la sociedad digital, demandando la integración de soluciones tecnológicas, marcos normativos y una cultura de seguridad en todos los niveles.

Una de las principales aportaciones para comprender la evolución de la ciberseguridad proviene de Dimitrov (2020), quien explica que los orígenes de la seguridad digital están estrechamente vinculados al desarrollo de la inteligencia militar y tecnológica durante la Guerra Fría. Este autor señala que los primeros sistemas informáticos no fueron diseñados con fines de seguridad, sino de procesamiento de información, lo que explica por qué las primeras amenazas eran limitadas y experimentales. Esta perspectiva complementa la etapa inicial descrita en el texto, reforzando la idea de que la ciberseguridad

surge como respuesta a la interconexión y al valor creciente de la información.

En relación con la evolución de los ciberataques, Raghuwanshi et al. (2025) destacan que el desarrollo del malware ha seguido una trayectoria paralela al avance tecnológico, pasando de programas experimentales a herramientas altamente sofisticadas utilizadas con fines económicos y estratégicos. Estos autores enfatizan que, especialmente a partir de los años 2000, los ataques comenzaron a estructurarse como actividades organizadas, lo que coincide con la profesionalización del cibercrimen mencionada en el texto. Además, subrayan la importancia del factor humano y de la ingeniería social como elementos clave en la efectividad de los ataques modernos.

Por su parte, Tzavara y Vassiliadis (2024) aportan un enfoque conceptual al introducir el término de resiliencia cibernética, que va más allá de la simple protección de sistemas. Según estos autores, la ciberseguridad actual no solo debe centrarse en prevenir ataques, sino también en la capacidad de las organizaciones para resistir, adaptarse y recuperarse ante incidentes. Esta idea amplía el apartado contemporáneo del texto, donde se menciona la necesidad de enfoques más integrales frente a amenazas avanzadas como las APT.

Desde una perspectiva más técnica y estructural, Aslan et al. (2023) ofrecen una clasificación detallada de vulnerabilidades, amenazas y tipos de ataques, lo que permite comprender mejor la complejidad del ecosistema actual de ciberseguridad. Su trabajo resalta que las debilidades pueden originarse tanto en el software como en el hardware o en los propios usuarios, lo que refuerza la idea de que la seguridad debe abordarse de manera multidimensional. Asimismo, destacan la necesidad de implementar soluciones combinadas, como sistemas de detección, cifrado y formación en seguridad.

En cuanto a las tendencias actuales, Zaid y Garai (2024) señalan que tecnologías emergentes como la inteligencia artificial, el Internet de las Cosas y la computación en la nube están redefiniendo el panorama de la ciberseguridad.

Estos autores advierten que, aunque estas tecnologías aportan beneficios significativos, también introducen nuevas superficies de ataque y riesgos más complejos. Esta aportación se vincula directamente con el apartado final del texto, donde se menciona la expansión de los riesgos debido a la transformación digital.

Finalmente, Cremer et al. (2022) realizan una contribución clave al analizar la disponibilidad de datos en el ámbito del riesgo cibernético. Su estudio pone de manifiesto que existe una falta significativa de datos estandarizados y accesibles sobre incidentes de ciberseguridad, lo que dificulta la evaluación del riesgo y la toma de decisiones informadas. Esta limitación afecta tanto a empresas como a gobiernos, y refuerza la necesidad de desarrollar marcos más sólidos de recopilación y análisis de información, aspecto fundamental en la gestión moderna de la ciberseguridad.

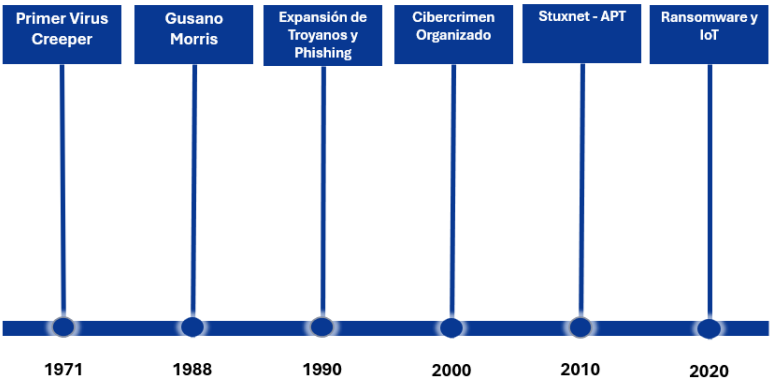


Figura 1.1. Línea del tiempo de la evolución de la ciberseguridad.

La Figura 1.1 ilustra de manera sintética la evolución histórica de la ciberseguridad, destacando los principales hitos que han marcado su desarrollo desde los inicios de la computación hasta la actualidad. En primer lugar, se observa que durante las décadas de 1960 y 1970 la seguridad se centraba fundamentalmente en el control físico de los sistemas, debido a que las computadoras eran recursos limitados y de acceso restringido. No obstante, la aparición de redes como ARPANET marcó el inicio de la interconectividad, introduciendo nuevas necesidades de protección en la transmisión de datos.

Posteriormente, la figura evidencia cómo en las décadas de 1980 y 1990 se produjo una expansión significativa del malware, impulsada por la popularización de las computadoras personales y el crecimiento de Internet. Durante este periodo surgieron virus, gusanos y troyanos que comenzaron a explotar vulnerabilidades en los sistemas, lo que obligó al desarrollo de las primeras soluciones antivirus y políticas básicas de seguridad. Este momento representa el tránsito desde una seguridad reactiva hacia una más estructurada.

En la etapa correspondiente a los años 2000, la línea del tiempo refleja la profesionalización del cibercrimen, caracterizada por la aparición de ataques más organizados y con fines económicos. El incremento del comercio electrónico y los servicios en línea favoreció el desarrollo de técnicas como el phishing y los ataques de denegación de servicio, consolidando un ecosistema delictivo digital cada vez más complejo y globalizado.

A partir de 2010, la figura destaca la transición hacia amenazas avanzadas, como las Amenazas Persistentes Avanzadas (APT), que se distinguen por su alto nivel de sofisticación y su enfoque dirigido a objetivos específicos. En este contexto, los ciberataques comienzan a tener implicaciones geopolíticas, afectando infraestructuras críticas y evidenciando la relevancia de la ciberseguridad en la seguridad nacional.

Finalmente, la figura muestra que en la actualidad la ciberseguridad se enfrenta a un entorno altamente dinámico, donde convergen tecnologías emergentes como la computación en la nube, el Internet de las Cosas y la inteligencia artificial. Estos avances han ampliado la superficie de ataque, incrementando la complejidad de las amenazas y requiriendo enfoques integrales que combinen tecnología, regulación y concienciación del usuario. En conjunto, la línea del tiempo permite comprender que la ciberseguridad ha evolucionado de manera paralela al desarrollo tecnológico, adaptándose constantemente a nuevos riesgos y desafíos.

Seguridad de la información vs. ciberseguridad

En la actualidad, una definición muy utilizada es ciberseguridad, la cual puede asociarse con otras palabras como ciberespacio, ciberamenazas, cibercrimen entre otros conceptos. Sin embargo, suele utilizarse como sinónimo de seguridad informática o seguridad de la información lo cual hace que se haga de manera errónea.

En este tema del capítulo uno, abordaremos las definiciones de seguridad de la información y ciberseguridad para poder apreciar sus diferencias en la práctica.

Seguridad de la Información

Es un término más general que se refiere a la protección de información en cualquier formato, ya sea digital, físico o verbal, y cuya relación es asegurar los tres pilares de la confidencialidad, integridad y disponibilidad (Figura 1.2).

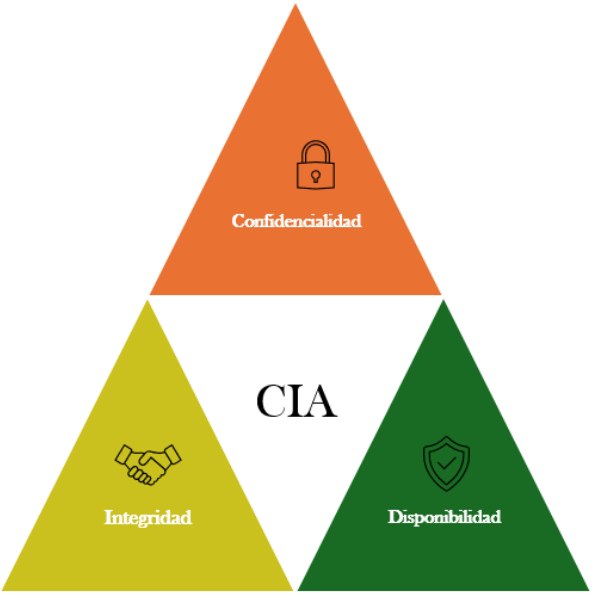


Figura 1.2. Confidentiality, Integrity and Availability (CIA).

- Confidencialidad:

Es garantizar que la información solo sea accesible para usuarios autorizados y así, poder evitar el acceso no autorizado a la información sensible.

- Integridad:

Es asegurar que la información sea la correcta y no haya sufrido ninguna alteración de manera no autorizada.

- Disponibilidad:

Es garantizar que la información y los sistemas sean accesibles de manera confiable y oportuna para los usuarios autorizados cuando la necesiten.

- Ciberseguridad

En cambio, se diferencia en lo que protege los sistemas informáticos, las redes y la información digital de daños, robos o alteraciones causados por personas no autorizadas.

La Tabla 1.1 presenta una comparación conceptual entre la seguridad de la información y la ciberseguridad, destacando sus principales diferencias en términos de enfoque, ámbito y objetivos. A partir de esta comparación, se evidencia que la seguridad de la información constituye un concepto más amplio, orientado a la protección de la información en cualquier formato, ya sea físico o digital. En contraste, la ciberseguridad se enfoca específicamente en la protección de la información en entornos digitales, lo que la posiciona como una subárea especializada dentro del campo general.

En relación con el ámbito, la tabla muestra que la seguridad de la información abarca tanto dimensiones físicas como digitales, lo que implica la implementación de controles diversos, incluyendo medidas administrativas, físicas y tecnológicas. Por su parte, la ciberseguridad se limita principalmente al entorno digital, centrando su atención en la protección de sistemas informáticos, redes y datos frente a amenazas cibernéticas. Esta diferencia resalta el carácter integral de la seguridad de la información frente al enfoque más técnico y específico de la ciberseguridad.

Finalmente, en cuanto a los objetivos, la seguridad de la información se fundamenta en la preservación de los principios de confidencialidad, integridad y disponibilidad de la información. La ciberseguridad,

aunque contribuye a estos principios, se orienta de manera más directa a la prevención, detección y respuesta ante ataques informáticos. En conjunto, la tabla permite comprender que la ciberseguridad actúa como un componente operativo dentro del marco más amplio de la seguridad de la información, siendo ambas disciplinas complementarias en la protección de los activos informacionales de una organización.

Tabla 1.1. Aspectos de la Seguridad de la información Vs. Ciberseguridad.

Aspecto	Seguridad de la Información	Ciberseguridad
Enfoque	Información en general	Información digital y sistemas
Ámbito	Físico y digital	Principalmente digital
Objetivo	CIA de la información	Prevención y respuesta a ataques

La protección de la información constituye un aspecto fundamental en la sociedad digital contemporánea, abarcando distintos niveles como el personal, organizacional y gubernamental. A nivel personal, es esencial resguardar la identidad del individuo, sus datos personales almacenados en línea y los dispositivos que utiliza, frente a posibles ciberataques. Entre los datos más sensibles se encuentran los registros médicos, educativos y financieros, cuya exposición puede generar consecuencias significativas para la privacidad y seguridad del individuo. En la actualidad, la digitalización de la información ha provocado que tanto empresas como instituciones públicas almacenen estos datos en plataformas digitales en lugar de medios físicos, lo que incrementa su accesibilidad, pero también su vulnerabilidad ante amenazas cibernéticas.

En el ámbito organizacional o empresarial, la protección de la información adquiere una dimensión estratégica, ya que involucra no solo la seguridad de los datos sensibles, sino también la reputación de la empresa y la confianza de sus clientes. Las organizaciones deben resguardar información crítica como datos de producción, procesos

de compra-venta, propiedad intelectual, estados financieros y declaraciones fiscales. La exposición o manipulación indebida de esta información puede afectar gravemente la competitividad, estabilidad y credibilidad de la empresa en el mercado.

Por su parte, a nivel gubernamental, la seguridad de la información se vincula directamente con la preservación de la seguridad nacional y el bienestar de la sociedad. Los gobiernos tienen la responsabilidad de proteger datos sensibles del Estado, garantizar la estabilidad económica y salvaguardar la seguridad de los ciudadanos en el entorno digital. La gestión adecuada de la información en este nivel es crucial para prevenir amenazas que puedan comprometer la infraestructura crítica, la gobernabilidad y la confianza pública.

Para garantizar una adecuada protección de la información, es necesario considerar diversos componentes clave. En primer lugar, la seguridad operativa comprende los procesos y decisiones relacionados con la gestión y protección de los recursos de datos, incluyendo la asignación de permisos de acceso a los usuarios y la definición de políticas sobre el almacenamiento y la compartición de la información. Asimismo, la recuperación ante desastres y la continuidad del negocio constituyen elementos esenciales, ya que permiten a las organizaciones establecer mecanismos de respuesta ante incidentes que interrumpan sus operaciones o generen pérdida de datos, asegurando la restauración oportuna de los servicios y la información.

Otro aspecto fundamental es la capacitación del personal, dado que el factor humano representa una de las principales vulnerabilidades en materia de ciberseguridad. La formación continua en buenas prácticas, como la identificación de correos electrónicos sospechosos o el uso adecuado de dispositivos externos, contribuye significativamente a la prevención de incidentes de seguridad.

En este contexto, el modelo del cubo de John McCumber (1991) proporciona un marco teórico integral para comprender la seguridad informática. Este modelo se

fundamenta en tres dimensiones interrelacionadas. La primera corresponde a los principios de seguridad, conocidos como la triada CIA: confidencialidad, integridad y disponibilidad. La confidencialidad implica garantizar que la información solo sea accesible para usuarios autorizados, mediante mecanismos como el cifrado y el control de acceso. La integridad se refiere a la protección de la exactitud y consistencia de los datos, evitando modificaciones no autorizadas. Por último, la disponibilidad asegura que la información esté accesible cuando sea requerida.

La segunda dimensión del modelo aborda el estado de la información en el ámbito digital, distinguiendo entre datos en procesamiento, almacenamiento y transmisión. Esto permite identificar los riesgos asociados a cada fase del ciclo de vida de la información y aplicar medidas de protección adecuadas en cada caso.

Finalmente, la tercera dimensión se centra en las contramedidas de ciberseguridad, que incluyen el uso de tecnologías como antivirus y sistemas de protección, la implementación de políticas y controles administrativos, y la participación activa de las personas mediante la capacitación continua. Estos elementos evidencian que la seguridad de la información y la ciberseguridad requieren un enfoque integral que combine tecnología, procesos y formación humana, con el fin de enfrentar de manera efectiva las amenazas del entorno digital actual.

Figura 1.3. El cubo de John MCcumber.

La Figura 1.3 representa el modelo conocido como el cubo de McCumber (1991), una de las aproximaciones conceptuales más influyentes en el ámbito de la ciberseguridad. Este modelo propone una visión tridimensional de la seguridad de la información, integrando tres dimensiones fundamentales que permiten comprender de manera estructurada cómo proteger los sistemas y los datos.

En primer lugar, una de las dimensiones del cubo está compuesta por los principios clásicos de la seguridad de la información: confidencialidad, integridad y disponibilidad. Estos elementos constituyen la base sobre la cual se diseñan las políticas y mecanismos de protección, asegurando que la información solo sea accesible para usuarios autorizados, que se mantenga sin alteraciones indebidas y que esté disponible cuando se necesite.

En segundo lugar, el modelo incorpora los estados de la información, que incluyen el almacenamiento, el procesamiento y la transmisión de los datos. Esta dimensión resalta que la información debe ser protegida en todo su ciclo de vida, ya que las vulnerabilidades pueden surgir en cualquiera de estas fases. De este modo, el cubo enfatiza que la seguridad no debe centrarse únicamente en un punto específico, sino abarcar todos los momentos en los que la información es utilizada o manipulada.

Por último, la tercera dimensión del cubo hace referencia a las medidas de seguridad, que se dividen en controles tecnológicos, organizativos y humanos. Esta clasificación pone de manifiesto que la ciberseguridad no depende exclusivamente de soluciones técnicas, sino también de políticas institucionales y del comportamiento de los usuarios. En este sentido, el modelo reconoce el papel crítico del factor humano como uno de los principales puntos de vulnerabilidad.

1.2. Análisis de riesgos, amenazas y vulnerabilidades en ciberseguridad

En el ámbito de la ciberseguridad, el análisis de riesgos, amenazas y vulnerabilidades constituye un

pilar fundamental para la protección de los sistemas de información y los activos digitales. La comprensión de estos conceptos permite identificar posibles escenarios de ataque, evaluar su impacto y establecer estrategias eficaces de mitigación. En este sentido, la ciberseguridad no solo se centra en la implementación de medidas tecnológicas, sino también en la gestión integral del riesgo, considerando factores técnicos, organizativos y humanos.

En línea con esta perspectiva, diversos estudios recientes destacan que el entorno actual se caracteriza por una creciente complejidad y sofisticación de los ataques, lo que exige enfoques multidimensionales para su análisis y mitigación. En particular, Li y Liu (2021) señalan que la evolución de los ciberataques ha incrementado la necesidad de integrar análisis predictivos y estrategias adaptativas, mientras que Admass et al. (2024) subrayan que la ciberseguridad moderna debe abordar simultáneamente desafíos técnicos, regulatorios y humanos para garantizar una protección efectiva.

El concepto de riesgo se refiere a la probabilidad de que una amenaza se materialice aprovechando una vulnerabilidad existente, generando un impacto negativo sobre un sistema de información. En otras palabras, el riesgo surge de la interacción entre amenazas y vulnerabilidades, y su nivel depende tanto de la probabilidad de ocurrencia como de las consecuencias derivadas del incidente. Por tanto, la evaluación del riesgo implica asumir la existencia de debilidades en los sistemas y la posibilidad de que estas sean explotadas.

En este contexto, Razaque et al. (2025) destacan que en entornos complejos como el Internet de las Cosas, la interconexión de dispositivos amplifica significativamente el riesgo, debido a la coexistencia de múltiples vulnerabilidades en distintos niveles (red, aplicación y hardware). Asimismo, Lallie et al. (2025) evidencian que sectores específicos, como el universitario, presentan patrones de riesgo particulares derivados de la diversidad de usuarios y sistemas, lo que incrementa la exposición a amenazas tanto internas como externas.

La gestión del riesgo se define como el proceso sistemático mediante el cual se identifican, analizan y evalúan los riesgos, con el objetivo de implementar medidas que permitan reducirlos a un nivel aceptable y mantenerlos bajo control. Este proceso se articula en torno a varios principios fundamentales y requiere un enfoque continuo y adaptativo. En este sentido, Ksibi et al. (2022) proponen modelos cuantitativos de gestión del riesgo que permiten medir de forma más precisa el impacto de las amenazas, especialmente en sectores críticos como la salud digital, donde la protección de los datos es esencial.

Por su parte, Awan y Alam (2025) destacan la importancia de adaptar las estrategias de gestión del riesgo a las características específicas de las organizaciones, particularmente en pequeñas y medianas empresas, donde las limitaciones de recursos requieren soluciones eficientes y escalables. El analisis de estas aportaciones evidencia que la gestión del riesgo en ciberseguridad debe ser un proceso dinámico, basado en la evaluación constante, la implementación de controles adecuados y la capacidad de adaptación frente a un entorno de amenazas en constante evolución.

Por otra parte, en el proceso de evaluación de riesgos y determinación de necesidades de protección, resulta fundamental, en primer lugar, identificar los activos de información críticos y establecer procedimientos de análisis alineados con los objetivos estratégicos de la organización. En segundo lugar, es necesario adoptar un enfoque centralizado de gestión, que implique la designación de responsables, la asignación adecuada de recursos y la garantía de una capacitación continua del personal.Asimismo, la gestión del riesgo requiere la implementación de políticas y controles de seguridad adecuados, alineados con los riesgos identificados y con los objetivos estratégicos de la organización. Estas políticas deben diferenciarse de las directrices operativas y estar respaldadas por mecanismos de supervisión efectivos. Otro aspecto clave es la promoción de la concienciación y la formación en ciberseguridad, ya que el factor humano constituye uno de los eslabones

más vulnerables en la cadena de seguridad. Finalmente, es imprescindible llevar a cabo un proceso continuo de supervisión y evaluación, que permita medir la eficacia de las medidas implementadas y realizar ajustes en función de la evolución del entorno de amenazas.

Por otro lado, el concepto de vulnerabilidad hace referencia a cualquier debilidad o fallo presente en un sistema de información que pueda ser explotado por una amenaza. Estas debilidades pueden comprometer la confidencialidad, integridad y disponibilidad de la información, y pueden originarse por múltiples factores, como errores de diseño, fallos de programación, configuraciones inadecuadas o falta de actualizaciones. La identificación y gestión de vulnerabilidades constituye una de las tareas esenciales en ciberseguridad, ya que permite anticipar posibles puntos de ataque y reducir la superficie de exposición.

Las vulnerabilidades pueden clasificarse en diferentes categorías en función de su naturaleza. Las vulnerabilidades físicas están relacionadas con el acceso no autorizado a infraestructuras tecnológicas, como centros de datos o equipos informáticos, lo que puede implicar robos, sabotajes o interrupciones del servicio. Las vulnerabilidades naturales hacen referencia a eventos derivados del entorno, como desastres naturales (inundaciones, incendios, terremotos), que pueden afectar la disponibilidad de los sistemas y la integridad de la información.

Asimismo, las vulnerabilidades de hardware se vinculan con defectos en los componentes físicos de los sistemas o con una configuración inadecuada de los mismos, lo que puede facilitar accesos no autorizados o fallos en el funcionamiento. Por su parte, las vulnerabilidades de software son especialmente críticas, ya que derivan de errores en el código, fallos en los sistemas operativos o aplicaciones, y pueden ser explotadas para ejecutar ataques, alterar datos o acceder a información sensible. Dentro de esta categoría también se incluyen las vulnerabilidades relacionadas con los medios de almacenamiento y los sistemas de comunicación, donde un uso inadecuado o la falta de protección puede

comprometer la seguridad de los datos durante su almacenamiento o transmisión.

Finalmente, las vulnerabilidades humanas constituyen uno de los factores más relevantes en el ámbito de la ciberseguridad. Estas se derivan de errores, negligencias o falta de formación por parte de los usuarios, así como de acciones malintencionadas por parte de individuos internos o externos a la organización. La falta de concienciación en seguridad, el uso inadecuado de los sistemas o la susceptibilidad a técnicas de ingeniería social son ejemplos claros de este tipo de vulnerabilidad.

En relación con el concepto de amenaza, esta se define como cualquier acción, evento o agente capaz de explotar una vulnerabilidad y causar un daño a un sistema de información. Las amenazas pueden tener diversos orígenes, incluyendo ataques deliberados (como malware, phishing o accesos no autorizados), eventos físicos (incendios, fallos eléctricos) o decisiones organizativas inadecuadas. En función de su procedencia, las amenazas pueden clasificarse en internas, cuando provienen de empleados o colaboradores de la organización, y externas, cuando son ejecutadas por actores ajenos, como ciberdelincuentes o grupos organizados.

Desde una perspectiva analítica, la relación entre riesgo, vulnerabilidad y amenaza puede abordarse desde diferentes enfoques. El enfoque de las ciencias aplicadas se centra en la estimación de daños potenciales, considerando tanto la probabilidad de ocurrencia de un ataque como la fragilidad de los sistemas afectados. Este enfoque permite cuantificar el riesgo y establecer prioridades en la gestión de la seguridad. Por otro lado, el enfoque de la reducción del riesgo pone énfasis en la prevención y mitigación, proponiendo intervenciones orientadas a disminuir las condiciones de vulnerabilidad y fortalecer los mecanismos de protección en todos los niveles.

La clasificación de las vulnerabilidades constituye un elemento esencial dentro del análisis de la seguridad de los sistemas de información, ya que permite identificar,

categorizar y abordar de manera estructurada las debilidades que pueden ser explotadas por diferentes amenazas. En este sentido, la Figura 1.4 presenta una tipología de vulnerabilidades que abarca diversas dimensiones del entorno tecnológico y organizativo, evidenciando que los riesgos en ciberseguridad no se limitan únicamente al ámbito digital, sino que también incluyen factores físicos, ambientales y humanos.

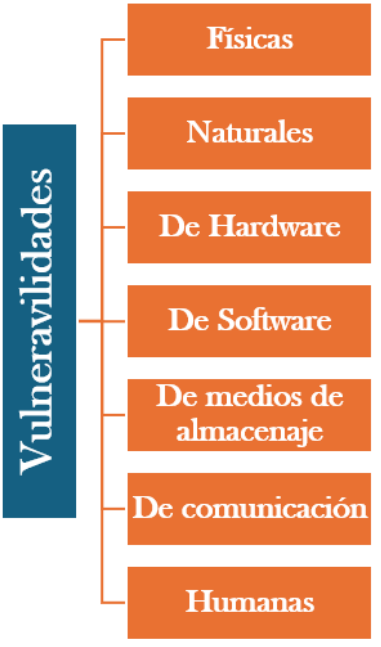


Figura 1.4. Clasificación de vulnerabilidades.

En primer lugar, las vulnerabilidades físicas hacen referencia a aquellas debilidades relacionadas con el entorno donde se alojan los sistemas de información, como centros de datos, salas de servidores o infraestructuras tecnológicas. Este tipo de vulnerabilidad implica riesgos derivados del acceso no autorizado a instalaciones, la manipulación directa de equipos, la interrupción del suministro eléctrico o el robo de dispositivos. La protección frente a estas amenazas requiere la implementación de medidas de seguridad física, tales como controles de acceso, sistemas de vigilancia, mecanismos de detección de intrusiones y planes de contingencia ante fallos en la infraestructura.

Por otra parte, las vulnerabilidades naturales están asociadas a fenómenos ambientales o desastres naturales que pueden comprometer la disponibilidad e integridad de los sistemas de información. Eventos como inundaciones, incendios, terremotos o variaciones extremas de temperatura pueden afectar gravemente a las infraestructuras tecnológicas. Para mitigar estos riesgos, resulta imprescindible adoptar estrategias como la realización de copias de seguridad periódicas, la implementación de sistemas de energía redundantes, el uso de infraestructuras resilientes y el diseño de planes de recuperación ante desastres.

En el ámbito técnico, las vulnerabilidades de hardware se refieren a defectos o fallos en los componentes físicos de los sistemas informáticos, así como a configuraciones inadecuadas o a la falta de mantenimiento. Estas debilidades pueden facilitar accesos no autorizados o provocar fallos en el funcionamiento de los equipos. Ejemplos representativos incluyen errores en la memoria RAM, problemas en los dispositivos de almacenamiento o vulnerabilidades en interfaces físicas como puertos USB o tecnologías de conexión avanzada. Asimismo, modificaciones en el firmware o en sistemas de arranque como BIOS o UEFI pueden comprometer gravemente la seguridad del sistema, permitiendo un control total por parte de un atacante.

En relación con lo anterior, las vulnerabilidades de software constituyen una de las principales fuentes de riesgo en los sistemas de información actuales. Estas se originan en errores de programación, fallos en los sistemas operativos, configuraciones incorrectas o ausencia de actualizaciones de seguridad. Tales debilidades pueden ser explotadas para ejecutar código malicioso, alterar la integridad de los datos o acceder a información confidencial sin autorización. Dentro de esta categoría también se incluyen las vulnerabilidades asociadas a los medios de almacenamiento, donde el uso inadecuado de soportes físicos o digitales puede comprometer la seguridad de la información, así como las vulnerabilidades en los sistemas de comunicación, que afectan la transmisión de datos a través de

redes cableadas o inalámbricas, exponiéndolos a interceptación o manipulación.

De igual forma, las vulnerabilidades humanas representan uno de los factores más críticos en la seguridad de la información. Estas se derivan tanto de errores involuntarios como de acciones malintencionadas por parte de los usuarios. La falta de formación en ciberseguridad, la escasa concienciación sobre buenas prácticas, la negligencia en el manejo de la información o la susceptibilidad a ataques de ingeniería social son algunos de los principales elementos que incrementan el riesgo. Asimismo, las amenazas internas, ya sean intencionales o accidentales, pueden tener un impacto significativo en la seguridad de los sistemas, lo que pone de manifiesto la necesidad de fomentar una cultura organizacional orientada a la seguridad.

En cuanto al concepto de amenaza, esta se define como cualquier acción o evento que tiene la capacidad de explotar una vulnerabilidad y causar daño a un sistema de información. Las amenazas pueden originarse a partir de ataques deliberados, como malware, accesos no autorizados o robo de información, así como de eventos físicos o decisiones organizativas inadecuadas. En función de su origen, se distinguen amenazas internas, procedentes de individuos con acceso legítimo al sistema, y amenazas externas, ejecutadas por actores ajenos a la organización, como ciberdelincuentes o grupos especializados.

Finalmente, la relación entre vulnerabilidades, amenazas y riesgo puede analizarse desde diferentes enfoques teóricos. El enfoque de las ciencias aplicadas se centra en la evaluación cuantitativa del riesgo, considerando la probabilidad de ocurrencia de un incidente y su impacto potencial sobre los sistemas. Este enfoque permite estimar pérdidas y establecer prioridades en la gestión de la seguridad. Por otro lado, el enfoque de la reducción del riesgo pone énfasis en la implementación de medidas preventivas y correctivas destinadas a disminuir las condiciones de vulnerabilidad, abordando tanto los factores técnicos como las causas estructurales que incrementan la exposición al riesgo.

1.3. Clasificación y tipología de amenazas en ciberseguridad

Las amenazas representan uno de los elementos clave en el análisis de la seguridad de los sistemas de información, ya que constituyen cualquier acción, evento o circunstancia capaz de explotar una vulnerabilidad y comprometer la integridad, confidencialidad o disponibilidad de los datos. Estas amenazas pueden originarse a partir de múltiples fuentes y responder a diversas causas, entre las que destacan los ataques maliciosos, los errores humanos y los factores externos.

La creciente digitalización de la sociedad y la interconectividad de los sistemas han contribuido a la aparición de un entorno cada vez más complejo, en el que las amenazas evolucionan constantemente en términos de sofisticación, alcance y capacidad de impacto. En este contexto, diversos estudios han subrayado la necesidad de abordar las amenazas desde una perspectiva estructurada y multidimensional. En particular, Jouini et al. (2014) proponen una clasificación sistemática de las amenazas en sistemas de información, destacando la importancia de distinguir entre amenazas accidentales y deliberadas, así como entre aquellas de origen interno y externo, lo que permite comprender mejor su naturaleza y facilitar su gestión.

Asimismo, investigaciones más recientes, como la de Almaiah et al. (2024), enfatizan que las amenazas en entornos específicos, como los sistemas de bases de datos, presentan características propias que requieren enfoques de protección especializados. Estos autores destacan que la creciente dependencia de los datos en las organizaciones ha convertido a las bases de datos en objetivos prioritarios para los atacantes, lo que incrementa la necesidad de implementar mecanismos avanzados de detección y respuesta. En una línea similar, Liu et al. (2022) señalan que en sectores como el comercio electrónico las amenazas son constantes y evolucionan rápidamente, configurando un desafío permanente que exige estrategias de seguridad dinámicas y adaptativas.

Por otro lado, Abdullah et al. (2025) analizan la evolución del cibercrimen a partir de datos históricos, evidenciando

que las amenazas han experimentado un crecimiento significativo tanto en volumen como en complejidad, impulsado por factores como la globalización digital y el desarrollo tecnológico. Este fenómeno se traduce en la aparición de nuevos vectores de ataque y en la profesionalización de los actores maliciosos. En este sentido, Darem et al. (2023) destacan que en sectores críticos como el bancario y financiero, las amenazas no solo buscan beneficios económicos, sino que también pueden comprometer la estabilidad de los sistemas y la confianza de los usuarios, lo que refuerza la necesidad de implementar contramedidas robustas y específicas.

De esta manera, el factor humano continúa siendo un elemento determinante en la materialización de las amenazas. Según Baltuttis et al. (2024), el comportamiento de los usuarios dentro de las organizaciones influye directamente en la exposición al riesgo, ya que las decisiones, hábitos y nivel de concienciación en seguridad pueden facilitar o dificultar la ejecución de ataques. Esta perspectiva pone de relieve que la ciberseguridad no puede abordarse únicamente desde un enfoque tecnológico, sino que debe integrar también dimensiones conductuales y organizativas.

Estos estudios evidencian que las amenazas en ciberseguridad no solo son cada vez más sofisticadas, sino también más diversas y dinámicas, lo que exige un enfoque integral que combine clasificación, análisis contextual y adaptación continua. La comprensión de su evolución y características resulta esencial para el diseño de estrategias de protección eficaces en un entorno digital en constante transformación.

Para facilitar su análisis y comprensión, las amenazas en ciberseguridad pueden clasificarse en diferentes categorías, siendo una de las más relevantes la de las amenazas humanas. Estas se refieren a aquellas acciones que provienen directamente de individuos, ya sea de manera intencionada o accidental, y que aprovechan debilidades en los sistemas para causar algún tipo de daño o acceder a información sensible. Este tipo de amenazas resulta especialmente significativo

debido a que el factor humano es considerado uno de los eslabones más débiles en la cadena de seguridad.

Dentro de las amenazas humanas, la ingeniería social constituye una de las técnicas más utilizadas por los atacantes. Esta estrategia se basa en la manipulación psicológica de las personas con el objetivo de obtener información confidencial o acceso a sistemas informáticos. A diferencia de otros tipos de ataques, la ingeniería social no requiere necesariamente conocimientos técnicos avanzados, sino que se fundamenta en el engaño, la persuasión y la explotación de la confianza del usuario. En este contexto, prácticas como el "trashing" o cartoneo evidencian cómo la falta de concienciación en seguridad puede derivar en incidentes graves, ya que la simple acción de desechar documentos con información sensible sin las medidas adecuadas puede facilitar el acceso no autorizado a sistemas.

Por otro lado, existen distintos perfiles de atacantes dentro de las amenazas humanas. Los terroristas informáticos, por ejemplo, buscan causar daño o desestabilizar sistemas sin un interés económico directo, mientras que los intrusos pagados actúan por encargo de terceros, generalmente con fines económicos o estratégicos, y suelen contar con un alto nivel de especialización. Asimismo, los exempleados representan una amenaza significativa debido a su conocimiento previo de los sistemas y procesos internos de la organización, lo que les permite explotar vulnerabilidades específicas con relativa facilidad. En muchos casos, la falta de una correcta gestión de accesos tras la finalización de la relación laboral permite que estos individuos mantengan privilegios indebidos.

A su vez, los denominados "curiosos" o usuarios con interés en la tecnología, aunque no siempre tengan intenciones maliciosas, pueden causar daños debido a su falta de experiencia o conocimiento. Por último, el personal interno constituye una fuente importante de riesgo, ya que los errores, omisiones o negligencias en el uso de los sistemas pueden generar incidentes de seguridad de gran impacto. Incluso acciones aparentemente simples, como un corte de energía

provocado por personal no especializado, pueden tener consecuencias graves en la integridad de los datos.

Otra categoría fundamental es la de las amenazas lógicas, que hacen referencia a aquellas que se materializan a través de software malicioso o herramientas informáticas diseñadas para comprometer los sistemas. Estas amenazas suelen ser creadas de manera intencionada y se caracterizan por su capacidad de propagación, ocultamiento y daño.

Entre los tipos más comunes de amenazas lógicas se encuentra el adware, que consiste en software diseñado para mostrar publicidad no deseada, generalmente instalado junto con otros programas. Aunque en muchos casos su impacto es limitado, puede afectar la experiencia del usuario y, en ocasiones, recopilar información sin su consentimiento. Por su parte, las puertas traseras o backdoors permiten a los atacantes acceder a sistemas de forma remota sin necesidad de autenticación, operando de manera oculta y dificultando su detección.

Las bombas lógicas representan otro tipo de amenaza relevante, ya que permanecen inactivas hasta que se cumple una condición específica, momento en el cual ejecutan acciones dañinas, como la eliminación de archivos o la alteración de sistemas. Los caballos de Troya, en cambio, se presentan como programas legítimos con el fin de engañar al usuario y facilitar la instalación de software malicioso. Este tipo de amenaza suele estar oculto en archivos aparentemente inofensivos, como imágenes, documentos o aplicaciones.

Asimismo, los exploits son herramientas diseñadas para aprovechar vulnerabilidades específicas en sistemas o aplicaciones, permitiendo la ejecución de código malicioso o el acceso no autorizado. Los gusanos o worms destacan por su capacidad de propagarse automáticamente a través de redes, replicándose sin necesidad de intervención del usuario, lo que los convierte en una amenaza altamente peligrosa en entornos interconectados.

En el ámbito de los ataques basados en el engaño, técnicas como el phishing y el pharming son especialmente relevantes. El phishing consiste en la suplantación de identidad mediante correos electrónicos o mensajes fraudulentos, mientras que el pharming redirige a los usuarios a sitios web falsos mediante la manipulación de sistemas de resolución de nombres. Por otro lado, el spam, aunque en apariencia menos dañino, puede actuar como medio de propagación de malware o fraudes.

El spyware se caracteriza por recopilar información sin el conocimiento del usuario, mientras que los virus informáticos requieren la intervención del usuario para activarse y replicarse en otros archivos. En los últimos años, el ransomware ha adquirido gran relevancia debido a su impacto económico, ya que cifra la información de las víctimas y exige un rescate para su recuperación. De igual forma, el scareware utiliza tácticas de intimidación para engañar a los usuarios, y los rootkits permiten ocultar la presencia de otros programas maliciosos, dificultando su detección.

Por otro lado, las amenazas físicas están relacionadas con el entorno donde se ubican los sistemas de información. Factores como incendios, inundaciones, terremotos o fallos eléctricos pueden comprometer gravemente la disponibilidad de los sistemas. Asimismo, elementos como interferencias electromagnéticas o deficiencias en la infraestructura pueden afectar el correcto funcionamiento de los equipos. Estas amenazas ponen de manifiesto la necesidad de implementar medidas de seguridad física y planes de contingencia adecuados.

Para comprender de manera integral las amenazas, es necesario analizar sus componentes. En primer lugar, se encuentran los agentes de la amenaza, es decir, los individuos o grupos que llevan a cabo los ataques. En segundo lugar, la capacidad del agente determina el nivel de sofisticación del ataque, incluyendo recursos, conocimientos y herramientas disponibles. Además, existen factores que influyen en el comportamiento de los atacantes, como los inhibidores, que actúan

como elementos disuasorios, y los amplificadores, que incentivan la actividad maliciosa.

Asimismo, los catalizadores representan circunstancias que pueden desencadenar un ataque, como cambios tecnológicos o situaciones específicas, mientras que los motivadores del agente incluyen factores políticos, económicos, ideológicos o personales. Estos elementos permiten comprender no solo cómo se producen las amenazas, sino también por qué ocurren.

La clasificación y tipología de amenazas en ciberseguridad evidencian la complejidad del entorno digital actual, caracterizado por la interacción de factores humanos, tecnológicos y físicos. El conocimiento detallado de estas amenazas resulta esencial para el desarrollo de estrategias de protección eficaces, orientadas a prevenir incidentes, reducir riesgos y garantizar la seguridad de los sistemas de información en un contexto de constante evolución.

1.4. Ciclo de vida y técnicas de los ataques informáticos

Los ciberataques, lejos de ser eventos aislados o aleatorios, suelen desarrollarse siguiendo un proceso estructurado y progresivo que permite a los atacantes maximizar sus probabilidades de éxito. Este proceso, conocido como ciclo de vida del ataque, comprende una serie de fases interrelacionadas que abarcan desde la recopilación inicial de información hasta la ejecución del ataque y la ocultación de evidencias. Comprender estas etapas resulta fundamental para el análisis de la seguridad en los sistemas de información, ya que permite identificar los puntos críticos en los que pueden implementarse mecanismos de detección y defensa. En este sentido, diversos estudios han profundizado en la naturaleza multietapa de los ciberataques, destacando que estos no solo siguen una secuencia lógica, sino que también presentan interdependencias entre fases. Ferrer-Oliva et al. (2025) evidencian, a partir de análisis empíricos, que los incidentes de ciberseguridad suelen desarrollarse como procesos encadenados, en los que

cada fase condiciona la siguiente, lo que refuerza la necesidad de abordar la defensa desde una perspectiva integral y no fragmentada.

En esta misma línea, Grigaliūnas et al. (2023) proponen métodos específicos para identificar el alcance y el impacto de cada fase del ataque, subrayando que no todas las etapas tienen el mismo peso en términos de riesgo. Según estos autores, la correcta identificación de las fases más críticas permite priorizar los esfuerzos de seguridad y optimizar los recursos destinados a la protección de los sistemas. Este enfoque resulta especialmente relevante en entornos complejos, donde la detección temprana en fases iniciales, como el reconocimiento o el escaneo, puede prevenir la materialización completa del ataque.

Por otro lado, el análisis del ciclo de vida del ataque no solo tiene implicaciones técnicas, sino también estratégicas y organizativas. Toivanen y Luoma-aho (2026) destacan que comprender la progresión de los ataques resulta clave para desarrollar estrategias de comunicación y gestión de crisis más eficaces, especialmente en contextos donde los incidentes de ciberseguridad tienen un alto impacto reputacional. Estos autores plantean que cada fase del ataque requiere una respuesta específica, no solo desde el punto de vista técnico, sino también en términos de coordinación organizativa y comunicación con los distintos actores implicados.

Asimismo, la evolución de las amenazas en el tiempo ha influido en la complejidad del ciclo de vida de los ataques. Dhanaraj (2025) señala que el avance de tecnologías como la inteligencia artificial ha permitido el desarrollo de ataques más sofisticados, capaces de adaptarse dinámicamente a los sistemas de defensa. Esta evolución ha ampliado el ciclo tradicional del ataque, incorporando nuevas fases o refinando las existentes, lo que dificulta su detección y mitigación. En este contexto, los atacantes no solo ejecutan acciones lineales, sino que pueden iterar entre fases, ajustando sus estrategias en función de las respuestas del sistema.

Por otra parte, Oppenheimer (2024) introduce una perspectiva crítica al señalar que la forma en que se detectan y analizan los ciberataques puede influir en la comprensión de su ciclo de vida. Según este autor, existe un sesgo en la identificación de las fases, ya que muchas veces solo se observan las etapas finales del ataque, cuando el daño ya ha sido causado, lo que limita la capacidad de aprendizaje y prevención. Esta reflexión pone de manifiesto la importancia de mejorar los sistemas de monitoreo y detección para obtener una visión más completa del proceso de ataque.

Finalmente, Ewoh y Vartiainen (2024) destacan la relevancia de considerar factores sociotécnicos en el análisis del ciclo de vida de los ataques, especialmente en sectores críticos como el sanitario. Estos autores subrayan que las vulnerabilidades no solo se encuentran en la tecnología, sino también en la interacción entre personas, procesos y sistemas, lo que influye directamente en la forma en que se desarrollan y evolucionan los ataques. En consecuencia, la comprensión del ciclo de vida del ataque debe integrar tanto aspectos técnicos como organizativos y humanos.

Estas investigaciones muestran que el ciclo de vida de los ciberataques es un proceso dinámico, complejo y multidimensional, cuya comprensión resulta esencial para el diseño de estrategias de defensa eficaces. El análisis detallado de sus fases no solo permite identificar puntos de intervención, sino también anticipar la evolución de las amenazas en un entorno digital en constante transformación.

En este sentido, el estudio del ciclo de vida de los ataques no solo facilita la comprensión del comportamiento de los actores maliciosos, sino que también contribuye al diseño de estrategias de protección más eficaces y proactivas en cada una de sus fases (Figura 1.5).

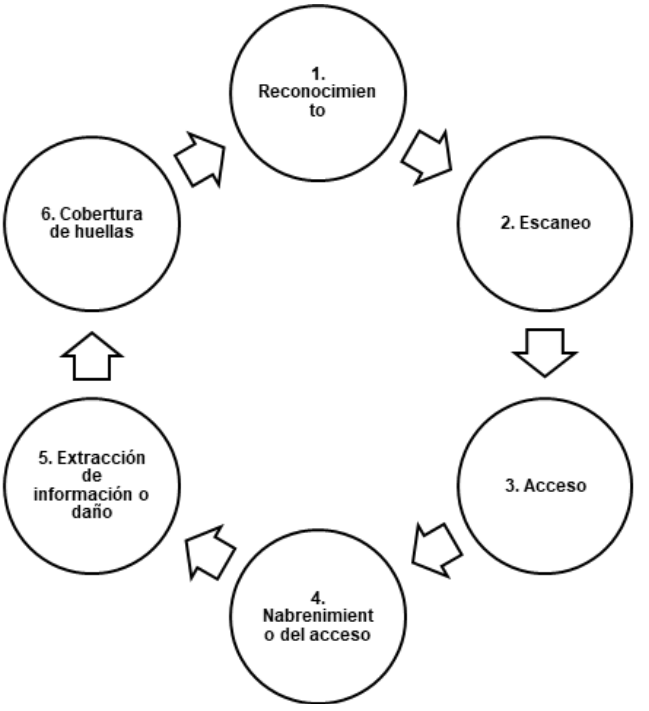


Figura 1.5. Ciclo de vida del ataque.

En el ámbito de la ciberseguridad, resulta fundamental comprender no solo las técnicas empleadas en los ataques, sino también los motivos que impulsan a los ciberdelincuentes a llevarlos a cabo, así como los efectos que estos generan y los métodos utilizados para infiltrarse en los sistemas. El análisis de estos elementos permite interpretar el comportamiento de los actores maliciosos, anticipar posibles escenarios de riesgo y diseñar estrategias de defensa más eficaces. En este contexto, los ciberataques no deben entenderse como eventos aislados, sino como fenómenos complejos que responden a múltiples factores psicológicos, económicos, tecnológicos y sociales.

Uno de los aspectos clave para comprender la naturaleza de los ciberataques son los motivos del ciberdelincuente. Estos pueden ser muy diversos y, en muchos casos, combinan diferentes intereses y objetivos. Entre ellos,

la autoconfianza constituye un factor relevante, ya que algunos atacantes buscan reforzar su autoestima demostrando sus habilidades técnicas. Este tipo de motivación suele estar presente en atacantes con menor nivel de experiencia, que realizan acciones poco agresivas con el objetivo de ganar reconocimiento o validación dentro de comunidades digitales.

Por otro lado, el poder asertivo representa una motivación más compleja, en la que el atacante busca ejercer control sobre sistemas o personas. En este caso, el objetivo no siempre es causar daño directo, sino demostrar superioridad o capacidad de dominio, lo que puede derivar en ataques más elaborados y persistentes. En contraste, la ira o represalia se manifiesta cuando el atacante actúa impulsado por sentimientos de frustración o injusticia, dirigiendo sus acciones contra una organización, institución o individuo específico. Este tipo de motivación suele estar asociado a ataques internos, como los realizados por exempleados descontentos.

Asimismo, el sadismo constituye una motivación en la que el ciberdelincuente obtiene satisfacción del daño causado, disfrutando del impacto negativo generado en las víctimas. Este tipo de comportamiento, aunque menos frecuente, puede dar lugar a ataques especialmente destructivos. Sin embargo, en la actualidad, el motivo más predominante es el económico o lucrativo, ya que muchos ataques buscan obtener beneficios financieros mediante el robo de información, la extorsión o la explotación de recursos. Este fenómeno ha dado lugar a la profesionalización del cibercrimen, donde los ataques se planifican y ejecutan como parte de actividades delictivas organizadas.

En relación con los efectos de los ciberataques, estos pueden clasificarse en función del tipo de daño que provocan en los sistemas de información. La interrupción se refiere a la pérdida de disponibilidad de los servicios, impidiendo el acceso a los recursos por parte de los usuarios. Este tipo de daño puede tener consecuencias críticas en entornos donde la continuidad operativa es esencial, como en sistemas financieros o infraestructuras críticas. Por otro lado, la modificación implica la alteración

no autorizada de la información, afectando su integridad y comprometiendo su fiabilidad.

La intercepción constituye otro tipo de daño relevante, ya que implica el acceso no autorizado a datos, comprometiendo la confidencialidad de la información. Este tipo de ataque es especialmente crítico cuando se trata de datos sensibles, como información personal o financiera. Finalmente, la fabricación se refiere a la inserción de información falsa en los sistemas, lo que puede generar confusión, errores en la toma de decisiones o pérdida de confianza en los sistemas.

En cuanto a los métodos de infiltración en redes informáticas, la ingeniería social destaca como una de las técnicas más efectivas y ampliamente utilizadas. Este método se basa en la manipulación de las personas para obtener información confidencial o acceso a sistemas, explotando factores como la confianza, la curiosidad o el miedo. Entre sus variantes se encuentra el pretexto, que consiste en asumir una identidad falsa para obtener datos sensibles; la infiltración o tailgating, que implica acceder a áreas restringidas siguiendo a una persona autorizada; y el quid pro quo, donde se ofrece un beneficio a cambio de información. Estas técnicas evidencian que la seguridad no depende únicamente de la tecnología, sino también del comportamiento de los usuarios.

Por otro lado, los ataques de denegación de servicio (DoS) constituyen una de las amenazas más comunes en entornos digitales. Estos ataques tienen como objetivo interrumpir el funcionamiento de sistemas, redes o aplicaciones mediante la saturación de recursos. Esto puede lograrse mediante el envío masivo de tráfico o mediante el uso de paquetes maliciosos que provocan fallos en el sistema. Las consecuencias de estos ataques incluyen la pérdida de disponibilidad, la degradación del rendimiento y, en algunos casos, daños físicos en los equipos.

Una evolución de estos ataques es la denegación de servicio distribuido (DDoS), que se caracteriza por el uso de múltiples dispositivos comprometidos para lanzar ataques coordinados. Estos dispositivos forman

parte de una botnet, es decir, una red de equipos infectados que son controlados remotamente por el atacante. El proceso de creación de una botnet implica la infección de dispositivos mediante software malicioso, su incorporación a una red controlada y su utilización para llevar a cabo diversas actividades maliciosas. Este modelo ha dado lugar a la aparición de mercados ilegales donde se alquilan botnets, lo que ha facilitado el acceso a este tipo de ataques incluso a individuos con pocos conocimientos técnicos.

Otro método relevante es el ataque man-in-the-middle (MITM), que consiste en la interceptación de las comunicaciones entre dos partes sin que estas sean conscientes. Este tipo de ataque permite al atacante acceder a información sensible, modificar datos o suplantar la identidad de uno de los interlocutores. Estos ataques son especialmente frecuentes en redes inalámbricas abiertas, donde los usuarios suelen conectarse sin las medidas de seguridad adecuadas. Una variante de este ataque es el man in the mobile, que afecta a dispositivos móviles y permite la captura de información como códigos de autenticación.

En el ámbito de la manipulación de la información, el envenenamiento SEO constituye una técnica que consiste en alterar el posicionamiento de sitios web en los motores de búsqueda para aumentar la visibilidad de páginas maliciosas. Esta estrategia busca atraer a un mayor número de usuarios hacia sitios fraudulentos, incrementando la probabilidad de éxito de los ataques.

Finalmente, el descifrado de contraseñas representa uno de los métodos más utilizados para obtener acceso no autorizado a sistemas. Entre las técnicas más comunes se encuentran la pulverización de contraseñas, los ataques de diccionario, la fuerza bruta y los ataques arcoíris. Estas técnicas se complementan con la interceptación de tráfico, que permite capturar credenciales cuando estas se transmiten sin cifrado. La eficacia de estos métodos pone de manifiesto la importancia de implementar políticas de seguridad robustas, como el uso de contraseñas seguras y mecanismos de autenticación multifactor.

El análisis de los motivos, los daños y los métodos de infiltración permite comprender la complejidad del fenómeno de los ciberataques. La interacción de factores humanos, tecnológicos y organizativos configura un entorno dinámico en el que las amenazas evolucionan constantemente. Por ello, resulta imprescindible adoptar un enfoque integral de la ciberseguridad, que no solo contemple la implementación de soluciones tecnológicas, sino también la formación de los usuarios, la gestión del riesgo y la adaptación continua a nuevas amenazas.

1.5. Clasificación de ciberataques y amenazas persistentes avanzadas en función de la tríada de seguridad de la información

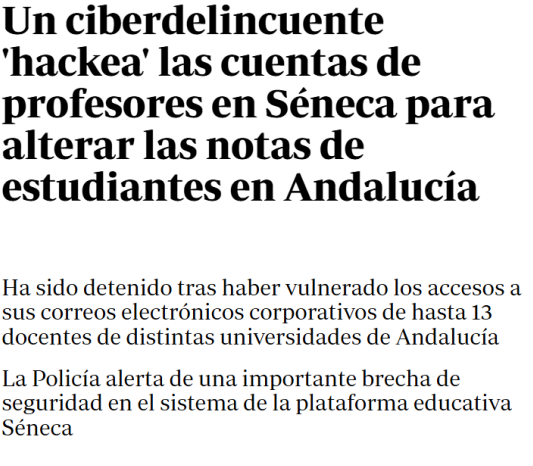
La clasificación de los ciberataques se fundamenta en los principios esenciales de la seguridad de la información, comúnmente conocidos como la tríada CIA (Confidencialidad, Integridad y Disponibilidad). Estos tres pilares constituyen la base sobre la cual se diseñan las políticas y mecanismos de protección en los sistemas informáticos. En este sentido, cada ciberataque puede afectar a uno o varios de estos principios, por lo que su análisis permite comprender mejor el alcance de los incidentes de seguridad, así como sus posibles consecuencias. Esta clasificación no solo facilita la identificación de los riesgos, sino que también contribuye a la planificación e implementación de medidas de prevención y mitigación más eficaces, adaptadas a la naturaleza de cada amenaza.

a) Ataques contra la integridad

La integridad de la información garantiza que los datos se mantengan completos, exactos y sin alteraciones no autorizadas a lo largo de su ciclo de vida. Un ataque dirigido contra este principio tiene como objetivo modificar, corromper o manipular la información, ya sea de forma directa o indirecta, con el fin de generar desconfianza en los sistemas o de obtener una ventaja indebida. Este tipo de ataques resulta especialmente crítico, ya que, aunque los sistemas continúen operativos, la pérdida de

fiabilidad en los datos puede afectar gravemente la toma de decisiones y la credibilidad de la organización.

Entre los ejemplos más comunes de ataques contra la integridad se encuentra la modificación de registros en sistemas informáticos, como la alteración de notas académicas en plataformas educativas o la manipulación de datos financieros en bases de datos empresariales. Asimismo, la inyección de código malicioso en aplicaciones web, como ocurre en los ataques de tipo SQL Injection, permite a los atacantes modificar o eliminar información almacenada en servidores, comprometiendo seriamente la consistencia de los datos. Este tipo de incidentes evidencia la importancia de implementar mecanismos de control de integridad, validación de entradas y auditorías periódicas que permitan detectar y prevenir alteraciones no autorizadas (Figura 1.6).



Un ciberdelincuente 'hackea' las cuentas de profesores en Séneca para alterar las notas de estudiantes en Andalucía

Ha sido detenido tras haber vulnerado los accesos a sus correos electrónicos corporativos de hasta 13 docentes de distintas universidades de Andalucía

La Policía alerta de una importante brecha de seguridad en el sistema de la plataforma educativa Séneca

Figura 1.6. Noticia sobre hackeo de cuentas de profesores.

Entre otros ejemplos relevantes de ataques contra la integridad se encuentra la alteración de registros financieros en bases de datos, donde los atacantes modifican transacciones, saldos o historiales contables con el objetivo de obtener beneficios económicos o encubrir actividades ilícitas. Este tipo de manipulación puede tener consecuencias graves, ya que afecta directamente a la veracidad de la información y puede comprometer la estabilidad financiera de una organización.

Asimismo, la inyección de código malicioso en aplicaciones web, como ocurre en los ataques de tipo SQL Injection, permite a los atacantes interactuar de manera indebida con las bases de datos, alterando, eliminando o insertando información sin autorización. Este tipo de ataque explota fallos en la validación de entradas y constituye una de las vulnerabilidades más críticas en entornos web.

El impacto de los ataques contra la integridad suele ser elevado, debido a que, aunque los sistemas permanezcan operativos, la información deja de ser confiable. Esto puede derivar en decisiones erróneas, pérdida de credibilidad institucional y, en casos más graves, en consecuencias legales o económicas significativas. Por ello, la protección de la integridad de los datos es esencial para garantizar la fiabilidad de los sistemas de información.

b) Ataques contra la privacidad (confidencialidad)

La privacidad o confidencialidad se relaciona con la protección de la información sensible frente a accesos no autorizados. Este principio garantiza que únicamente las personas o sistemas debidamente autorizados puedan acceder a determinados datos. Los ataques dirigidos contra la confidencialidad tienen como objetivo principal la obtención de información restringida, como datos personales, credenciales de acceso, información financiera o propiedad intelectual, lo que los convierte en una de las amenazas más frecuentes y peligrosas en el entorno digital actual.

Entre los ejemplos más comunes se encuentra el robo de credenciales mediante técnicas de phishing, en las que los atacantes se hacen pasar por entidades legítimas para engañar a los usuarios y obtener sus datos de acceso. Asimismo, la interceptación de comunicaciones, conocida como sniffing, permite capturar información transmitida a través de redes inseguras, especialmente en conexiones Wi-Fi abiertas o mal protegidas. Otro caso frecuente es la filtración de bases de datos que contienen información personal de usuarios, lo que puede derivar

en fraudes, robo de identidad o uso indebido de los datos.

Este tipo de ataques representa un riesgo significativo no solo desde el punto de vista técnico, sino también legal y reputacional para las organizaciones. La exposición de datos sensibles puede implicar sanciones por incumplimiento de normativas de protección de datos, como la Ley Orgánica de Protección de Datos Personales en Ecuador, así como una pérdida de confianza por parte de los usuarios y clientes. En consecuencia, la protección de la confidencialidad requiere la implementación de medidas como el cifrado de la información, el control de accesos y la concienciación de los usuarios.

c) Ataques contra la disponibilidad

La disponibilidad es un principio fundamental que garantiza que los sistemas, servicios y recursos estén accesibles para los usuarios autorizados en el momento en que los necesiten. Los ataques dirigidos contra este principio tienen como objetivo interrumpir, degradar o impedir el acceso a dichos recursos, afectando directamente la operatividad de las organizaciones.

Entre los ejemplos más comunes se encuentran los ataques de denegación de servicio (DoS) y su variante distribuida (DDoS), en los que se envía una gran cantidad de solicitudes a un sistema con el fin de saturarlo y provocar su caída. Asimismo, la saturación de servidores mediante tráfico masivo puede reducir significativamente el rendimiento de las aplicaciones, afectando la experiencia de los usuarios. Por otro lado, el uso de ransomware permite a los atacantes cifrar los archivos de una organización, impidiendo el acceso a la información hasta que se pague un rescate, lo que constituye una forma directa de afectar la disponibilidad.

En entornos como instituciones académicas o empresas, este tipo de ataques puede tener consecuencias especialmente graves, ya que puede paralizar servicios esenciales como plataformas de aprendizaje en línea, sistemas de correo electrónico o herramientas administrativas. La interrupción de estos servicios no solo afecta la productividad, sino que también puede generar pérdidas económicas y daños a la reputación institucional. Por ello, garantizar la disponibilidad de los

sistemas requiere la implementación de mecanismos de redundancia, planes de contingencia y sistemas de monitoreo que permitan detectar y mitigar estos ataques de manera oportuna (Figura 1.7).

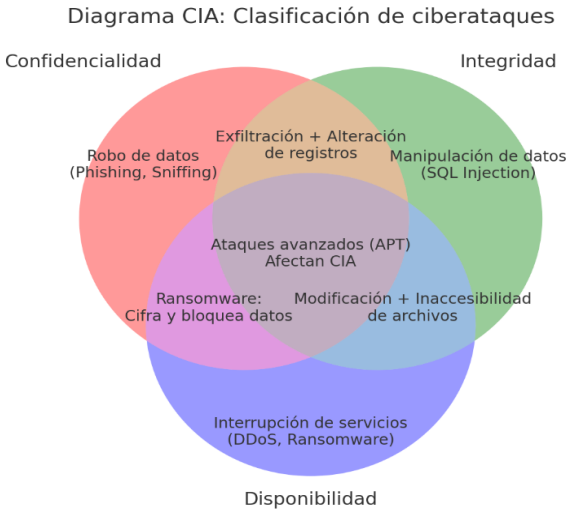


Figura 1.7. Diagrama gráfico de la tríada CIA.

Las Amenazas Persistentes Avanzadas (APT, por sus siglas en inglés *Advanced Persistent Threats*) representan una de las formas más sofisticadas, complejas y peligrosas de ataque en el ámbito de la ciberseguridad actual. A diferencia de otros ataques más directos, visibles o de ejecución inmediata, las APT se caracterizan por su alto nivel de planificación, su enfoque estratégico y su capacidad para permanecer ocultas dentro de los sistemas comprometidos durante largos periodos de tiempo. Estas particularidades las convierten en una amenaza crítica no solo para gobiernos y grandes corporaciones, sino también para instituciones educativas, organismos públicos y cualquier organización que gestione información sensible o de alto valor.

En este contexto, las APT no deben entenderse únicamente como incidentes técnicos, sino como operaciones complejas que combinan habilidades avanzadas, recursos significativos y objetivos claramente definidos. En esta línea, Salim et al. (2023) destacan que las APT requieren enfoques avanzados de detección basados

en la conciencia situacional cibernética, debido a su capacidad para adaptarse dinámicamente al entorno y evadir los sistemas tradicionales de seguridad. Según estos autores, la detección efectiva de este tipo de amenazas implica analizar patrones de comportamiento a lo largo del tiempo, en lugar de depender únicamente de firmas o eventos aislados.

Asimismo, el carácter sofisticado de las APT ha impulsado el desarrollo de técnicas avanzadas de detección basadas en inteligencia artificial y aprendizaje automático. Chu et al. (2019) señalan que el uso de modelos como las máquinas de soporte vectorial permite identificar patrones complejos en el comportamiento de los ataques, facilitando la detección temprana de amenazas persistentes que de otro modo pasarían desapercibidas. Este enfoque resulta especialmente relevante en entornos donde los atacantes emplean técnicas de evasión avanzadas, lo que dificulta su identificación mediante métodos tradicionales.

Desde una perspectiva más amplia, el análisis de las APT también debe considerar su impacto en los principios fundamentales de la seguridad de la información. En este sentido, Jones et al. (2025) proponen una taxonomía basada en la tríada CIA que permite clasificar los ataques en función de su impacto sobre la confidencialidad, integridad y disponibilidad. Esta aproximación resulta útil para comprender cómo las APT pueden afectar simultáneamente a múltiples dimensiones de la seguridad, lo que incrementa su peligrosidad y complejidad.

Por otro lado, la evaluación de este tipo de amenazas también presenta desafíos en términos de análisis experto. Shoufan y Damiani (2017) destacan la importancia de la fiabilidad en la evaluación de riesgos en ciberseguridad, señalando que la interpretación de incidentes complejos como las APT puede variar significativamente entre expertos. Esta variabilidad pone de manifiesto la necesidad de desarrollar metodologías estandarizadas que permitan mejorar la consistencia en la identificación y análisis de este tipo de ataques.

En entornos críticos, como las infraestructuras inteligentes o sistemas industriales, las APT adquieren una dimensión aún más preocupante. Bouramdane (2023) subraya que en sistemas como las redes eléctricas inteligentes, este tipo de ataques puede comprometer no solo la información, sino también el funcionamiento físico de los sistemas, generando impactos a gran escala. De manera similar, Alanazi et al. (2023) destacan que las vulnerabilidades en sistemas SCADA pueden ser explotadas mediante APT para acceder a infraestructuras críticas, lo que evidencia la necesidad de reforzar las medidas de seguridad en estos entornos.

Finalmente, desde el punto de vista de la respuesta a incidentes, Gilbert et al. (2025) señalan que la gestión de las APT requiere estrategias específicas que integren detección temprana, respuesta coordinada y recuperación efectiva. Estos autores enfatizan que, debido a la naturaleza persistente de estos ataques, no basta con eliminar la amenaza inicial, sino que es necesario implementar mecanismos que impidan su reaparición y fortalezcan la resiliencia de los sistemas.

Una amenaza persistente avanzada puede definirse como un ataque dirigido y prolongado en el tiempo contra una organización específica, en el que los atacantes buscan infiltrarse de manera sigilosa en los sistemas, evadir los mecanismos de detección tradicionales y mantener un acceso continuo con el fin de extraer información valiosa o llevar a cabo acciones de espionaje, sabotaje o manipulación estratégica.

A diferencia de otros tipos de ataques cuyo propósito principal es causar una interrupción inmediata del servicio, las APT se centran en la permanencia y el acceso sostenido, lo que permite a los atacantes operar de forma encubierta, recopilando información de manera progresiva y evitando levantar sospechas. Este tipo de amenazas suele estar asociado a actores altamente especializados, como grupos organizados de ciberdelincuencia o entidades respaldadas por Estados, lo que incrementa su nivel de sofisticación y peligrosidad.

Entre las características que definen a las APT destaca, en primer lugar, la persistencia, que se manifiesta en la capacidad de los atacantes para mantener el acceso a los sistemas durante semanas, meses o incluso años. Esta permanencia prolongada les permite analizar en profundidad el entorno de la organización, identificar activos críticos y planificar acciones de manera estratégica. En segundo lugar, el avance gradual del ataque implica que este no se produce de forma repentina, sino que sigue una secuencia de fases cuidadosamente estructuradas, en las que cada etapa cumple una función específica dentro del objetivo global. Asimismo, estas amenazas se caracterizan por el uso de técnicas avanzadas, como exploits de día cero, malware desarrollado específicamente para el objetivo, mecanismos de cifrado de comunicaciones y sofisticadas estrategias de ingeniería social que dificultan su detección.

Otra característica fundamental es su orientación hacia objetivos estratégicos, ya que las APT no se dirigen de manera indiscriminada, sino que seleccionan cuidadosamente a sus víctimas en función del valor de la información que poseen. En este sentido, suelen enfocarse en sectores como el gubernamental, financiero, tecnológico o educativo, donde el acceso a datos sensibles puede generar ventajas económicas, políticas o estratégicas. A ello se suma la discreción con la que operan, ya que están diseñadas para actuar en segundo plano, evitando generar alertas en los sistemas de seguridad convencionales y prolongando así su permanencia dentro de la infraestructura comprometida.

El desarrollo de una APT responde a un proceso estructurado que se inicia con la fase de reconocimiento, en la cual los atacantes recopilan información detallada sobre la organización objetivo, incluyendo su infraestructura tecnológica, sus sistemas de seguridad y el comportamiento de sus usuarios. Esta fase resulta clave, ya que permite diseñar un ataque adaptado a las características específicas del entorno. Posteriormente, se produce la intrusión inicial, que suele llevarse a cabo mediante técnicas como el phishing, la explotación de

vulnerabilidades de software o la instalación de malware. Una vez dentro del sistema, los atacantes buscan escalar privilegios, obteniendo permisos de mayor nivel que les permitan moverse con mayor libertad dentro de la red y acceder a recursos más sensibles.

A continuación, se desarrolla el movimiento lateral, mediante el cual los atacantes se desplazan dentro de la red de la organización para comprometer otros sistemas y ampliar su acceso. Esta fase es especialmente crítica, ya que permite identificar y alcanzar los activos de mayor valor. Posteriormente, se lleva a cabo la exfiltración de datos, que consiste en la extracción de información de manera sigilosa, evitando ser detectados por los sistemas de monitoreo. Finalmente, los atacantes establecen mecanismos de persistencia, como la instalación de puertas traseras, que les permiten mantener el acceso al sistema incluso si se detecta y neutraliza parte del ataque.

A lo largo de los últimos años, se han registrado diversos ejemplos de APT que han tenido un impacto significativo a nivel global. Entre ellos destaca el caso de Stuxnet, un gusano informático diseñado para sabotear instalaciones nucleares, considerado uno de los primeros ataques dirigidos contra infraestructuras críticas. Asimismo, grupos como APT28 y APT29 han sido vinculados a actividades de ciberespionaje, dirigidas contra instituciones gubernamentales y organizaciones estratégicas. Estos casos evidencian el alto nivel de sofisticación de este tipo de amenazas y su capacidad para generar consecuencias de gran alcance.

En el ámbito educativo y en el sector público, las APT representan un riesgo creciente debido a la naturaleza de la información que gestionan estas organizaciones. Las instituciones educativas, por ejemplo, almacenan datos personales de estudiantes y docentes, información administrativa y resultados de investigaciones científicas que pueden tener un alto valor estratégico. De igual forma, las entidades públicas manejan información sensible relacionada con la gestión gubernamental y los servicios ciudadanos. Un ataque de este tipo puede derivar en el robo de identidad, la filtración de información

confidencial, el chantaje o el deterioro de la reputación institucional, afectando tanto a la organización como a los individuos involucrados.

Frente a este tipo de amenazas, resulta imprescindible adoptar un enfoque integral de ciberseguridad que combine medidas técnicas, organizativas y humanas. La implementación de estrategias de defensa en profundidad, que integren múltiples capas de protección como firewalls, sistemas de detección y prevención de intrusiones, autenticación multifactor y cifrado de la información, constituye una de las principales líneas de defensa. Asimismo, el monitoreo constante de los sistemas permite identificar comportamientos anómalos y detectar posibles intrusiones en etapas tempranas. La capacitación y concienciación de los usuarios es igualmente fundamental, ya que muchas de estas amenazas se inician mediante técnicas de ingeniería social. Por otro lado, la actualización periódica de los sistemas y la aplicación de parches de seguridad permiten reducir las vulnerabilidades que pueden ser explotadas por los atacantes.

Las amenazas persistentes avanzadas representan un desafío significativo para la ciberseguridad moderna, debido a su capacidad de adaptación, su complejidad y su impacto potencial. Su análisis y comprensión resultan esenciales para el diseño de estrategias de protección eficaces en un entorno digital cada vez más interconectado y dinámico, donde la prevención, la detección temprana y la respuesta coordinada constituyen elementos clave para garantizar la seguridad de los sistemas de información.

02.

Gobernanza, normativa y gestión del riesgo en seguridad de la información

2.1. Seguridad de datos y fundamentos de la criptografía aplicada

La información constituye uno de los activos más valiosos en el mundo moderno, ya que sustenta la toma de decisiones, la operación de sistemas críticos y el desarrollo de actividades económicas, sociales y gubernamentales. En la actualidad, empresas, instituciones públicas y usuarios particulares generan, procesan y almacenan grandes volúmenes de datos de manera constante, tales como registros financieros, credenciales de acceso, contratos digitales, correos electrónicos, historiales médicos y configuraciones de sistemas. En este contexto, la seguridad de los datos se convierte en un elemento esencial, cuyo objetivo principal es proteger la información frente a accesos no autorizados, pérdidas, filtraciones o alteraciones indebidas,

garantizando al mismo tiempo su disponibilidad para los usuarios legítimos.

Un dato que no cuenta con mecanismos adecuados de protección puede convertirse rápidamente en una vulnerabilidad crítica dentro de un sistema. Por ejemplo, el almacenamiento de contraseñas en texto plano en un equipo de cómputo representa un riesgo significativo, ya que, en caso de que el sistema sea comprometido, un atacante podría acceder fácilmente a dicha información y utilizarla para escalar privilegios o acceder a otros sistemas. Este tipo de situaciones evidencia la importancia de implementar medidas de seguridad robustas, como el cifrado de datos, la correcta gestión de credenciales y la aplicación de políticas de control de acceso. En este sentido, la seguridad de los datos no solo se limita a la protección técnica, sino que también implica una adecuada gestión del riesgo y la adopción de buenas prácticas organizacionales.

Desde una perspectiva más avanzada, Roman (2023) destaca que la protección de los datos no solo debe centrarse en evitar accesos no autorizados, sino también en preservar el equilibrio entre privacidad y utilidad de la información, especialmente en entornos donde los datos son utilizados para análisis y toma de decisiones. Esto implica que los mecanismos de protección, como los algoritmos de perturbación o anonimización, deben diseñarse cuidadosamente para evitar la exposición de información sensible sin comprometer su valor funcional. De manera complementaria, Tariq et al. (2023) señalan que en entornos altamente interconectados como el Internet de las Cosas (IoT), la falta de protección adecuada de los datos puede amplificar las vulnerabilidades, ya que los dispositivos suelen tener limitaciones de seguridad y pueden ser explotados como puntos de entrada para ataques más complejos.

En este contexto, el uso de técnicas criptográficas adquiere un papel fundamental. Dizon y Meehan (2024) subrayan que los principios técnicos de la encriptación no solo son esenciales para proteger la confidencialidad de los datos, sino que también tienen implicaciones en la regulación tecnológica, ya que determinan cómo se

gestionan los derechos de privacidad y el acceso a la información. Por su parte, Urooj et al. (2023) enfatizan que en sistemas distribuidos, como las redes de sensores inalámbricos, la seguridad criptográfica es indispensable para garantizar la transmisión confiable de datos, evitando alteraciones y accesos indebidos que puedan comprometer la integridad del sistema.

Asimismo, la seguridad de los datos debe abordarse desde una perspectiva integral que incluya factores humanos y organizacionales. Smith et al. (2021) destacan que las políticas públicas y organizacionales en ciberseguridad deben considerar los valores de los usuarios, ya que el comportamiento humano puede influir significativamente en la efectividad de las medidas de protección. Esto refuerza la idea de que la seguridad no depende únicamente de la tecnología, sino también de la concienciación, formación y cumplimiento de buenas prácticas por parte de los usuarios y administradores.

Estas aportaciones evidencian que la protección de los datos requiere un enfoque multidimensional que combine técnicas criptográficas avanzadas, gestión de riesgos, regulación adecuada y formación de los usuarios. Solo a través de esta integración es posible mitigar las vulnerabilidades y garantizar la seguridad de la información en entornos digitales cada vez más complejos y dinámicos.

Los principios fundamentales de la seguridad de la información se basan en la denominada tríada CIA, que constituye el marco conceptual sobre el cual se desarrollan las estrategias de protección. En primer lugar, la confidencialidad establece que la información solo debe ser accesible para aquellas personas o sistemas que cuenten con la debida autorización, evitando así la divulgación indebida de datos sensibles. En segundo lugar, la integridad garantiza que los datos se mantengan exactos, completos y libres de modificaciones no autorizadas, lo cual resulta esencial para asegurar la confiabilidad de la información. Finalmente, la disponibilidad asegura que los datos y recursos estén accesibles en el momento en que los

usuarios legítimos los requieran, evitando interrupciones que puedan afectar la operación de los sistemas.

La aplicación de estos principios puede observarse en distintos escenarios prácticos. Por ejemplo, en la gestión de redes, un archivo de configuración de un router debe ser accesible únicamente por el administrador autorizado, lo que responde al principio de confidencialidad. Asimismo, cualquier modificación realizada en dicho archivo debe ser controlada y verificada para evitar cambios que comprometan el funcionamiento del sistema, lo que se relaciona con la integridad. Por último, es fundamental contar con mecanismos de respaldo y recuperación que permitan restaurar la configuración en caso de fallos o incidentes, garantizando así la disponibilidad del sistema.

Un caso representativo que evidencia la importancia de la seguridad de los datos es el ataque de ransomware conocido como WannaCry, ocurrido en el año 2017. Este incidente afectó a más de 150 países y comprometió millones de equipos al cifrar sus archivos, impidiendo el acceso a la información hasta que se pagara un rescate. Entre los sectores afectados se encontraban hospitales, empresas y sistemas de transporte, lo que generó graves consecuencias a nivel operativo y económico. Este caso puso de manifiesto cómo una vulnerabilidad no corregida puede ser explotada a gran escala, afectando simultáneamente a la confidencialidad, integridad y disponibilidad de los datos.

En este contexto, la criptografía desempeña un papel fundamental en la protección de la información. La criptografía puede definirse como la disciplina encargada de desarrollar técnicas para proteger la información mediante la transformación de los datos en formatos que resulten incomprensibles para personas no autorizadas. Esta disciplina se compone de dos elementos principales. Por un lado, el cifrado, que consiste en ocultar el contenido de los mensajes o datos mediante algoritmos matemáticos, de tal manera que solo puedan ser interpretados por quienes poseen la clave correspondiente. Por otro lado, la autenticación e integridad, que permiten verificar la identidad de los

usuarios y garantizar que la información no ha sido alterada durante su transmisión o almacenamiento.

La encriptación, también conocida como cifrado, es uno de los mecanismos más utilizados para proteger la información en entornos digitales. Este proceso transforma los datos en un formato ilegible, denominado texto cifrado, que solo puede ser revertido a su forma original mediante el uso de una clave específica. De esta manera, incluso si un atacante logra interceptar la información, no podrá interpretarla sin contar con la clave de descifrado. La encriptación se aplica en múltiples contextos, como la protección de comunicaciones en internet, el almacenamiento seguro de datos y la autenticación de usuarios, constituyendo una de las principales defensas frente a accesos no autorizados.

La seguridad de los datos representa un componente esencial dentro de la ciberseguridad, ya que permite proteger uno de los recursos más críticos de la sociedad digital. Su adecuada implementación requiere no solo el uso de tecnologías avanzadas, sino también la adopción de políticas de seguridad, la formación de los usuarios y la evaluación constante de los riesgos. En un entorno donde las amenazas evolucionan de manera constante, garantizar la protección de la información se convierte en un desafío permanente que exige un enfoque integral y proactivo.

ENCRIPTAR = CIFRAR

Cuando se sustituyen las letras o los caracteres que conforman un mensaje original para crear uno distinto (Figura 2.1).

Figura 2.1. Proceso de encriptación.

Los objetivos principales de la encriptación se centran en garantizar la protección integral de la información en entornos digitales, especialmente frente a amenazas como el acceso no autorizado, la manipulación de datos y la suplantación de identidad. En este sentido, uno de

los objetivos fundamentales es la confidencialidad, que busca impedir que terceros puedan acceder o interpretar los datos sin autorización. Esto es especialmente importante en contextos donde se manejan datos sensibles, como información financiera, médica o credenciales de acceso.

Otro objetivo clave es la autenticidad, que permite verificar que la información proviene realmente de la fuente declarada, evitando así ataques de suplantación o falsificación. Este principio es esencial en sistemas de comunicación, donde es necesario garantizar que el emisor es quien dice ser. Asimismo, la integridad de la información constituye un pilar fundamental, ya que asegura que los datos no han sido alterados durante su transmisión o almacenamiento. Esto permite detectar cualquier modificación indebida que pueda comprometer la validez de la información.

Finalmente, el no repudio es un objetivo crítico en entornos digitales, especialmente en transacciones electrónicas, ya que impide que el emisor niegue haber enviado un mensaje o realizado una acción. Este principio se apoya en mecanismos criptográficos que permiten demostrar de manera verificable la autoría de la información, lo que resulta fundamental en procesos legales y comerciales.

La función principal de la encriptación consiste en transformar la información en un formato ilegible para cualquier persona que no posea la clave adecuada para su descifrado. De este modo, incluso si los datos son interceptados durante su transmisión o accedidos de forma indebida, no podrán ser comprendidos ni utilizados por terceros no autorizados. Esta capacidad de proteger la información frente a interceptaciones convierte a la encriptación en una de las herramientas más importantes en la seguridad de la información, especialmente en entornos como redes de comunicación, almacenamiento en la nube y sistemas distribuidos.

En cuanto a los tipos de encriptación, estos pueden clasificarse en tres grandes categorías, cada una con características, ventajas y aplicaciones específicas. El cifrado simétrico es uno de los métodos más utilizados

y se caracteriza por emplear una única clave tanto para cifrar como para descifrar la información. En este esquema, todas las partes involucradas comparten la misma clave secreta, lo que permite un proceso de cifrado rápido y eficiente. Un ejemplo ampliamente utilizado es el estándar AES (Advanced Encryption Standard), que se aplica en múltiples sistemas de seguridad. Su principal ventaja radica en su alto rendimiento, lo que lo hace ideal para el manejo de grandes volúmenes de datos, como el cifrado de discos duros en dispositivos corporativos. Sin embargo, su principal desventaja es la necesidad de compartir la clave entre las partes, lo que puede generar riesgos durante su transmisión si no se emplean canales seguros.

Por otro lado, el cifrado asimétrico introduce un modelo más avanzado al utilizar dos claves diferentes: una clave pública, que puede ser compartida libremente, y una clave privada, que se mantiene en secreto. En este esquema, la información cifrada con la clave pública solo puede ser descifrada con la clave privada correspondiente. Este modelo resuelve el problema del intercambio seguro de claves, ya que no es necesario compartir información sensible previamente. Entre los algoritmos más conocidos se encuentran RSA y ECC (Elliptic Curve Cryptography). No obstante, este tipo de cifrado es más lento en comparación con el simétrico, por lo que suele utilizarse en combinación con este último en sistemas híbridos. Su uso es fundamental en procesos de autenticación y en protocolos de seguridad como HTTPS o SSH, donde se garantiza la identidad de las partes involucradas y la protección de la comunicación.

Finalmente, el hashing representa un enfoque diferente dentro de la seguridad criptográfica. A diferencia del cifrado, el hashing es un proceso unidireccional, lo que significa que no es posible revertir el resultado para obtener el dato original. Este proceso transforma la información en un valor único de longitud fija, conocido como hash, que actúa como una huella digital del dato original. Algoritmos como SHA-256 o MD5 son ejemplos comunes de esta técnica. El hashing se utiliza principalmente para verificar la integridad de la

información y proteger contraseñas, ya que permite comprobar si un dato ha sido modificado sin necesidad de conocer su contenido original. Un caso práctico de uso es la verificación de archivos descargados desde internet, donde el proveedor publica el hash correspondiente para asegurar que el archivo no ha sido alterado durante su distribución.

La Tabla 2.1 presenta un análisis comparativo claro de los principales métodos de encriptación, destacando sus diferencias en términos de uso de claves, velocidad, nivel de seguridad y aplicaciones. En primer lugar, el cifrado simétrico, representado por AES, se caracteriza por utilizar una única clave y ofrecer una alta velocidad, lo que lo convierte en una opción eficiente para el manejo de grandes volúmenes de datos, como en discos duros o copias de seguridad. Sin embargo, su principal limitación radica en la necesidad de compartir la clave de forma segura.

Por otro lado, el cifrado asimétrico, como RSA, emplea un par de claves (pública y privada), lo que incrementa significativamente la seguridad al eliminar el problema del intercambio de claves. Aunque su velocidad es menor en comparación con el cifrado simétrico, su alto nivel de seguridad lo hace ideal para aplicaciones críticas como protocolos SSL/TLS y firmas digitales.

Finalmente, el hashing, ejemplificado por SHA-256, no utiliza claves y destaca por su muy alta velocidad. Su función no es cifrar información reversible, sino garantizar la integridad de los datos, siendo ampliamente utilizado en la validación de archivos y almacenamiento seguro de contraseñas.

Tabla 2.1. Cuadro comparativo de tipos de encriptación.

Método	Clave usada	Velocidad	Seguridad	Aplicaciones comunes
Simétrico (AES)	Una sola	Alta	Alta	VPNs, discos duros, backups
Asimétrico (RSA)	Par pública/privada	Media	Muy alta	SSL/TLS, firmas digitales

Método	Clave usada	Velocidad	Seguridad	Aplicaciones comunes
Hashing (SHA-256)	No aplica	Muy alta	Alta	Validación de integridad, contraseñas

Los algoritmos de cifrado constituyen el núcleo de la criptografía moderna, ya que definen el procedimiento mediante el cual los datos originales son transformados en texto cifrado. De manera simplificada, pueden entenderse como una "receta matemática" que reorganiza y altera la información de tal forma que resulta incomprensible para cualquier persona que no disponga de la clave adecuada. Esta transformación no es arbitraria, sino que sigue reglas estrictas basadas en principios matemáticos complejos que garantizan la seguridad del proceso.

Entre los algoritmos de encriptación más utilizados se encuentran, en el ámbito del cifrado simétrico, AES, 3-DES y SNOW, mientras que en el cifrado asimétrico destacan RSA y la criptografía de curva elíptica (ECC). Cada uno de estos algoritmos ha sido diseñado para cumplir con estándares de seguridad específicos y responder a diferentes necesidades en cuanto a velocidad, nivel de protección y consumo de recursos.

En el caso de los algoritmos asimétricos, como RSA, su funcionamiento se basa en problemas matemáticos complejos, como la factorización de números primos de gran tamaño. En este esquema, dos números primos muy grandes son multiplicados para generar una clave pública, mientras que su factorización resulta extremadamente difícil sin conocer los valores originales. Esta dificultad computacional es lo que garantiza la seguridad del sistema, ya que descifrar una clave RSA mediante fuerza bruta requeriría una cantidad de tiempo y recursos prácticamente inalcanzable con la tecnología actual. Por su parte, la criptografía de curva elíptica ofrece niveles de seguridad similares con claves más pequeñas, lo que la hace más eficiente en dispositivos con recursos limitados.

Es importante destacar que los algoritmos de cifrado son públicos y están estandarizados, lo que permite su

análisis y validación por parte de la comunidad científica. Esto no representa una debilidad, ya que la seguridad del sistema no depende del secreto del algoritmo, sino de la confidencialidad de la clave utilizada. La clave, que suele ser una cadena de caracteres o números, es el elemento verdaderamente secreto y su correcta gestión resulta fundamental para garantizar la seguridad de los datos.

En cuanto a sus aplicaciones prácticas, la encriptación está presente en prácticamente todos los ámbitos tecnológicos. En el ámbito de las redes, se utiliza en protocolos como HTTPS para la navegación segura, VPNs para conexiones remotas protegidas y WPA3 para asegurar redes inalámbricas. En los sistemas operativos, herramientas como BitLocker en Windows o LUKS en Linux permiten cifrar discos completos, protegiendo la información incluso en caso de pérdida o robo del dispositivo. Asimismo, en el ámbito de las comunicaciones, tecnologías como PGP o GPG permiten cifrar correos electrónicos, garantizando la confidencialidad de los mensajes. En el almacenamiento en la nube, servicios como OneDrive o Google Drive implementan cifrado de extremo a extremo para proteger los datos almacenados y transmitidos.

Un ejemplo práctico se puede observar en el diseño de redes seguras, donde protocolos como IPsec cifran el tráfico entre dos puntos, asegurando que, aunque la información sea interceptada, no pueda ser interpretada por terceros. Este tipo de implementación resulta esencial en entornos corporativos, donde la protección de la información es crítica.

Para garantizar una adecuada seguridad de los datos, es fundamental seguir una serie de buenas prácticas. Entre ellas destaca el uso de algoritmos de cifrado reconocidos y actualizados, como AES-256 o RSA-2048, que ofrecen altos niveles de seguridad. Asimismo, es esencial implementar políticas de contraseñas robustas, realizar copias de seguridad cifradas y monitorear continuamente los intentos de acceso no autorizado. La aplicación periódica de parches y actualizaciones

también resulta clave para corregir vulnerabilidades y mantener la integridad de los sistemas.

Sin embargo, existen errores comunes que pueden comprometer la seguridad de los datos, incluso cuando se utilizan tecnologías avanzadas. Uno de los más frecuentes es el uso de algoritmos obsoletos, como DES o MD5, que han sido vulnerados y ya no ofrecen garantías de seguridad. Otro error crítico es reutilizar la misma clave de cifrado durante largos periodos de tiempo, lo que incrementa el riesgo de compromiso. Asimismo, la falta de protección de la clave privada en sistemas asimétricos o el almacenamiento de contraseñas en texto plano representan fallos graves que pueden facilitar el acceso no autorizado a los sistemas.

En entornos técnicos como el ensamblaje de equipos y la administración de redes, la encriptación desempeña un papel fundamental. En el ensamblaje de equipos, es habitual configurar discos con cifrado para proteger la información almacenada y garantizar la seguridad desde el nivel más básico del sistema. Además, durante la instalación de sistemas operativos, se implementan mecanismos de protección de credenciales que impiden accesos no autorizados. En el ámbito de redes, la configuración de VPNs permite establecer conexiones seguras para el acceso remoto, mientras que la implementación de estándares como WPA3 protege las redes inalámbricas frente a intrusiones. Asimismo, el uso de protocolos como SSH garantiza una administración remota segura de dispositivos como routers y switches.

2.2. Gestión segura de credenciales y llaves criptográficas

Las credenciales y las llaves criptográficas constituyen la primera línea de defensa en la protección de sistemas, redes y dispositivos, ya que permiten controlar el acceso y garantizar la identidad de los usuarios y de los recursos dentro de un entorno digital. Una credencial puede definirse como un conjunto de datos que permite autenticar a un usuario o dispositivo, como es el caso de un nombre de usuario y una contraseña. Por su parte, las llaves criptográficas (keys) son elementos

fundamentales dentro de los mecanismos de cifrado y autenticación, ya que permiten proteger la información mediante procesos matemáticos que garantizan su confidencialidad e integridad. En este sentido, la gestión de llaves criptográficas se ha convertido en un componente crítico dentro de las infraestructuras modernas, dado que su correcta administración influye directamente en la seguridad global de los sistemas. De acuerdo con Rana et al. (2023), los sistemas de gestión de claves deben garantizar no solo la generación y almacenamiento seguro de las mismas, sino también su distribución, rotación y revocación, lo que evidencia la complejidad de su manejo en entornos empresariales.

En el contexto del ensamblaje de equipos y el diseño de redes, el manejo seguro de credenciales y llaves resulta crítico, debido a que cualquier debilidad en su gestión puede comprometer la seguridad de toda la infraestructura tecnológica. Por ejemplo, un router configurado con credenciales por defecto puede ser fácilmente explotado por un atacante, permitiéndole acceder a la red y modificar su configuración. De igual manera, la exposición de una clave privada en un servidor puede facilitar accesos remotos no autorizados, comprometiendo la información almacenada y los servicios ofrecidos. Asimismo, el uso inadecuado de contraseñas administrativas, como su reutilización o compartición entre usuarios, puede generar vulnerabilidades que afecten a toda la red corporativa. En este contexto, Fehis et al. (2021) destacan que la gestión segura de claves debe apoyarse en políticas estrictas de control de acceso, como el modelo de seguridad "Chinese Wall", que limita el acceso a la información sensible en función de los roles y relaciones dentro de la organización.

Adicionalmente, el avance de las tecnologías emergentes ha incrementado la necesidad de implementar mecanismos más robustos para la protección de credenciales. Darmawan y Cahyono (2025) proponen el uso de cifrado de conocimiento cero (*zero-knowledge encryption*) en gestores de contraseñas, lo que permite que incluso los proveedores del servicio no tengan

acceso a las credenciales almacenadas, reforzando así la privacidad del usuario. De igual manera, Kumar y Goel (2025) señalan que la integración de técnicas de aprendizaje automático en sistemas de cifrado permite adaptar dinámicamente los mecanismos de seguridad frente a nuevas amenazas, especialmente en entornos distribuidos como el *fog computing.*

Por otro lado, en sistemas inteligentes y entornos en la nube, la protección de datos mediante técnicas criptográficas resulta esencial para evitar accesos no autorizados durante la transmisión y almacenamiento de la información. Qureshi et al. (2022) enfatizan que el uso de técnicas avanzadas de encriptación en sistemas conectados permite mitigar riesgos asociados a la externalización de datos, especialmente cuando estos son procesados fuera del entorno local. Asimismo, Kuzminykh et al. (2021) destacan que los sistemas modernos de gestión de claves criptográficas deben ser escalables, eficientes y capaces de integrarse con múltiples aplicaciones, lo que resulta fundamental en infraestructuras con un gran número de usuarios y servicios.

Estos estudios evidencian que la gestión de credenciales y llaves criptográficas no solo constituye un elemento técnico, sino también estratégico dentro de la ciberseguridad. Su correcta implementación requiere la integración de tecnologías avanzadas, políticas organizacionales sólidas y una gestión continua del riesgo, con el fin de garantizar la protección efectiva de los sistemas en entornos digitales cada vez más complejos.

Las credenciales pueden clasificarse en diferentes tipos según su naturaleza y nivel de seguridad. Las contraseñas tradicionales, que consisten en combinaciones de caracteres, siguen siendo el método más común de autenticación, aunque su efectividad depende en gran medida de su complejidad y gestión adecuada. Los PINs o códigos cortos se utilizan frecuentemente en dispositivos móviles o en sistemas básicos como el acceso al BIOS, ofreciendo un nivel de seguridad limitado. Por otro lado, los tokens de seguridad, ya sean dispositivos

físicos o aplicaciones, generan códigos dinámicos que incrementan significativamente la seguridad al añadir un factor adicional de autenticación.

La biometría, que incluye mecanismos como el reconocimiento de huellas dactilares, facial o de voz, representa una alternativa avanzada basada en características únicas del usuario, lo que dificulta su suplantación. Finalmente, los certificados digitales permiten validar identidades en redes y servicios mediante infraestructuras de clave pública (PKI), siendo ampliamente utilizados en entornos corporativos y en protocolos seguros de comunicación. Un ejemplo aplicado de este tipo de autenticación es el acceso de un administrador de red a un switch mediante el protocolo SSH utilizando un certificado digital en lugar de una contraseña, lo que incrementa significativamente el nivel de seguridad.

La gestión de contraseñas constituye un aspecto fundamental dentro de la seguridad de la información, ya que una mala práctica en este ámbito puede anular incluso los sistemas de protección más avanzados. Entre las buenas prácticas recomendadas se encuentra el uso de contraseñas con una longitud mínima de 12 caracteres, que combinen letras mayúsculas, minúsculas, números y símbolos, lo que incrementa su complejidad y dificulta ataques de fuerza bruta. Asimismo, es importante evitar el uso de palabras comunes, secuencias predecibles o información personal, ya que estos elementos son fácilmente identificables por los atacantes.

Otra medida esencial es el cambio periódico de contraseñas, especialmente en caso de incidentes de seguridad o sospechas de compromiso. Además, la implementación de autenticación multifactor (MFA) en accesos críticos añade una capa adicional de protección, combinando diferentes factores como algo que el usuario sabe (contraseña), algo que posee (token) o algo que es (biometría).

No obstante, existen errores comunes que comprometen la seguridad de las credenciales y que deben ser evitados. Entre ellos destaca el uso de contraseñas por

defecto, como “admin/admin”, que suelen ser conocidas públicamente y explotadas en ataques automatizados. Asimismo, la reutilización de contraseñas en múltiples servicios incrementa el riesgo de que una brecha en un sistema afecte a otros. La práctica de compartir credenciales entre usuarios también representa un riesgo significativo, ya que dificulta la trazabilidad de las acciones y aumenta la probabilidad de uso indebido.

Otro error crítico es el almacenamiento de contraseñas en texto plano o en medios inseguros, como correos electrónicos o archivos sin protección. En el ámbito técnico, es común encontrar casos en los que personal de soporte almacena credenciales en archivos de texto dentro de los equipos, lo que constituye una grave vulnerabilidad. La práctica recomendada en estos casos es el uso de gestores de contraseñas seguros, que permiten almacenar y administrar credenciales de manera cifrada, garantizando su protección frente a accesos no autorizados (Figura 2.2).

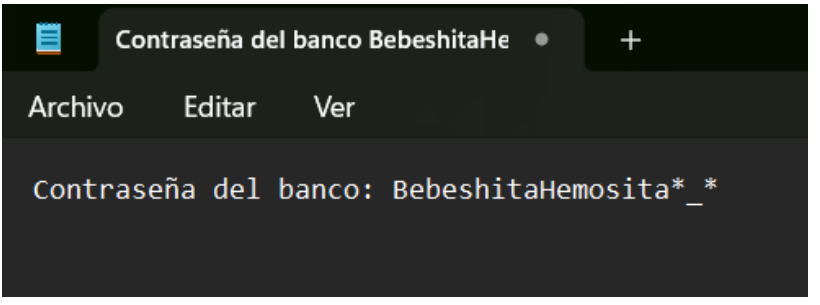


Figura 2.2. Almacenamiento de contraseña en texto plano.

Un caso representativo que evidencia las consecuencias de una gestión inadecuada de credenciales ocurrió en 2019, cuando un ataque masivo comprometió cámaras IP en distintos países. Este incidente se produjo debido a que numerosos dispositivos mantenían las contraseñas de fábrica, las cuales además se encontraban publicadas en manuales técnicos accesibles al público. Esta situación permitió a los atacantes acceder fácilmente a los dispositivos, demostrando cómo la falta de medidas básicas de seguridad puede derivar en vulnerabilidades críticas a gran escala.

En este contexto, las llaves criptográficas desempeñan un papel fundamental dentro de los sistemas de seguridad, ya que constituyen fragmentos de información utilizados por algoritmos criptográficos para cifrar, descifrar y autenticar datos. Su correcta gestión es esencial para garantizar la confidencialidad, integridad y autenticidad de la información en entornos digitales.

Las llaves pueden clasificarse en diferentes tipos según su función y estructura. En primer lugar, las llaves simétricas utilizan una única clave tanto para cifrar como para descifrar la información, como ocurre en algoritmos como AES-256. Este tipo de llaves destaca por su eficiencia y rapidez, siendo ampliamente utilizado en el cifrado de grandes volúmenes de datos. En segundo lugar, las llaves asimétricas funcionan mediante un par de claves: una pública, que puede compartirse libremente, y una privada, que debe mantenerse en secreto. Ejemplos de este tipo de criptografía incluyen RSA-2048 y ECC, que son fundamentales en procesos de autenticación y comunicaciones seguras. Finalmente, las llaves de sesión son claves temporales generadas para una comunicación específica, lo que permite reducir el riesgo de exposición al limitar su tiempo de uso.

La gestión adecuada de las llaves criptográficas requiere la implementación de buenas prácticas que garanticen su seguridad a lo largo de todo su ciclo de vida. En primer lugar, la generación de llaves debe realizarse en entornos confiables, como módulos de seguridad de hardware (HSM) o software especializado, que aseguren su aleatoriedad y fortaleza criptográfica. Asimismo, su almacenamiento debe efectuarse en ubicaciones seguras, evitando prácticas como el uso de texto plano o la inclusión en repositorios públicos, ya que esto podría facilitar su exposición. Otra medida clave es la rotación periódica de llaves, que consiste en renovarlas regularmente para reducir el impacto en caso de compromiso. Finalmente, es fundamental establecer mecanismos de revocación inmediata ante cualquier sospecha de vulneración, con el fin de impedir accesos no autorizados.

Un ejemplo práctico del uso de llaves criptográficas se observa en el protocolo HTTPS, donde el servidor utiliza su clave privada para autenticarse ante el cliente, mientras que este emplea la clave pública del servidor para cifrar la comunicación. Este proceso garantiza que la información transmitida permanezca protegida frente a posibles interceptaciones.

En entornos de redes y hardware, el manejo de llaves criptográficas es igualmente esencial. Las claves SSH permiten la administración segura de servidores, routers y switches, eliminando la necesidad de contraseñas tradicionales y reduciendo el riesgo de ataques. Por su parte, los certificados digitales son utilizados para autenticar identidades en redes privadas virtuales (VPN) y otros sistemas, asegurando la legitimidad de los participantes en la comunicación. Además, la rotación periódica de claves se presenta como una práctica fundamental para minimizar la exposición ante posibles ataques.

En la actualidad, la gestión de credenciales y llaves se apoya en tecnologías modernas que buscan mejorar la seguridad y la experiencia del usuario. El Single Sign-On (SSO) permite a los usuarios acceder a múltiples servicios mediante una única autenticación, reduciendo la necesidad de gestionar múltiples credenciales. Los sistemas de Identity and Access Management (IAM) ofrecen un control centralizado sobre los accesos a los recursos, permitiendo definir políticas de seguridad basadas en roles y permisos. Por su parte, la autenticación multifactor (MFA) añade capas adicionales de seguridad al combinar distintos factores de verificación, como contraseñas, tokens y biometría. Finalmente, los gestores de contraseñas permiten almacenar credenciales de forma cifrada, facilitando su administración y reduciendo el riesgo asociado al uso de contraseñas débiles o repetidas.

La Tabla 2.2 presenta una comparación clara de los distintos tipos de credenciales, destacando el equilibrio existente entre seguridad, usabilidad y riesgos asociados. En primer lugar, las contraseñas se caracterizan por su facilidad de implementación y amplio uso; sin embargo,

su principal debilidad radica en su vulnerabilidad frente a ataques de fuerza bruta o malas prácticas como el uso de claves débiles o repetidas.

Por otro lado, la autenticación multifactor (MFA) incrementa significativamente el nivel de seguridad al requerir múltiples formas de verificación, aunque puede afectar la experiencia del usuario al añadir pasos adicionales en el proceso de acceso. En contraste, los tokens físicos ofrecen un nivel de seguridad muy alto, ya que implican la posesión de un dispositivo específico, pero presentan el inconveniente de que pueden perderse o dañarse.

Finalmente, la biometría destaca por su rapidez y comodidad, al basarse en características únicas del usuario; no obstante, no está exenta de limitaciones, como la posibilidad de falsos positivos o negativos, lo que puede afectar su fiabilidad en ciertos contextos.

Tabla 2.2. Cuadro comparativo de tipos de credenciales.

Método	Ventajas	Desventajas	Ejemplo de uso
Contraseña	Simple de implementar	Vulnerable a ataques de fuerza bruta	Inicio de sesión en SO
MFA	Alta seguridad	Menor comodidad para el usuario	Acceso a banca en línea
Token físico	Muy seguro, portátil	Puede perderse	Llaves YubiKey
Biometría	Conveniente, rápida	Posibles falsos positivos/negativos	Desbloqueo de laptops

Los escenarios de riesgo asociados a la gestión inadecuada de credenciales y llaves criptográficas representan una de las principales fuentes de vulnerabilidad en los sistemas de información. Uno de los casos más comunes es la configuración de routers y switches con credenciales predeterminadas, lo que facilita el acceso no autorizado por parte de atacantes que conocen estas configuraciones estándar. Asimismo, la exposición de llaves privadas en repositorios públicos, como GitHub, constituye un riesgo crítico, ya que puede comprometer certificados SSL o accesos mediante

claves SSH, permitiendo el control indebido de sistemas y servicios.

Otro escenario frecuente es el uso de contraseñas débiles en servidores, lo que abre la puerta a ataques de fuerza bruta que buscan descubrir credenciales mediante intentos automatizados. De igual manera, la ausencia de autenticación multifactor (MFA) incrementa significativamente la vulnerabilidad frente a ataques de phishing, donde los atacantes logran obtener credenciales legítimas mediante engaño, accediendo posteriormente a sistemas sensibles sin mayores restricciones.

En cuanto a los casos de uso, en el ámbito del ensamblaje de equipos es fundamental implementar medidas de seguridad desde las primeras etapas de configuración. Esto incluye la protección del BIOS mediante contraseñas seguras, evitando modificaciones no autorizadas en el arranque del sistema. Asimismo, la implementación de tecnologías de cifrado de discos, como BitLocker en entornos Windows o LUKS en sistemas Linux, permite proteger la información almacenada frente a accesos indebidos, especialmente en caso de pérdida o robo de los dispositivos.

En el ámbito de las redes, la adopción de mecanismos de autenticación seguros resulta esencial. El uso de SSH con llaves públicas, en lugar de contraseñas simples, incrementa significativamente la seguridad en la administración remota de dispositivos. Además, la implementación de protocolos como RADIUS o TACACS+ permite centralizar la autenticación de usuarios en routers y switches, facilitando el control de accesos y la aplicación de políticas de seguridad. Por otro lado, el uso de certificados digitales en redes privadas virtuales (VPN) garantiza la autenticación de los participantes y la protección de las comunicaciones.

Para mitigar estos riesgos, es necesario aplicar buenas prácticas en la gestión de credenciales y llaves. Entre ellas destaca la prohibición de almacenar claves privadas en dispositivos compartidos o inseguros, así como el uso de gestores de contraseñas cifrados que

permitan administrar credenciales de manera segura. También es fundamental documentar políticas claras de manejo de credenciales, monitorear de forma constante los intentos de acceso fallidos y establecer políticas de rotación periódica de contraseñas y llaves criptográficas, reduciendo así la probabilidad de exposición prolongada.

No obstante, existen errores frecuentes que continúan comprometiendo la seguridad de los sistemas. Uno de ellos es el denominado hardcoding, práctica en la que los desarrolladores incluyen contraseñas directamente en el código fuente, lo que puede exponerlas si el código es compartido o vulnerado. Otro problema relevante es el fenómeno del Shadow IT, que implica el uso de aplicaciones no autorizadas dentro de una organización, muchas de las cuales gestionan credenciales sin cumplir estándares de seguridad adecuados. Finalmente, la generación de copias de seguridad sin cifrado representa un riesgo significativo, ya que estas pueden contener información sensible, incluyendo contraseñas en texto plano, que podrían ser explotadas en caso de acceso no autorizado.

2.3. Modelo COBIT para la gobernanza y gestión integral de Tecnologías de la Información

En el ámbito de la seguridad informática, no es suficiente con implementar medidas técnicas aisladas como firewalls, cifrado de discos duros o el uso de contraseñas robustas. Aunque estas herramientas son fundamentales para la protección de los sistemas, por sí solas no garantizan una seguridad efectiva si no se integran dentro de un enfoque global. Por ello, resulta imprescindible establecer un marco organizativo que permita dirigir, supervisar y controlar de manera adecuada la gestión de los activos tecnológicos y de la información. Este enfoque integral se conoce como gobernanza de la ciberseguridad, y constituye un elemento clave para asegurar la coherencia entre la estrategia empresarial y la gestión de la tecnología.

La gobernanza de la ciberseguridad establece las directrices estratégicas y define una estructura clara de

roles y responsabilidades orientadas a la protección de la información y los sistemas. En este contexto, marcos de referencia como COBIT 2019 ofrecen modelos estructurados que permiten a las organizaciones gestionar de forma eficaz sus recursos tecnológicos, mejorar la toma de decisiones y asegurar el cumplimiento normativo. Diversos estudios han demostrado que la aplicación de estos marcos favorece la alineación entre los objetivos de negocio y los procesos de tecnologías de la información, al mismo tiempo que incrementa la transparencia y la eficiencia en la gestión. No obstante, también se identifican retos importantes, como la resistencia al cambio organizacional, la complejidad en la adopción de nuevos modelos de gestión o la falta de madurez en los procesos internos (Antariksa et al., 2025; Vaya-Arboledas et al., 2025).

A partir de este marco de gobernanza se derivan las políticas de ciberseguridad, que constituyen el conjunto de normas, reglas y procedimientos que deben seguir todos los miembros de la organización, desde usuarios hasta administradores de sistemas. Estas políticas no solo regulan el uso adecuado de los recursos tecnológicos, sino que también establecen mecanismos de control y supervisión que permiten prevenir incidentes y responder de manera eficaz ante posibles amenazas. En este sentido, la integración de marcos como COBIT 2019 con otros enfoques como ITIL 4 contribuye a mejorar la gestión de servicios de TI, optimizar los procesos operativos y reforzar los controles internos. Investigaciones recientes evidencian que esta combinación permite incrementar el rendimiento organizacional, mejorar la calidad de los servicios tecnológicos y fortalecer la capacidad de respuesta frente a incidentes de seguridad (Mohamed et al., 2024; Nachrowi et al., 2020).

Además, la literatura señala que la gobernanza de la ciberseguridad no solo tiene un impacto en la protección de los activos digitales, sino también en el desempeño global de la organización. En sectores como el comercio minorista, por ejemplo, se ha comprobado que la implementación de marcos de gobernanza basados en COBIT 2019 permite mejorar la eficiencia operativa,

optimizar la gestión de riesgos y facilitar la adaptación a entornos tecnológicos en constante evolución (Viriyatama Lim y Indah Fianty, 2023). Esto demuestra que la gobernanza no debe entenderse únicamente como un mecanismo de control, sino también como una herramienta estratégica que impulsa la competitividad y la innovación.

De esta forma la gobernanza de la ciberseguridad puede definirse como el conjunto de mecanismos, políticas, procedimientos y estructuras organizativas que permiten gestionar de manera segura los activos de información y los recursos tecnológicos. Su correcta implementación no solo contribuye a reducir riesgos y prevenir incidentes, sino que también favorece el cumplimiento normativo, mejora la eficiencia organizacional y refuerza la confianza de clientes, socios y otras partes interesadas en un entorno digital cada vez más complejo y exigente.

En esencia, la gobernanza en ciberseguridad comprende un conjunto de estructuras, procesos y políticas que permiten dirigir y controlar la seguridad de la información dentro de una organización, asegurando que esta se gestione en coherencia con los objetivos estratégicos del negocio. Entre sus principales propósitos se encuentran la definición clara de roles y responsabilidades, la alineación de la seguridad con los objetivos organizacionales, el cumplimiento de normativas legales y estándares internacionales, el establecimiento de mecanismos de control y mejora continua, así como la promoción de una cultura de seguridad entre los usuarios.

En la práctica, esto se materializa a través de políticas concretas aplicables a redes y equipos. Por ejemplo, una política de acceso a dispositivos puede establecer que únicamente el personal técnico autorizado tenga la capacidad de configurar routers, switches o servidores. Asimismo, una política de respaldo define la realización periódica de copias de seguridad de configuraciones de red y sistemas críticos, mientras que una política de actualizaciones exige la instalación regular de parches de seguridad en sistemas operativos y firmware. Por su parte, la política de uso aceptable delimita el

comportamiento esperado en el uso de dispositivos y redes. La ausencia de estas políticas implica que decisiones críticas queden en manos de los usuarios, lo que puede comprometer seriamente la seguridad de toda la infraestructura tecnológica.

Dentro de los modelos de gobernanza más reconocidos se encuentra COBIT (Control Objectives for Information and Related Technologies), un marco internacional ampliamente utilizado para la gobernanza y gestión de las tecnologías de la información. Este modelo fue desarrollado por ISACA y tiene como objetivo ayudar a las organizaciones a alcanzar sus metas estratégicas mediante una gestión eficiente de las TI. COBIT establece una clara diferenciación entre gobernanza y gestión, no impone soluciones específicas, sino que proporciona lineamientos generales adaptables a distintos contextos, y permite su integración con otros estándares y marcos de referencia.

La implementación de COBIT aporta múltiples beneficios, entre los que destacan la mejora de la eficiencia operativa mediante la optimización de procesos, la alineación de la tecnología con los objetivos del negocio, y la adopción de un enfoque estructurado para la gestión de riesgos. Además, contribuye a generar valor para la organización, fortalecer la toma de decisiones, establecer responsabilidades claras, facilitar el cumplimiento normativo y promover la mejora continua. Asimismo, incrementa la transparencia y la rendición de cuentas en la gestión de las tecnologías de la información.

En cuanto a su estructura, COBIT 5 está conformado por varios componentes que permiten un enfoque integral de la gobernanza de TI. Incluye principios que establecen las bases conceptuales del modelo, un conjunto de 37 procesos organizados en cinco dominios, y una serie de habilitadores que apoyan la implementación de dichos procesos. Todo ello se orienta a garantizar la alineación con los objetivos del negocio, la generación de valor, la adecuada gestión de riesgos y el uso eficiente de los recursos tecnológicos.

La Figura 2.3 presenta los principios fundamentales de COBIT 5, los cuales constituyen la base conceptual para implementar una gobernanza y gestión eficaz de las tecnologías de la información en una organización. Estos principios no solo orientan la toma de decisiones, sino que también permiten asegurar que la tecnología contribuya de manera directa al logro de los objetivos estratégicos, generando valor y minimizando riesgos.

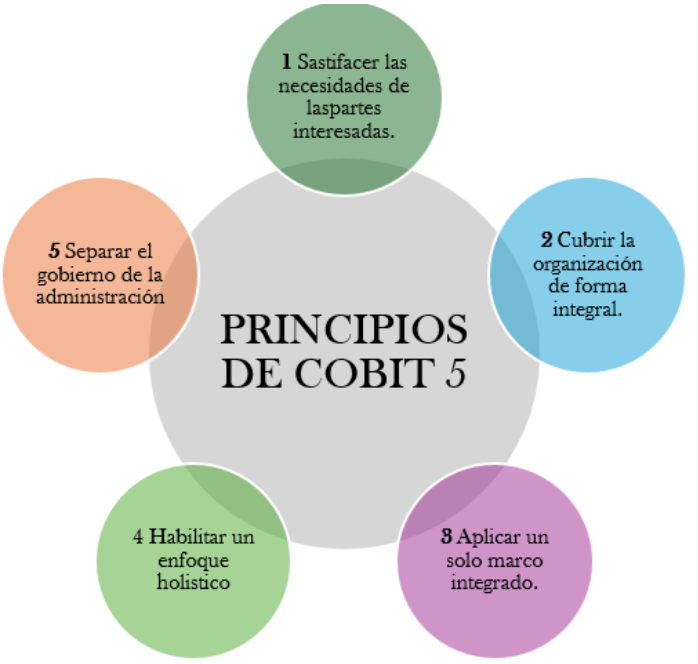


Figura 2.3. Principios Fundamentales del COBIT.

En primer lugar, el principio de satisfacer las necesidades de las partes interesadas establece que toda actividad relacionada con las tecnologías de la información debe enfocarse en responder a las expectativas de los stakeholders. Esto implica identificar quiénes son las partes interesadas, comprender sus necesidades y prioridades, y equilibrar sus intereses, que en muchos casos pueden ser distintos o incluso contrapuestos. A partir de este análisis, la organización debe tomar decisiones que maximicen el valor generado por las TI, considerando simultáneamente los beneficios, los riesgos y el uso eficiente de los recursos. Este principio refuerza la idea de que la tecnología no es un fin en sí mismo, sino un medio para apoyar los objetivos del negocio.

El segundo principio, cubrir la empresa de extremo a extremo, propone una visión integral de la gobernanza de TI. Esto significa que la gestión de la tecnología no debe limitarse al área informática, sino que debe abarcar toda la organización, incluyendo procesos, departamentos y funciones. De esta manera, se logra una alineación completa entre la estrategia empresarial y el uso de las tecnologías, evitando silos organizacionales y asegurando que todas las áreas trabajen bajo un mismo enfoque. Además, este principio incorpora tanto los procesos internos como las relaciones con terceros, lo que permite una gestión más completa y coherente.

En tercer lugar, el principio de aplicar un solo marco integrado busca unificar las distintas prácticas, estándares y normativas en un único sistema coherente. En este sentido, COBIT 5 se alinea con marcos internacionales como ISO/IEC 9000 e ISO/IEC 31000, lo que facilita su adopción y compatibilidad con otros sistemas de gestión ya implementados en la organización. Esta integración evita duplicidades, reduce la complejidad y permite una gestión más eficiente, al proporcionar un lenguaje común y una estructura unificada para la gobernanza de TI.

El cuarto principio, habilitar un enfoque holístico integrado, destaca la importancia de considerar todos los elementos que influyen en la gobernanza y gestión de TI como un sistema interrelacionado. Esto incluye procesos, estructuras organizativas, cultura, información, servicios, infraestructura y personas. COBIT denomina a estos elementos "habilitadores", los cuales deben funcionar de manera coordinada para lograr los objetivos de la organización. Este enfoque permite comprender que los problemas o mejoras en un área pueden impactar en otras, por lo que se requiere una visión global para una gestión efectiva.

Finalmente, el principio de separar la gobernanza de la gestión establece una clara distinción entre las funciones estratégicas y las operativas dentro de la organización. La gobernanza se encarga de definir la dirección, establecer objetivos y supervisar el cumplimiento, mientras que la gestión se ocupa de planificar, ejecutar y monitorear las actividades necesarias para alcanzar

dichos objetivos. Esta separación permite una mayor claridad en los roles y responsabilidades, mejora la rendición de cuentas y favorece una toma de decisiones más efectiva, al evitar conflictos de interés y asegurar que cada nivel cumpla su función específica.

Estos principios conforman un marco sólido que permite a las organizaciones gestionar sus tecnologías de la información de manera eficiente, alineada con el negocio y orientada a la mejora continua. Su correcta aplicación contribuye a optimizar el uso de los recursos, fortalecer la gestión de riesgos, mejorar la transparencia organizacional y garantizar que las TI generen valor real para la organización.

El análisis en cascada, representado en la Figura 2.4, es un concepto clave dentro de COBIT 5, ya que permite traducir de manera estructurada las necesidades de las partes interesadas en acciones concretas dentro de la organización. Este enfoque parte del principio fundamental de satisfacer las necesidades de los interesados y se desarrolla progresivamente hasta llegar a la definición de metas específicas en los procesos y actividades de TI.

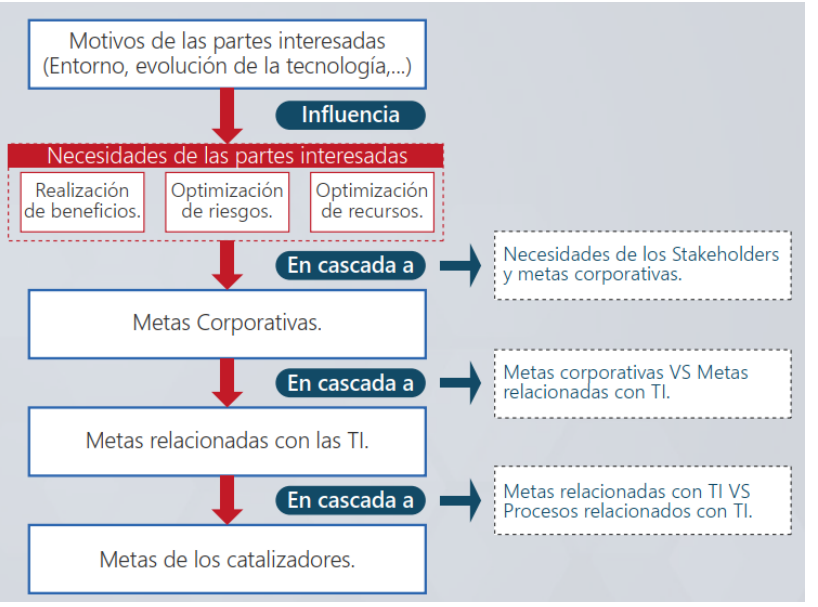


Figura 2.4. Análisis en cascada.

El modelo de cascada comienza con la identificación de las necesidades y expectativas de los stakeholders, las cuales constituyen el punto de partida para toda la gobernanza de TI. Estas necesidades suelen estar relacionadas con la generación de valor, la optimización de riesgos y el uso eficiente de los recursos. A partir de este análisis inicial, dichas expectativas se transforman en metas empresariales, que reflejan los objetivos estratégicos que la organización busca alcanzar.

Posteriormente, estas metas empresariales se traducen en metas relacionadas con las tecnologías de la información, las cuales establecen cómo las TI deben contribuir al cumplimiento de los objetivos del negocio. Este paso es fundamental, ya que permite alinear la estrategia tecnológica con la estrategia organizacional, asegurando que ambas trabajen de manera coordinada.

Finalmente, las metas de TI se descomponen en metas de los procesos catalizadores (habilitadores), que corresponden a actividades específicas, procesos operativos y controles necesarios para implementar la estrategia. En este nivel se concretan las acciones prácticas que deben ejecutarse para lograr los objetivos definidos en los niveles superiores.

Este análisis en cascada permite establecer una trazabilidad clara entre lo que esperan las partes interesadas y lo que se ejecuta en los niveles operativos de la organización. Además, facilita la alineación estratégica, mejora la toma de decisiones y asegura que cada proceso contribuya de manera directa a la generación de valor. En consecuencia, se convierte en una herramienta esencial para garantizar que la gobernanza de TI sea coherente, medible y orientada a resultados.

El principio 4 de COBIT 5, denominado “habilitar un enfoque holístico integrado”, representa uno de los pilares más profundos y trascendentales dentro del modelo de gobernanza de las tecnologías de la información. Este principio reconoce que la realidad organizacional no puede ser comprendida ni gestionada a partir de elementos aislados, sino como un sistema complejo,

dinámico e interdependiente, en el que múltiples factores interactúan de manera simultánea para alcanzar los objetivos estratégicos.

En este contexto, COBIT establece que las metas de alto nivel solo pueden alcanzarse cuando se consideran de forma conjunta los siete habilitadores, los cuales constituyen los componentes esenciales que hacen posible la gobernanza y la gestión efectiva de las TI. Estos habilitadores no operan de manera independiente, sino que forman una red integrada en la que cada uno influye y es influenciado por los demás. De esta forma, la organización deja de concebir la tecnología como un conjunto de herramientas aisladas y la entiende como un ecosistema en el que convergen procesos, estructuras, cultura organizacional, información, servicios, infraestructura y personas.

El carácter holístico de este enfoque implica que cualquier iniciativa, cambio o mejora en el ámbito de las tecnologías de la información debe ser evaluado considerando su impacto en todos los habilitadores. Por ejemplo, la implementación de una nueva tecnología no solo requiere ajustes técnicos, sino también cambios en los procesos, capacitación del personal, adaptación de la cultura organizacional y redefinición de responsabilidades. Ignorar alguno de estos elementos puede generar desequilibrios que comprometan el logro de los objetivos.

Asimismo, este principio enfatiza la necesidad de integración, entendida como la capacidad de articular coherentemente todos los habilitadores hacia un propósito común. La integración evita la fragmentación de esfuerzos, reduce redundancias y permite una utilización más eficiente de los recursos. En este sentido, la gobernanza de TI se convierte en un mecanismo de coordinación estratégica que asegura que todas las partes del sistema trabajen de manera alineada.

La Figura 2.5 ilustra cómo estos habilitadores actúan como medios fundamentales para alcanzar los objetivos deseados, mostrando que el éxito organizacional no

depende de un único factor, sino de la interacción equilibrada de todos ellos. Este enfoque refuerza la idea de que la excelencia en la gestión de TI no se logra mediante soluciones parciales, sino a través de una visión integral que considere la complejidad del entorno organizacional.

Figura 2.5. Habilitadores para alcanzar los objetivos deseados.

Habilitar un enfoque holístico integrado implica adoptar una perspectiva sistémica de la gobernanza de TI, en la que cada decisión se analiza en función de su impacto global. Este principio no solo fortalece la capacidad de la organización para alcanzar sus objetivos, sino que también promueve la resiliencia, la adaptabilidad y la mejora continua en un entorno tecnológico cada vez más complejo y cambiante.

El principio 5 de COBIT 5, denominado "separar la gobernanza de la gestión", constituye un elemento esencial para garantizar la claridad organizacional, la eficiencia en la toma de decisiones y la correcta asignación de responsabilidades dentro de las estructuras de tecnología de la información. Este principio establece una distinción explícita y rigurosa entre dos ámbitos fundamentales: el gobierno (governance) y la administración o gestión (management), los cuales, aunque complementarios, cumplen funciones claramente diferenciadas.

El gobierno se sitúa en el nivel estratégico de la organización y, por lo general, es responsabilidad de la alta dirección, particularmente de la junta directiva bajo el liderazgo de su presidente. Su función principal consiste en asegurar que se alcancen los objetivos organizacionales, estableciendo la dirección, definiendo prioridades y evaluando continuamente el desempeño. En este nivel se toman decisiones clave orientadas al largo plazo, se determinan las políticas generales y se supervisa que las acciones ejecutadas estén alineadas con la visión y los intereses de las partes interesadas. Además, el gobierno tiene un rol crítico en la supervisión y control, asegurando que los resultados obtenidos correspondan con lo esperado y que los riesgos estén adecuadamente gestionados.

Por otro lado, la administración o gestión opera en el nivel táctico y operativo, y recae en la gerencia ejecutiva, liderada por el director general o CEO. Su responsabilidad es llevar a la práctica las directrices establecidas por el gobierno, mediante la planificación, construcción, ejecución y monitoreo de las actividades necesarias para cumplir los objetivos definidos. En este ámbito se desarrollan e implementan procesos, se asignan recursos, se gestionan equipos de trabajo y se supervisa el desempeño diario de las operaciones. La gestión, por tanto, traduce la estrategia en acciones concretas, asegurando que las decisiones estratégicas se materialicen de manera eficiente y efectiva.

La Figura 2.6 ilustra claramente esta división funcional, mostrando cómo el gobierno se enfoca en dirigir y evaluar, mientras que la gestión se encarga de planificar y ejecutar. Esta separación no implica aislamiento, sino una relación de complementariedad en la que ambos niveles deben interactuar de manera coordinada. Sin embargo, mantener esta distinción es crucial para evitar conflictos de interés, duplicidad de funciones y falta de rendición de cuentas.

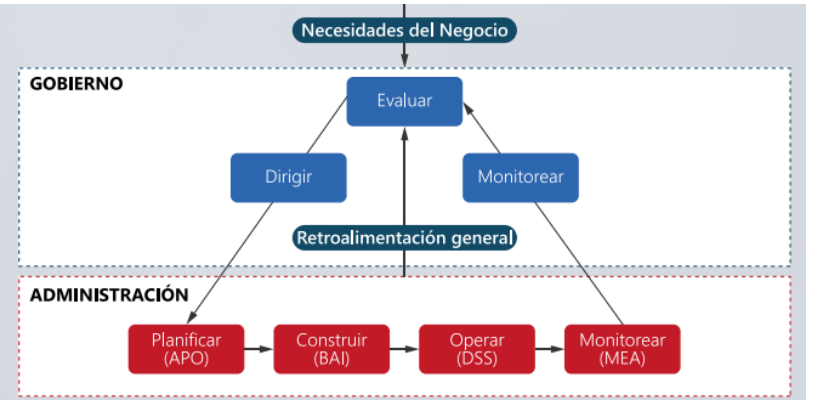


Figura 2.6. Funciones del gobierno y la administración.

Este principio refuerza la necesidad de establecer límites claros entre quienes definen el rumbo estratégico y quienes se encargan de implementarlo. Al hacerlo, se promueve una gobernanza más transparente, una gestión más eficiente y una organización mejor preparada para responder a los desafíos del entorno tecnológico, garantizando así que las tecnologías de la información generen valor real y sostenible para el negocio.

COBIT 5 define un marco integral de gobierno y gestión de tecnologías de la información compuesto por 37 procesos, organizados en cinco dominios que permiten estructurar de manera eficiente las actividades relacionadas con la dirección, control y operación de TI dentro de una organización. Este modelo facilita la alineación entre los objetivos del negocio y las tecnologías de la información, asegurando una adecuada gestión de recursos, riesgos y generación de valor.

Tal como se observa en la Figura 2.7, los procesos se distribuyen de manera estructurada en torno a un eje central que diferencia claramente el gobierno de TI de su gestión operativa. En la parte superior del modelo se ubica el dominio EDM (Evaluar, Dirigir y Supervisar), que agrupa cinco procesos orientados al gobierno de TI empresarial. Este dominio se encarga de establecer las directrices estratégicas, evaluar el desempeño y asegurar que las decisiones relacionadas con TI estén alineadas con los objetivos organizacionales.

Por debajo de este nivel, la figura muestra los cuatro dominios correspondientes a la gestión de TI, los cuales concentran los 32 procesos restantes. El dominio APO (Alinear, Planificar y Organizar) se enfoca en la planificación estratégica y la definición de políticas, evidenciando en la figura su papel como base para la organización de los recursos tecnológicos. A su vez, el dominio BAI (Construir, Adquirir e Implementar) refleja en la figura la fase de desarrollo e implementación de soluciones tecnológicas, conectando la planificación con la ejecución.

El dominio DSS (Entregar, Dar Servicio y Soporte), representado en la figura como la capa operativa, se encarga de la prestación de servicios y el soporte continuo, garantizando el funcionamiento diario de los sistemas. Finalmente, el dominio MEA (Monitorear, Evaluar y Valorar) aparece como un componente transversal en la figura, encargado de supervisar y evaluar el desempeño de todos los procesos, asegurando la mejora continua.

El análisis de la Figura 2.7 permite comprender que COBIT 5 no solo organiza procesos de manera jerárquica, sino que también establece una relación cíclica entre planificación, implementación, operación y evaluación. Esta estructura evidencia un enfoque integral, en el que el gobierno define el rumbo estratégico, mientras que la gestión ejecuta, controla y optimiza las actividades de TI, garantizando así un equilibrio entre control, eficiencia y generación de valor.

Figura 2.7. Procesos COBIT 5.

Los procesos de gobierno de TI empresarial en COBIT 5 constituyen el nivel estratégico encargado de asegurar

que las tecnologías de la información generen valor para la organización, gestionen adecuadamente los riesgos y se alineen con los objetivos del negocio. Este conjunto de procesos se agrupa en el dominio EDM (Evaluate, Direct and Monitor – Evaluar, Dirigir y Supervisar), el cual representa la capa más alta dentro del modelo de gobierno.

Tal como se observa en la Figura 2.8, los procesos de gobierno se presentan como un bloque estructurado que guía y supervisa todas las actividades relacionadas con la gestión de TI. La figura evidencia que estos procesos no operan de manera aislada, sino que actúan como un eje central de control que influye directamente en los dominios operativos. En este sentido, el gobierno de TI no ejecuta tareas técnicas, sino que establece lineamientos, define políticas y supervisa el desempeño de la gestión.

El dominio EDM está compuesto por cinco procesos fundamentales. El primero, EDM01 (Asegurar el establecimiento y mantenimiento del marco de gobierno), se enfoca en definir la estructura y los principios que regirán la gestión de TI dentro de la organización. En segundo lugar, EDM02 (Asegurar la entrega de beneficios) busca garantizar que las inversiones en TI generen valor real para el negocio. El proceso EDM03 (Asegurar la optimización del riesgo) se centra en identificar, evaluar y gestionar los riesgos asociados al uso de las tecnologías.

Asimismo, EDM04 (Asegurar la optimización de recursos) tiene como objetivo garantizar el uso eficiente de los recursos tecnológicos, humanos y financieros. Finalmente, EDM05 (Asegurar la transparencia hacia las partes interesadas) promueve la rendición de cuentas y la comunicación efectiva sobre el desempeño de TI.

El análisis de la Figura 2.8 permite comprender que estos procesos de gobierno actúan como un mecanismo de supervisión continua, estableciendo un ciclo en el que se evalúan las necesidades, se dirigen las acciones y se monitorean los resultados. Esta estructura refuerza la idea de que el gobierno de TI no solo define el rumbo

estratégico, sino que también asegura el control y la mejora continua de las actividades tecnológicas dentro de la organización.

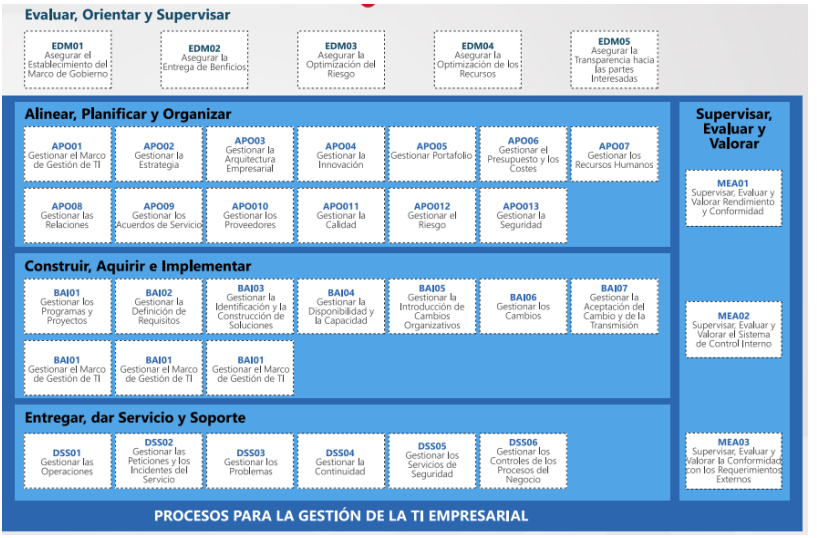


Figura 2.8. Procesos de gobierno de TI empresarial.

Los habilitadores en el marco COBIT 5 constituyen los elementos fundamentales que permiten a las organizaciones alcanzar sus objetivos estratégicos y garantizar un gobierno y gestión eficaz de las tecnologías de la información. Estos habilitadores no solo influyen en el éxito de los procesos organizacionales, sino que también determinan la manera en que se estructuran, ejecutan y supervisan las actividades relacionadas con TI. En este sentido, representan factores clave que integran tanto aspectos técnicos como organizativos y humanos dentro de un enfoque holístico de la ciberseguridad y la gestión tecnológica.

Tal como se observa en la Figura 2.9, los habilitadores se presentan como un conjunto de siete componentes interrelacionados que abarcan diferentes dimensiones de la organización. En primer lugar, los principios, directrices y marcos de referencia establecen las bases normativas y estratégicas que guían la toma de decisiones. A continuación, los procesos definen las actividades necesarias para alcanzar los objetivos organizacionales, evidenciando su papel central en la ejecución de la estrategia.

Figura 2.9. Habilitadores.

Asimismo, la figura destaca las estructuras organizativas, que representan la distribución de roles y responsabilidades dentro de la empresa, permitiendo una adecuada gobernanza de TI. De igual manera, la cultura, ética y comportamiento reflejan la importancia del factor humano en la adopción de buenas prácticas y en el cumplimiento de las políticas de seguridad. Por otro lado, la información se presenta como un recurso esencial, ya que constituye el activo principal que debe ser protegido y gestionado de manera eficiente.

En los niveles más operativos, la figura incluye los servicios, infraestructura y aplicaciones, que representan los componentes tecnológicos necesarios para el funcionamiento de la organización. Finalmente, las habilidades y competencias del personal destacan como un elemento crítico, ya que el conocimiento y la capacitación de los empleados influyen directamente en la correcta implementación de los controles y procesos de seguridad.

El análisis de la Figura 2.9 permite comprender que los habilitadores no funcionan de manera aislada, sino que están interconectados y se influyen mutuamente. Esta interdependencia evidencia que el éxito en la gestión de TI no depende únicamente de la tecnología, sino también de factores organizacionales y humanos. En consecuencia, COBIT 5 propone un enfoque integral en el que todos los habilitadores deben alinearse para

garantizar una gestión eficiente, segura y orientada al logro de los objetivos empresariales.

2.4. Marcos de referencia en ciberseguridad y gestión de Tecnologías de la Información: NIST, ITIL y políticas organizacionales

El **NIST Cybersecurity Framework (CSF)** se ha consolidado como uno de los marcos de referencia más influyentes en el ámbito de la ciberseguridad a nivel internacional. Fue desarrollado por el National Institute of Standards and Technology, organismo dependiente del Departamento de Comercio de los Estados Unidos, con el propósito de proporcionar una guía estructurada que permita a las organizaciones gestionar de forma eficaz los riesgos relacionados con la seguridad de la información. Su diseño responde a la necesidad de contar con un modelo flexible, adaptable y alineado con distintos contextos organizativos y regulatorios.

Este marco se fundamenta en un conjunto de estándares, directrices y buenas prácticas que facilitan a organizaciones de cualquier tamaño y sector la identificación, evaluación y mitigación de riesgos de ciberseguridad. Según diversos estudios, el NIST CSF destaca por su enfoque práctico y su capacidad de integración con otros marcos y normativas internacionales, lo que lo convierte en una herramienta clave para la gestión del riesgo y el cumplimiento regulatorio (McIntosh et al., 2024; Taherdoost, 2022). Además, su estructura permite a las organizaciones establecer un lenguaje común en materia de ciberseguridad, facilitando la comunicación entre áreas técnicas y de negocio.

Uno de los aspectos más relevantes del NIST CSF es su enfoque basado en cinco funciones principales: identificar, proteger, detectar, responder y recuperar. Estas funciones proporcionan un ciclo completo de gestión de la seguridad, que abarca desde la comprensión de los activos y riesgos hasta la capacidad de respuesta ante incidentes y la recuperación de las operaciones. Investigaciones recientes destacan que este enfoque integral contribuye significativamente a mejorar la resiliencia organizacional frente a ciberamenazas, al

permitir no solo prevenir ataques, sino también minimizar su impacto y acelerar la recuperación (Salas Riega et al., 2025).

Asimismo, el NIST CSF no se limita a la protección de sistemas y datos, sino que promueve una visión más amplia de la ciberseguridad, incluyendo aspectos como la gobernanza, la gestión del riesgo y la coordinación entre diferentes actores. En entornos complejos y altamente interconectados, como el sector sanitario o los ecosistemas digitales globales, este tipo de marcos resulta esencial para garantizar la interoperabilidad y la protección de la información compartida (Luidold y Jungbauer, 2024).

El NIST Cybersecurity Framework constituye una herramienta estratégica que permite a las organizaciones adoptar un enfoque estructurado, flexible y orientado a la mejora continua en la gestión de la ciberseguridad. Su aplicación no solo facilita la reducción de riesgos, sino que también fortalece la capacidad de adaptación ante amenazas emergentes, contribuyendo a la sostenibilidad y seguridad de los entornos digitales en un contexto cada vez más dinámico y exigente.

Una de las principales características del NIST CSF es su estructura basada en cinco funciones fundamentales que representan el ciclo de vida de la gestión de la ciberseguridad: Identificar, Proteger, Detectar, Responder y Recuperar. Estas funciones permiten organizar las actividades de seguridad de manera lógica y progresiva. La función Identificar se enfoca en comprender el entorno organizacional, los activos críticos y los riesgos asociados; Proteger establece las medidas necesarias para salvaguardar los sistemas; Detectar se orienta a identificar oportunamente incidentes de seguridad; Responder define las acciones a ejecutar ante un incidente; y Recuperar se centra en restablecer las capacidades operativas y garantizar la continuidad del negocio.

El diseño del NIST CSF destaca por su alta flexibilidad, lo que permite su integración con otros marcos y estándares existentes, así como su adaptación a distintos contextos

organizacionales e industriales. Esta característica lo convierte en una herramienta versátil que puede ser utilizada tanto por grandes corporaciones como por pequeñas y medianas empresas, independientemente de su nivel de madurez en ciberseguridad. Además, el marco es especialmente relevante en entornos altamente dependientes de la tecnología, como los sistemas de información (TI), sistemas de control industrial (ICS), sistemas ciberfísicos (CPS) y dispositivos conectados en el ecosistema del Internet de las Cosas.

Desde su creación, el NIST CSF ha evolucionado como respuesta a la creciente complejidad de las amenazas cibernéticas. Su origen se remonta a la Orden Ejecutiva 13636 emitida en 2013 en los Estados Unidos, la cual impulsó la colaboración entre el sector público y privado para fortalecer la ciberseguridad de la infraestructura crítica del país. Como resultado de esta iniciativa, se desarrolló la versión inicial del marco (CSF 1.0), que integró diversas normas y buenas prácticas existentes en un modelo unificado. Posteriormente, la Ley de Mejora de la Ciberseguridad de 2014 consolidó el papel del NIST en el desarrollo de estándares y directrices en esta materia.

Actualmente, la versión 1.1 del NIST CSF, publicada en abril de 2018, sigue siendo ampliamente utilizada a nivel global y se ha consolidado como un referente clave para la gestión de riesgos de ciberseguridad. Su adopción en múltiples sectores demuestra su efectividad como herramienta para fortalecer la resiliencia organizacional, mejorar la toma de decisiones y establecer un enfoque estructurado y proactivo frente a las amenazas digitales. El marco consta de tres componentes principales: 1) el núcleo del marco o framework core; 2) los niveles de implementación o tiers; y 3) y los perfiles del marco

El Framework Core del NIST Cybersecurity Framework constituye el componente central del modelo, ya que integra un conjunto estructurado de actividades, resultados y objetivos de ciberseguridad alineados con estándares internacionales y buenas prácticas de la industria. Este núcleo proporciona una base común que permite a las organizaciones organizar, evaluar y

mejorar sus capacidades de ciberseguridad de manera sistemática y coherente.

El Framework Core se encuentra estructurado en tres niveles jerárquicos: funciones, categorías y subcategorías. En el nivel superior se ubican las funciones, que representan los grandes dominios de actividad en ciberseguridad y organizan las acciones en torno al ciclo de gestión del riesgo. En un nivel intermedio se encuentran las categorías, que agrupan objetivos específicos dentro de cada función y abarcan aspectos técnicos, organizacionales y humanos. Finalmente, en el nivel más detallado se sitúan las subcategorías, las cuales consisten en declaraciones orientadas a resultados que proporcionan directrices concretas para la implementación y mejora de un programa de ciberseguridad. En total, el marco incluye 23 categorías y 108 subcategorías, lo que evidencia su alto nivel de detalle y aplicabilidad práctica.

Tal como se ilustra en la Figura 2.10, el Framework Core se organiza en torno a cinco funciones principales que constituyen los pilares fundamentales de un programa de ciberseguridad integral y eficaz. Estas funciones no deben entenderse como etapas aisladas, sino como componentes interrelacionados que conforman un ciclo continuo de gestión del riesgo.

Figura 2.10. Estructura - funciones del núcleo.

La función Identificar permite desarrollar un entendimiento integral del contexto organizacional, incluyendo los activos, sistemas, datos, capacidades y riesgos asociados. En esta fase se realiza un inventario de recursos, se analizan las amenazas y se establecen las bases para la gestión del riesgo, lo que resulta esencial para una adecuada toma de decisiones.

Por su parte, la función Proteger se centra en la implementación de medidas de seguridad destinadas a salvaguardar los sistemas y la información. Esto incluye el desarrollo de políticas de ciberseguridad, la

definición de roles y responsabilidades, la capacitación de los usuarios y la adopción de controles técnicos que permitan prevenir incidentes o limitar su impacto.

La función Detectar está orientada a la identificación oportuna de eventos de ciberseguridad. En esta etapa se implementan mecanismos de monitoreo continuo, análisis de comportamiento y sistemas de alerta que permiten detectar accesos no autorizados, anomalías o actividades sospechosas dentro de la infraestructura tecnológica.

En cuanto a la función Responder, esta abarca las acciones necesarias para gestionar incidentes de ciberseguridad una vez que han sido detectados. Incluye actividades como la contención del incidente, la notificación a las partes interesadas, la investigación de las causas y la implementación de medidas correctivas. Asimismo, esta fase incorpora la retroalimentación necesaria para mejorar las políticas y procedimientos existentes.

Finalmente, la función Recuperar se enfoca en restablecer las capacidades operativas afectadas por un incidente de seguridad. Esto implica la ejecución de planes de recuperación, la restauración de sistemas y datos, y la garantía de la continuidad del negocio. Además, se enfatiza la importancia de realizar pruebas periódicas de los planes de recuperación y mantener una comunicación transparente con empleados, clientes y otras partes interesadas, con el fin de preservar la confianza y fortalecer la resiliencia organizacional.

El análisis de la Figura 2.10 permite comprender que estas cinco funciones no solo estructuran las actividades de ciberseguridad, sino que también establecen un enfoque cíclico y dinámico orientado a la mejora continua. De este modo, el NIST CSF proporciona una herramienta eficaz para gestionar los riesgos de manera proactiva, adaptándose a un entorno tecnológico en constante evolución.

Las 23 categorías del NIST Cybersecurity Framework abarcan una amplia gama de objetivos relacionados con la ciberseguridad, proporcionando una estructura intermedia que conecta las funciones generales con acciones más específicas. Estas categorías están diseñadas para abordar de manera integral los distintos componentes de la seguridad, incluyendo aspectos técnicos, organizacionales y humanos, con un enfoque orientado a resultados medibles y alineados con la gestión del riesgo. Entre las categorías más relevantes se encuentran la gestión de riesgos, la gestión de identidades y control de acceso, la protección de datos, la detección de eventos de seguridad, la respuesta a incidentes y la recuperación de la continuidad del negocio, entre otras, lo que evidencia la amplitud y profundidad del marco (Tabla 2.3).

Por su parte, las subcategorías constituyen el nivel más detallado dentro del *Framework Core*, desglosándose en un total de 108 elementos. Estas subcategorías se presentan como declaraciones basadas en resultados que describen prácticas específicas de ciberseguridad, proporcionando una guía clara para el diseño, implementación y mejora continua de los programas de seguridad en una organización. Su nivel de detalle permite traducir los objetivos estratégicos en acciones concretas, facilitando la evaluación del estado actual de la ciberseguridad y la identificación de áreas de mejora.

Tabla 2. 3. Categorías.

FUNC-TION	CATEGORY	ID
Identificar	Gestión de activos.	ID.AM
	Entorno Empresarial.	ID.BE
	Gobernanza.	ID.GV
	Evaluación de riesgos.	ID.RA
	Estrategia de gestión de riesgo.	ID.RM
	Gestión de riesgos en la cadena de suministros.	ID.SC

FUNC-TION	CATEGORY	ID
Proteger	Gestión de identidades y control de acceso.	PR.AC
	Concienciación y formación.	PR.AT
	Seguridad de datos.	PR.DS
	Procesos y procedimientos de protección de la información.	PR.IP
	Mantenimiento.	PR.MA
	Tecnología de protección.	PR.PT
Detectar	Anomalías y eventos.	DE.AE
	Vigilancia continua de la seguridad.	DE.CM
	Procesos de detección.	DE.DP
Responder	Planificación de la respuesta.	RS.RP
	Comunicaciones.	RS.CO
	Análisis.	RS.AN
	Mitigación.	RS.MI
	Mejoras.	RS.IM
Recuperar	Planificación de la recuperación.	RC.RP
	Mejoras.	RC.IM
	Comunicaciones.	RC.CO

Los niveles de implementación del NIST Cybersecurity Framework describen el grado de madurez con el que una organización adopta y aplica las prácticas de ciberseguridad, así como el nivel de integración de estas dentro de sus procesos organizacionales. Estos niveles permiten evaluar qué tan formalizadas, gestionadas y optimizadas se encuentran las actividades relacionadas con la gestión del riesgo de ciberseguridad (Figura 2.11).

En el nivel Parcial, la organización ha iniciado la adopción del marco, pero aún no ha desarrollado completamente las actividades requeridas. En esta etapa, existe una comprensión limitada de los activos de información y de los riesgos asociados. Aunque pueden existir algunos controles de seguridad, estos no forman parte de una estrategia estructurada ni se aplican de manera consistente o sistemática.

El nivel Riesgo Informado se caracteriza por una mayor comprensión de los activos y los riesgos de la organización. En este punto, comienzan a desarrollarse políticas y

estrategias de seguridad más definidas, aunque todavía no se encuentran completamente integradas en todos los procesos. Asimismo, se promueve la concienciación del personal mediante capacitaciones y se establecen las bases para una gestión formal del riesgo.

En el nivel Repetible, la organización dispone de procesos de seguridad documentados, estandarizados y aplicados de manera consistente. Se implementan herramientas de monitoreo y análisis que permiten supervisar continuamente los sistemas, detectar anomalías y responder de forma más eficaz ante incidentes. La gestión de la ciberseguridad se vuelve más proactiva y estructurada.

Finalmente, el nivel Adaptado representa el grado más alto de madurez. En este nivel, las prácticas de ciberseguridad están completamente integradas en todos los procesos organizacionales y forman parte de la cultura empresarial. La gestión del riesgo se apoya en procesos automatizados, existe un conocimiento integral de los activos de información y se promueve la mejora continua mediante la evaluación constante de la efectividad de los controles implementados.

Figura 2.11. Estructura - niveles de implementación.

Los perfiles del marco tienen como propósito alinear las funciones, categorías y subcategorías con los objetivos y requisitos del negocio. Su finalidad es describir

tanto el estado actual como el estado deseado de la ciberseguridad dentro de una organización, permitiendo una gestión estratégica orientada a resultados.

El uso de perfiles facilita la comparación entre la situación presente y los objetivos futuros, lo que permite establecer una hoja de ruta clara para la mejora continua. De esta manera, las organizaciones pueden identificar, priorizar y planificar sus actividades de seguridad en función de sus necesidades específicas.

Existen dos tipos principales de perfiles:

• Perfil actual: representa el estado real de la ciberseguridad en la organización, incluyendo los controles implementados y los resultados alcanzados.
• Perfil objetivo: define el estado deseado que la organización busca alcanzar en términos de gestión del riesgo y madurez en ciberseguridad.

El NIST CSF propone una serie de pasos estructurados que permiten desarrollar o fortalecer un programa de ciberseguridad:

• Paso 1: Priorizar y determinar el alcance

Identificar los objetivos estratégicos del negocio y establecer las prioridades organizacionales para definir el alcance del programa.

• Paso 2: Orientación

Establecer lineamientos claros en materia de ciberseguridad alineados con los objetivos de la organización.

• Paso 3: Creación del perfil actual

Analizar el estado actual, identificando activos críticos, amenazas, vulnerabilidades y controles existentes.

• Paso 4: Determinación del perfil objetivo

Definir los objetivos de ciberseguridad considerando necesidades, recursos y nivel de riesgo aceptable.

• Paso 5: Análisis de brechas

Comparar el perfil actual con el perfil objetivo para identificar diferencias y áreas de mejora.

• Paso 6: Acción

Desarrollar e implementar un plan para cerrar las brechas identificadas.

• Paso 7: Revisión

Evaluar continuamente el estado de la ciberseguridad y repetir el proceso para asegurar la mejora continua.

Por su parte, ITIL es un marco de mejores prácticas para la gestión de servicios de tecnologías de la información. Fue desarrollado en la década de 1980 por la Office of Government Commerce del Reino Unido con el objetivo de estandarizar la administración de la infraestructura tecnológica.

Con el tiempo, ITIL se ha consolidado como un referente internacional, aplicable a organizaciones de distintos sectores (Figura 2.12). Su enfoque permite alinear los servicios tecnológicos con las necesidades del negocio, mejorar la calidad del soporte, optimizar la gestión operativa y garantizar la continuidad de los servicios, contribuyendo así a una gestión más eficiente y orientada al valor.

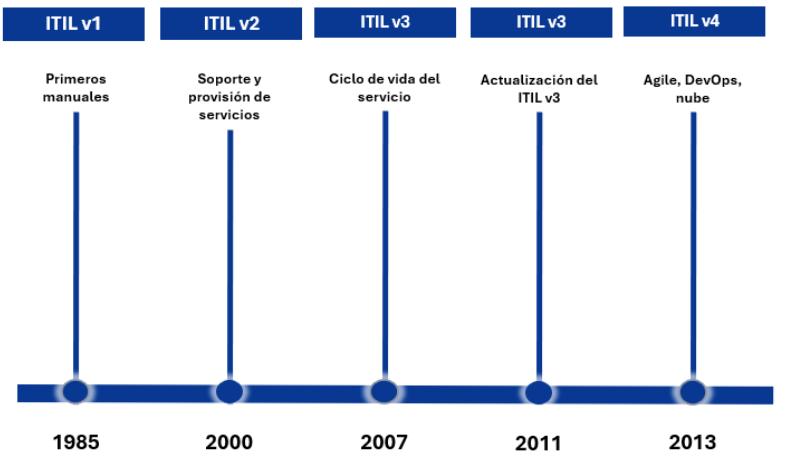


Figura 2.12. Evolución histórica de ITIL.

La versión más reciente de ITIL se fundamenta en siete principios rectores que orientan la toma de decisiones dentro de la gestión de servicios de TI. Estos principios permiten a las organizaciones adoptar un enfoque flexible, eficiente y centrado en la generación de valor:

• Enfocarse en el valor: todas las actividades y servicios deben estar orientados a generar valor tanto para el cliente como para la organización.

• Comenzar donde se está: se deben aprovechar los procesos, recursos y capacidades existentes antes de implementar cambios o crear nuevos sistemas.

• Progresar de manera iterativa con retroalimentación: las mejoras deben implementarse de forma gradual, evaluando continuamente los resultados obtenidos.

• Colaborar y promover la visibilidad: es fundamental fomentar el trabajo en equipo y la transparencia en la información para mejorar la toma de decisiones.

• Pensar y trabajar de manera holística: se deben considerar de forma integral todos los elementos del sistema, incluyendo personas, procesos, tecnología y socios.

• Mantenerlo simple y práctico: los procesos deben ser claros, eficientes y evitar complejidades innecesarias.

• Optimizar y automatizar: se deben utilizar herramientas tecnológicas para mejorar la eficiencia operativa y reducir errores humanos.

En cuanto al ciclo de vida del servicio, aunque ITIL 4 introduce una estructura más flexible, el modelo de ITIL v3 continúa siendo una referencia ampliamente utilizada en el ámbito académico y técnico. Este modelo se organiza en cinco etapas que representan el desarrollo completo de un servicio:

• Estrategia del servicio: se definen los objetivos, el público objetivo y el valor que se desea ofrecer.

• Diseño del servicio: se planifican los componentes necesarios, incluyendo niveles de servicio (SLA), seguridad, disponibilidad y capacidad.

• Transición del servicio: se implementa el servicio mediante procesos de cambio, pruebas y despliegue controlado.

• Operación del servicio: se gestionan las actividades diarias, incluyendo incidencias, problemas y solicitudes de usuarios.

• Mejora continua: se evalúa el desempeño del servicio y se implementan mejoras constantes para optimizar su funcionamiento.

Un ejemplo aplicado de este ciclo puede observarse en la implementación de un servicio de red corporativa. En primer lugar, se identifica la necesidad de conectividad (estrategia); posteriormente, se diseña la infraestructura adecuada (diseño); luego se implementa la solución tecnológica (transición); se supervisa su funcionamiento diario (operación); y finalmente se aplican mejoras, como la optimización del ancho de banda o la segmentación de la red (mejora continua).

Dentro de ITIL, existen procesos clave que permiten garantizar la eficiencia y continuidad de los servicios de TI:

• Gestión de incidencias: orientada a restaurar el servicio en el menor tiempo posible ante interrupciones.

• Gestión de problemas: enfocada en identificar y eliminar las causas raíz de los incidentes.
• Gestión de cambios: permite implementar modificaciones controladas sin afectar la continuidad del servicio.

• Gestión de configuración: asegura el control de los activos de TI y sus relaciones dentro del entorno tecnológico.

• Gestión de niveles de servicio (SLA): establece acuerdos claros entre proveedores y clientes respecto a la calidad del servicio.

Un ejemplo técnico ilustra la aplicación de estos procesos: si un switch en una red crítica falla y provoca una interrupción, ITIL propone gestionar inicialmente la incidencia para restaurar el servicio, posteriormente analizar el problema para identificar su causa (por ejemplo, una sobrecarga eléctrica), y finalmente implementar un cambio, como la instalación de un sistema de respaldo energético (UPS) y un switch redundante.

La aplicación de ITIL en entornos de ciberseguridad, ensamblaje y redes ofrece múltiples beneficios:

• Mejora de la disponibilidad de los servicios, reduciendo los tiempos de inactividad en sistemas y redes.

• Estandarización de procesos, facilitando la instalación, configuración y mantenimiento de la infraestructura tecnológica.

• Gestión proactiva de riesgos, permitiendo detectar fallos en hardware y software de manera anticipada.

• Optimización de recursos, mediante un uso más eficiente del personal, herramientas y presupuesto.

• Cumplimiento normativo, apoyando la implementación de estándares como ISO/IEC 27001.

• Mejora continua, a través de mecanismos de retroalimentación que permiten optimizar configuraciones y controles de seguridad.

Un caso práctico permite evidenciar estos beneficios. En una empresa de telecomunicaciones que implementa ITIL en su centro de soporte técnico, se establece como objetivo estratégico alcanzar una disponibilidad del 99.9% en sus servicios de Internet corporativo. Posteriormente, se diseñan acuerdos de nivel de servicio (SLA) y se implementa redundancia en los equipos. Durante la fase de transición, se prueban configuraciones antes de su despliegue. En la operación, un centro de monitoreo de red (NOC) supervisa continuamente la infraestructura. Finalmente, tras enfrentar un ataque DDoS, la organización incorpora nuevas medidas de seguridad, como firewalls avanzados y reglas de mitigación, dentro de un proceso de mejora continua. Como resultado, se logra una reducción significativa de incidentes críticos y un aumento en la satisfacción de los clientes.

Finalmente, ITIL mantiene una estrecha relación con otros marcos y estándares de referencia en tecnologías de la información:

• ISO/IEC 20000: estándar internacional basado en ITIL para la gestión de servicios de TI.

• COBIT: enfocado en el gobierno de TI, complementando a ITIL, que se orienta a la operación y entrega de servicios.

• NIST y CIS Controls: marcos más especializados en ciberseguridad, mientras que ITIL ofrece un enfoque más amplio en la gestión de servicios tecnológicos.

De esta manera ITIL se posiciona como un marco fundamental (Tabla 2.4) para la gestión eficiente de servicios de TI, permitiendo integrar buenas prácticas operativas con estrategias de ciberseguridad y mejora continua.

Tabla 2.4. Comparación de marcos de gobernanza.

Marco	Enfoque principal	Aplicación
COBIT	Control y auditoría de TI	Políticas de seguridad y métricas
NIST CSF	Gestión de riesgos y ciberseguridad	Empresas privadas y sectores críticos
ITIL	Gestión de servicios TI	Procesos de soporte y operación

Las políticas de seguridad constituyen documentos normativos fundamentales dentro de la gobernanza de la ciberseguridad, ya que establecen las reglas, directrices y responsabilidades que deben seguir tanto los usuarios como los administradores de sistemas. Estas políticas funcionan como un marco interno que define cómo debe protegerse la información, qué prácticas están permitidas y cuáles están estrictamente prohibidas, garantizando así un comportamiento coherente y alineado con los objetivos de seguridad de la organización.

Para que una política sea eficaz, debe cumplir con ciertas características esenciales. En primer lugar, debe ser clara y comprensible, de modo que todos los usuarios puedan interpretarla sin ambigüedades. Asimismo, debe ser aplicable a toda la organización, independientemente del nivel jerárquico o función del usuario. Es fundamental que esté basada en un análisis real de riesgos, lo que permite que las medidas establecidas respondan a amenazas concretas y no a supuestos teóricos. Finalmente, debe estar respaldada por mecanismos de control y sanción en caso de incumplimiento, asegurando su cumplimiento efectivo.

Dentro de una organización, existen diversos tipos de políticas que abordan distintos aspectos de la seguridad. Entre las más comunes se encuentra la política de contraseñas, que establece requisitos de longitud, complejidad y periodicidad de cambio. La política de uso de equipos define las prácticas permitidas en los dispositivos corporativos, como la restricción de instalación de software no autorizado. Por su parte, la política de redes inalámbricas regula aspectos como el uso de protocolos seguros (por ejemplo, WPA3) y la segmentación de redes para invitados. La política de copias de seguridad establece la frecuencia, los métodos y los mecanismos de almacenamiento seguro de los datos. Asimismo, la política de acceso remoto define los protocolos y mecanismos de autenticación necesarios para conexiones externas, mientras que la política de respuesta a incidentes establece los procedimientos a seguir ante eventos de seguridad.

Un ejemplo representativo es la política de contraseñas, cuyo objetivo es proteger las cuentas de usuario frente a accesos indebidos. Esta puede incluir lineamientos como una longitud mínima de 12 caracteres, el uso combinado de mayúsculas, minúsculas, números y símbolos, la prohibición de utilizar información personal, la obligatoriedad de cambiar la contraseña cada cierto periodo, el bloqueo de cuentas tras múltiples intentos fallidos y la implementación de autenticación multifactor. Estas medidas, en conjunto, reducen significativamente el riesgo de accesos no autorizados.

La implementación de una adecuada gobernanza y políticas de seguridad aporta múltiples beneficios a la organización. En primer lugar, mejora la seguridad técnica al prevenir configuraciones inseguras y accesos indebidos. Además, facilita el cumplimiento normativo, permitiendo a la organización alinearse con estándares internacionales como ISO/IEC 27001 y superar auditorías externas. También contribuye a reducir los riesgos asociados al factor humano, minimizando errores comunes como el uso de contraseñas débiles o la instalación de software no autorizado. Asimismo, fortalece la cultura organizacional, promoviendo la

concienciación en materia de seguridad entre todos los miembros de la organización.

Por el contrario, la ausencia de gobernanza y políticas de seguridad puede generar consecuencias graves. Entre ellas se encuentran el uso de credenciales por defecto en dispositivos críticos, la pérdida de información sensible debido a la falta de copias de seguridad, la vulnerabilidad frente a ataques por ausencia de segmentación de redes y el incumplimiento de normativas legales de protección de datos. Además, estas deficiencias pueden afectar la reputación de la organización, generando desconfianza entre clientes y usuarios.

Un caso real ilustra estas consecuencias: diversas empresas han sido víctimas de ataques de ransomware debido a la ausencia de políticas adecuadas de copias de seguridad. En muchos casos, aunque existían respaldos, estos no estaban cifrados ni actualizados, lo que impidió la recuperación de la información y generó pérdidas significativas.

Es importante destacar que la gobernanza no sustituye las tareas técnicas realizadas por administradores de sistemas o redes, sino que las complementa. Mientras los técnicos implementan configuraciones específicas, como reglas en un firewall, la gobernanza asegura que dichas configuraciones estén documentadas, auditadas y alineadas con las políticas organizacionales.

Finalmente, la implementación efectiva de políticas de seguridad requiere la adopción de buenas prácticas. Entre ellas se encuentra la participación de todas las áreas de la organización en su definición, la actualización periódica de las políticas en función de nuevas amenazas, la capacitación continua de usuarios y técnicos, la documentación clara de roles y responsabilidades, y la realización de auditorías internas que permitan verificar su cumplimiento.

2.5. Normas internacionales, análisis y auditoría de la seguridad de la información

La seguridad de la información se ha consolidado como un elemento esencial para el funcionamiento, la

competitividad y la sostenibilidad de las organizaciones modernas. En un entorno marcado por la rápida evolución de las tecnologías digitales, la globalización de los mercados y el aumento en la complejidad y frecuencia de los ciberataques, resulta imprescindible adoptar enfoques estructurados que permitan proteger los activos más críticos: la información y los datos. En este contexto, los estándares internacionales adquieren un papel fundamental, ya que proporcionan marcos de referencia, metodologías y buenas prácticas que facilitan una gestión sistemática y eficaz de la seguridad. Diversos estudios destacan que la adopción de estos estándares no solo mejora la protección frente a amenazas, sino que también contribuye a la generación de valor a partir de los datos y al fortalecimiento de la confianza en los entornos digitales (Kitsios et al., 2023; Reuben-Owoh y Haig, 2025).

Dentro de estos estándares, destacan especialmente las normas ISO/IEC 27001 e ISO/IEC 27032, que ofrecen un enfoque complementario para abordar la seguridad desde diferentes perspectivas. La ISO/IEC 27001 se centra en la implantación de un sistema de gestión de seguridad de la información basado en el análisis y tratamiento de riesgos, lo que permite a las organizaciones establecer controles eficaces, mejorar la toma de decisiones y optimizar la gestión de sus activos informacionales. Investigaciones recientes evidencian que su implementación tiene un impacto positivo en la organización, mejorando la gestión del conocimiento, la protección de la información y el aprovechamiento estratégico de los datos (Kitsios et al., 2023: Qusef y Alkilani, 2022).

Por su parte, la ISO/IEC 27032 aborda de manera específica la ciberseguridad, centrándose en la protección frente a amenazas propias del entorno digital, como ataques maliciosos, fraude en línea o ingeniería social. Este enfoque resulta especialmente relevante en contextos altamente interconectados, donde los riesgos no solo afectan a sistemas individuales, sino también a ecosistemas completos de información. En este sentido, la literatura subraya la importancia de

combinar estándares de gestión como ISO/IEC 27001 con marcos específicos de ciberseguridad para lograr una protección más completa y adaptada a las nuevas amenazas (Amine et al., 2023).

Asimismo, la integración de estos estándares facilita el desarrollo de procesos de auditoría y evaluación más robustos, permitiendo a las organizaciones medir el nivel de cumplimiento, identificar debilidades y establecer mejoras continuas en sus sistemas de seguridad. En línea con esto, algunos enfoques recientes proponen marcos de auditoría centrados en el usuario y en la gestión integral del riesgo, lo que refuerza la eficacia de las políticas de seguridad y su alineación con los objetivos organizacionales (Antunes et al., 2022).

La norma ISO/IEC 27001 es un estándar internacional desarrollado por la Organización Internacional de Normalización y la Comisión Electrotécnica Internacional, que establece los requisitos necesarios para implementar, mantener y mejorar un Sistema de Gestión de Seguridad de la Información. Este sistema se basa en un enfoque sistemático que permite identificar, analizar y tratar los riesgos relacionados con la información, garantizando así la protección de su confidencialidad, integridad y disponibilidad (International Organization for Standardization & International Electrotechnical Commission, 2022).

Entre los principales objetivos de la ISO/IEC 27001 se encuentran la protección de la información frente a accesos no autorizados, la gestión adecuada de los riesgos de seguridad, la definición de políticas y controles claros, y el fortalecimiento de la confianza entre las partes interesadas, incluyendo clientes, proveedores y socios comerciales. Para lograr estos objetivos, la norma establece una serie de elementos fundamentales, entre los que destacan la definición de políticas de seguridad, la evaluación continua de riesgos, la implementación de controles de seguridad organizados en múltiples dominios, la realización de auditorías internas y externas, y la mejora continua del sistema.

Uno de los aspectos más relevantes de esta norma es su enfoque basado en el ciclo de mejora continua conocido como Planificar, Hacer, Verificar y Actuar. Este ciclo permite a las organizaciones gestionar la seguridad de manera dinámica y adaptativa. En la fase de planificación se identifican los riesgos y se establecen las políticas de seguridad; en la fase de ejecución se implementan los controles necesarios; en la fase de verificación se evalúa el cumplimiento y la eficacia de las medidas adoptadas; y en la fase de actuación se aplican mejoras y acciones correctivas que fortalecen el sistema.

La implementación de la ISO/IEC 27001 aporta múltiples beneficios a las organizaciones, entre los que destacan la reducción de riesgos de seguridad, el cumplimiento de normativas legales y regulatorias, el incremento de la confianza de los clientes y socios comerciales, la protección frente a pérdidas económicas derivadas de incidentes de seguridad y una mayor eficiencia en la gestión de las tecnologías de la información. Por ejemplo, una empresa del sector de telecomunicaciones puede aplicar esta norma para proteger sus bases de datos mediante técnicas como el cifrado de la información, la segmentación de redes, el control de accesos y la realización periódica de auditorías.

Por su parte, la norma ISO/IEC 27032 se enfoca específicamente en la ciberseguridad, entendida como la protección del ciberespacio frente a amenazas digitales como ataques maliciosos, espionaje, fraude y otras formas de ciberdelincuencia. A diferencia de la ISO/IEC 27001, que tiene un enfoque más amplio sobre la seguridad de la información en general, la ISO/IEC 27032 profundiza en los riesgos asociados al entorno en línea y en la protección de los sistemas interconectados (International Organization for Standardization & International Electrotechnical Commission, 2012).

Esta norma aborda áreas clave como la protección contra ciberataques, la seguridad en el intercambio de información a través de redes, la prevención de ataques de ingeniería social, la protección frente a fraudes en Internet y la promoción de la cooperación internacional en materia de ciberseguridad. Su enfoque reconoce

que las amenazas actuales no solo afectan a sistemas individuales, sino que pueden tener un impacto global, por lo que resulta esencial adoptar medidas coordinadas entre organizaciones, sectores y países.

Un ejemplo de aplicación de la ISO/IEC 27032 puede observarse en una empresa que busca fortalecer su protección frente a ataques de phishing dirigidos a sus empleados. En este caso, la organización puede implementar programas de concienciación, filtros avanzados de correo electrónico y mecanismos de detección temprana de amenazas, reduciendo así el riesgo de comprometer credenciales o información sensible.

Las normas ISO/IEC 27001 e ISO/IEC 27032 proporcionan un enfoque integral (Tabla 2.5) que combina la gestión estructurada de la seguridad de la información con la protección específica frente a amenazas del ciberespacio. Su implementación permite a las organizaciones no solo prevenir incidentes, sino también mejorar su capacidad de respuesta y recuperación, fortaleciendo su resiliencia en un entorno digital cada vez más complejo y desafiante.

Tabla 2.5. Comparación entre ISO/IEC 27001 e ISO/IEC 27032.

Aspecto	ISO/IEC 27001	ISO/IEC 27032
Enfoque	Seguridad de la información en general	Ciberseguridad y amenazas en línea
Objetivo principal	Confidencialidad, integridad, disponibilidad	Protección del ciberespacio
Alcance	Organizaciones y procesos internos	Usuarios, sistemas y entornos digitales
Certificación	Sí, certificable	No, es guía de referencia

Aunque las normas ISO/IEC 27001 e ISO/IEC 27032 presentan enfoques distintos, ambas son altamente complementarias y, en conjunto, proporcionan un marco integral para la protección de la información en entornos organizacionales. La ISO/IEC 27001 establece las bases para la gestión estructurada de la seguridad de la información mediante la implementación de políticas, controles y procesos orientados a la mitigación de

riesgos. Por su parte, la ISO/IEC 27032 amplía esta visión al abordar de manera específica los riesgos asociados al ciberespacio, incluyendo amenazas externas como el malware, el phishing, el cibercrimen y los ataques dirigidos a infraestructuras digitales.

La integración de ambas normas permite a las organizaciones adoptar un enfoque más completo, que no solo considera la protección interna de la información, sino también la defensa frente a amenazas externas cada vez más sofisticadas. Por ejemplo, una empresa que ha implementado un sistema de gestión de seguridad basado en la ISO/IEC 27001 puede apoyarse en la ISO/IEC 27032 para fortalecer áreas críticas como la seguridad del correo electrónico, la protección en redes sociales y la gestión de servicios en la nube.

Beneficios de la aplicación conjunta

La adopción combinada de ambas normas ofrece múltiples ventajas para las organizaciones:

- Permite una mejora integral de la seguridad de la información, abarcando tanto aspectos internos como externos.
- Reduce significativamente las vulnerabilidades frente a amenazas, tanto internas como provenientes del entorno digital.
- Facilita el cumplimiento de normativas legales y estándares internacionales en materia de protección de datos.
- Incrementa la confianza de clientes, proveedores y socios comerciales.
- Fortalece la capacidad de respuesta ante ciberamenazas avanzadas, como ataques persistentes, ransomware o fraudes en línea.

En el ámbito empresarial, la combinación de ambas normas puede observarse en distintos escenarios. Por ejemplo, una empresa de desarrollo de software puede implementar la ISO/IEC 27001 para proteger su código fuente mediante controles de acceso, cifrado y auditorías, mientras que utiliza la ISO/IEC 27032 para prevenir

ataques de phishing dirigidos a sus desarrolladores mediante programas de concienciación y sistemas de filtrado.

De igual manera, un proveedor de servicios en la nube puede aplicar la ISO/IEC 27001 para asegurar sus centros de datos mediante mecanismos como redundancia, cifrado y control de accesos, y complementar esta protección con la ISO/IEC 27032 para mitigar riesgos asociados a ataques de ingeniería social o intentos de robo de información a través de medios digitales.

El análisis de seguridad y las auditorías constituyen herramientas esenciales dentro de la gestión de la ciberseguridad, ya que permiten identificar, evaluar y mitigar riesgos en sistemas de información y redes. Su principal objetivo es garantizar que las medidas de seguridad implementadas sean adecuadas, eficaces y estén alineadas con estándares, políticas y normativas vigentes.

Es importante diferenciar ambos conceptos. El análisis de seguridad se enfoca en la identificación de riesgos, amenazas y vulnerabilidades presentes en los sistemas, mientras que la auditoría de seguridad corresponde a un proceso sistemático de revisión que evalúa el grado de cumplimiento de políticas internas, normativas externas y buenas prácticas internacionales.

El análisis de seguridad consiste en un estudio detallado de los activos de información, las amenazas potenciales y las vulnerabilidades existentes, con el propósito de evaluar los riesgos y establecer controles adecuados. Este proceso se desarrolla generalmente a través de una serie de etapas estructuradas:

- Identificación de activos críticos, como servidores, routers, switches y estaciones de trabajo.
- Detección de amenazas, incluyendo malware, accesos no autorizados o fallos físicos.
- Identificación de vulnerabilidades, tales como configuraciones inseguras, contraseñas débiles o falta de actualizaciones.

- Evaluación del riesgo, considerando tanto el impacto como la probabilidad de ocurrencia.
- Definición de controles de seguridad, como el uso de firewalls, sistemas de detección de intrusos o segmentación de redes.

Un ejemplo práctico de este proceso se presenta cuando un análisis de seguridad detecta que varios dispositivos de red aún utilizan credenciales por defecto. Esta situación representa un riesgo elevado, ya que facilita accesos no autorizados. Como medida correctiva inmediata, se establece el cambio de credenciales y la implementación de mecanismos de autenticación centralizada, reduciendo significativamente la exposición al Riesgo (Tabla 2.6).

Tabla 2.6. Matriz de riesgos en auditoría.

Activo	Vulnerabilidad	Amenaza	Impacto	Probabilidad	Nivel de riesgo	Control recomendado
Router principal	Credenciales por defecto	Acceso no autorizado	Alto	Alta	Crítico	Cambiar credenciales y activar SSH
Servidor web	Sin parches de seguridad	Exploits conocidos	Alto	Media	Alto	Actualización y parches
Estación de trabajo	USB sin control	Malware externo	Medio	Alta	Alto	Desactivar puertos USB
Switch de acceso	Sin VLANs configuradas	Sniffing de tráfico	Medio	Media	Medio	Segmentación de red

Existen diversos marcos y metodologías reconocidos a nivel internacional que permiten llevar a cabo análisis de seguridad de manera estructurada y sistemática. Estas metodologías proporcionan lineamientos, herramientas y enfoques que facilitan la identificación de riesgos, la evaluación de vulnerabilidades y la implementación de controles adecuados para la protección de los sistemas de información.

Entre las metodologías más destacadas se encuentran:

• OCTAVE (Operationally Critical Threat, Asset, and Vulnerability Evaluation): desarrollada por la Universidad Carnegie Mellon, esta metodología se centra en la identificación de activos críticos y en el análisis de las amenazas y vulnerabilidades que pueden afectar a una organización. Su enfoque está orientado a la gestión estratégica del riesgo, involucrando tanto aspectos técnicos como organizacionales.

• MAGERIT: es una metodología de origen español, ampliamente utilizada en el ámbito de la administración pública. Su objetivo principal es analizar y gestionar los riesgos en los sistemas de información mediante un enfoque estructurado que combina la identificación de activos, amenazas y salvaguardas, con un fuerte énfasis en la documentación y la trazabilidad.

• NIST SP 800-30: se trata de una guía desarrollada por el Instituto Nacional de Estándares y Tecnología de los Estados Unidos, que proporciona un enfoque detallado y práctico para la gestión de riesgos de seguridad. Esta metodología describe paso a paso cómo identificar amenazas, evaluar vulnerabilidades, estimar impactos y determinar niveles de riesgo.

• ISO/IEC 27005: es una norma internacional que complementa la ISO/IEC 27001 y se enfoca específicamente en la gestión de riesgos de seguridad de la información. Proporciona un marco metodológico que permite a las organizaciones integrar el análisis de riesgos dentro de su sistema de gestión de seguridad, asegurando coherencia con los estándares internacionales.

Estas metodologías ofrecen diferentes enfoques que pueden adaptarse a las necesidades específicas de cada organización, permitiendo establecer procesos de análisis de seguridad más robustos, coherentes y alineados con buenas prácticas internacionales (Tabla 2.7).

Tabla 2.7. Comparación de metodologías.

Metodología	Enfoque principal	Ámbito típico de uso
OCTAVE	Identificación de activos críticos y amenazas	Empresas privadas y militares
MAGERIT	Evaluación de riesgos con énfasis documental	Gobiernos y sector público
NIST SP 800-30	Guía práctica paso a paso	Todo tipo de organizaciones
ISO/IEC 27005	Complemento normativo a ISO 27001	Organizaciones que buscan certificación ISO

La auditoría de seguridad es un proceso sistemático, independiente y documentado cuyo objetivo es obtener evidencias objetivas para evaluar si los sistemas, procesos y controles de una organización cumplen con:

- Las políticas internas establecidas.
- Las normas y leyes aplicables, como ISO/IEC 27001 o la legislación sobre protección de datos.
- Los estándares internacionales y las buenas prácticas en materia de seguridad.

Existen distintos tipos de auditorías, según su enfoque y propósito:

- Auditoría interna: realizada por personal de la propia organización para revisar y mejorar sus controles de seguridad.
- Auditoría externa: llevada a cabo por una entidad independiente, lo que aporta mayor objetividad.

- Auditoría de cumplimiento: verifica si la organización cumple con normativas, regulaciones y requisitos legales.
- Auditoría técnica: se centra en aspectos operativos y tecnológicos, como configuraciones, pruebas de penetración y análisis de redes.

Ejemplo técnico

Durante una auditoría de red, se detecta que varios switches aún utilizan la contraseña predeterminada admin/admin. Esta situación representa una vulnerabilidad crítica, ya que facilita accesos no autorizados y requiere una corrección inmediata.

Para llevar a cabo auditorías de seguridad, se emplean diversas herramientas especializadas, entre las que destacan:

- Nmap: permite escanear puertos y detectar servicios abiertos en equipos y redes.
- Nessus / OpenVAS: se utilizan para identificar vulnerabilidades en sistemas, dispositivos y redes.
- Wireshark: facilita el análisis del tráfico de red en tiempo real.
- Metasploit: framework orientado a la realización de pruebas de penetración.
- Kali Linux: distribución especializada que integra múltiples herramientas para auditoría y análisis de seguridad.

Ejemplo práctico

Un técnico utiliza Nmap para identificar los puertos abiertos de un router corporativo y descubre que el servicio Telnet, considerado inseguro, sigue habilitado. Como resultado, la auditoría recomienda desactivarlo y sustituirlo por SSH, un protocolo más seguro para la administración remota.

El proceso de auditoría de seguridad comienza con una fase de planificación, en la que se define el alcance

del análisis. En esta etapa se determinan los sistemas, redes, aplicaciones y procesos que serán evaluados, estableciendo así los límites y objetivos de la auditoría.

A continuación, se lleva a cabo la recopilación de información, que consiste en identificar y documentar los activos de la organización, así como sus configuraciones, usuarios, servicios y controles de seguridad existentes. Esta fase proporciona una visión completa del entorno a analizar.

Posteriormente, se desarrollan las pruebas técnicas, donde se aplican herramientas y técnicas como el análisis de vulnerabilidades, las pruebas de penetración y la revisión de registros o logs. El objetivo es detectar posibles fallos o debilidades en los sistemas.

En la fase de evaluación, los resultados obtenidos se analizan y se comparan con las políticas internas de la organización, así como con las normativas y estándares de seguridad aplicables, con el fin de determinar el nivel de cumplimiento.

Seguidamente, se elabora el informe de auditoría, en el que se recogen los hallazgos detectados, se evalúan los riesgos asociados y se proponen recomendaciones y medidas correctivas para mejorar la seguridad.

Finalmente, se realiza el seguimiento, cuyo propósito es verificar que las acciones correctivas han sido implementadas adecuadamente y comprobar su eficacia en la reducción de riesgos.

La seguridad de la información es esencial para las organizaciones actuales debido al aumento de amenazas digitales y la dependencia tecnológica. Para gestionarla de forma eficaz, se utilizan estándares internacionales como ISO/IEC 27001, centrada en la gestión global de la seguridad, y ISO/IEC 27032, orientada a la protección frente a riesgos del ciberespacio. Ambas normas, complementarias, permiten proteger sistemas, datos y usuarios mediante un enfoque estructurado y adaptativo basado en la gestión de riesgos y la mejora continua.

El análisis de seguridad y la auditoría son procesos clave dentro de esta gestión. El primero identifica activos,

amenazas y vulnerabilidades para evaluar riesgos y definir controles, mientras que la auditoría verifica el cumplimiento de políticas, normativas y buenas prácticas mediante revisiones sistemáticas.

Para realizar estos procesos se emplean metodologías reconocidas como OCTAVE, MAGERIT, NIST SP 800-30 e ISO/IEC 27005, que aportan marcos estructurados para analizar riesgos. Asimismo, se utilizan herramientas técnicas como Nmap, Wireshark o Metasploit para detectar vulnerabilidades y evaluar sistemas.

03.

Enfoques ofensivos y defensivos en ciberseguridad y su aplicación en la protección digital

3.1. Hacking ético y pruebas de penetración: fundamentos, técnicas y aplicaciones

El término *hacker* ha sido objeto de múltiples interpretaciones a lo largo de las últimas décadas. En el discurso mediático, suele asociarse con actividades ilícitas, delincuencia informática y ataques cibernéticos. Sin embargo, en los ámbitos académico y profesional, este concepto adquiere un significado distinto, aludiendo a un perfil altamente capacitado en informática, con amplios conocimientos en sistemas, redes y seguridad. Comprender con precisión qué implica ser un hacker resulta fundamental para diferenciar entre prácticas ilegales y aquellas orientadas al fortalecimiento de la seguridad de los sistemas.

Desde una perspectiva histórica, el concepto de hacker

ha experimentado una evolución significativa. Durante las décadas de 1960 y 1970, los primeros hackers surgieron en entornos universitarios, particularmente en instituciones como el MIT, donde se destacaban por su creatividad e interés en mejorar sistemas informáticos. En la década de 1980, con la expansión de las computadoras personales y el desarrollo de redes de comunicación, el término comenzó a popularizarse, aunque ya acompañado de connotaciones negativas. Posteriormente, en la década de 1990, el auge de Internet consolidó la asociación del hacker con el acceso no autorizado a sistemas. En la actualidad, dentro del campo de la ciberseguridad, el concepto de *hacker* ético se refiere a un profesional autorizado que identifica y corrige vulnerabilidades antes de que puedan ser explotadas por actores maliciosos.

En cuanto a su clasificación, los hackers pueden agruparse en distintos niveles según sus capacidades, motivaciones y métodos. En un nivel básico se encuentran los denominados *script kiddies*, individuos con escasas habilidades técnicas que utilizan herramientas desarrolladas por otros para ejecutar ataques simples, generalmente motivados por reconocimiento o notoriedad.

En un nivel intermedio se ubican los hackers tradicionalmente clasificados según su ética y propósito. Los hackers de sombrero blanco (*white hat*) emplean sus conocimientos con fines legales y preventivos, realizando auditorías de seguridad y pruebas de penetración. Por el contrario, los hackers de sombrero negro (*black hat*) explotan vulnerabilidades con fines delictivos, tales como el robo de información, fraude o espionaje. Entre ambos se encuentran los hackers de sombrero gris (*grey hat*), quienes operan en una zona intermedia, identificando fallos sin autorización, aunque en ocasiones los revelan sin intenciones maliciosas.

En un nivel avanzado se identifican actores con motivaciones más complejas y estructuradas. Los hacktivistas utilizan técnicas de hacking con fines políticos o sociales, buscando visibilizar causas o denunciar vulnerabilidades. Asimismo, existen

ciberdelincuentes respaldados por Estados, orientados a la obtención de información estratégica y ventajas geopolíticas. Por último, los ciberterroristas emplean herramientas digitales con el objetivo de generar caos, afectar infraestructuras críticas y provocar temor en la sociedad, en consonancia con los fines del terrorismo tradicional.

El hacking ético se ha consolidado como una disciplina profesional esencial dentro del ámbito de la ciberseguridad, cuyo propósito principal es anticiparse a posibles ataques mediante la identificación y corrección de vulnerabilidades en los sistemas de información. En un entorno digital cada vez más complejo y expuesto a amenazas sofisticadas, las organizaciones recurren a profesionales especializados para simular ataques controlados que permitan evaluar el nivel de seguridad de sus infraestructuras tecnológicas. De esta manera, el hacking ético se define como la práctica de realizar pruebas de penetración autorizadas con el fin de detectar debilidades antes de que puedan ser explotadas por actores maliciosos. Estas evaluaciones no solo abarcan aspectos técnicos, sino también configuraciones, políticas de seguridad y posibles fallos humanos que podrían comprometer la integridad de los sistemas.

Diversos estudios destacan que estas prácticas permiten fortalecer significativamente las estrategias de defensa al adoptar un enfoque proactivo y preventivo frente a las amenazas emergentes (Mandru, 2020; Zhang et al., 2026). En particular, las metodologías de pruebas de penetración han evolucionado hacia modelos más estructurados que incluyen fases como reconocimiento, escaneo, explotación y post explotación, lo que facilita una evaluación integral del sistema. Asimismo, la literatura reciente resalta que las pruebas de penetración constituyen un componente clave en la gestión de riesgos en redes, ya que permiten medir el nivel de exposición ante ataques reales y priorizar acciones correctivas basadas en el impacto potencial (Alhamed y Rahman, 2023). En este sentido, el hacking ético no solo actúa como un mecanismo de detección, sino también como

una herramienta estratégica para la mejora continua de la seguridad organizacional.

Para desempeñarse eficazmente en este campo, un hacker ético debe poseer un conjunto de habilidades técnicas y personales altamente desarrolladas. Entre ellas destacan el dominio de herramientas especializadas como escáneres de vulnerabilidades, marcos de explotación y analizadores de tráfico, así como la capacidad de desarrollar soluciones propias adaptadas a contextos específicos. Además, es fundamental contar con conocimientos sólidos en diversos lenguajes de programación, sistemas operativos, redes y arquitectura de sistemas, lo que permite comprender en profundidad el funcionamiento de los entornos evaluados. Esta combinación de habilidades técnicas facilita la identificación de vulnerabilidades complejas que no siempre son detectadas por herramientas automatizadas.

A nivel cognitivo, se requiere pensamiento analítico, atención al detalle y una alta capacidad de resolución de problemas, así como creatividad para identificar posibles vectores de ataque no convencionales. Los hackers éticos deben ser capaces de adoptar la perspectiva de un atacante, anticipando sus estrategias y adaptándose a escenarios dinámicos y cambiantes. Asimismo, la confiabilidad y la ética profesional son fundamentales, dado que estos profesionales manejan información altamente sensible y, en muchos casos, acceden a sistemas críticos. En este sentido, la ética adquiere un papel central, especialmente en áreas como la ingeniería social, donde las pruebas pueden involucrar la interacción directa con usuarios. Estas prácticas deben equilibrar la efectividad técnica con el respeto a los principios morales, la privacidad y la dignidad de las personas (Hatfield, 2019).

Además, se ha señalado que el valor del hacker ético radica en su capacidad de pensar como un atacante sin comprometer la integridad profesional, lo cual implica actuar dentro de marcos legales y normativos estrictos. Este aspecto resulta particularmente relevante en sectores críticos como el de la salud, donde la protección de datos sensibles es prioritaria y cualquier

vulnerabilidad puede tener consecuencias graves para los usuarios (Lorenzini et al., 2022). En consecuencia, el hacking ético no solo requiere competencias técnicas avanzadas, sino también un fuerte compromiso con la responsabilidad social, la confidencialidad y el cumplimiento de estándares éticos y legales.

Los beneficios del hacking ético son múltiples y de gran relevancia para las organizaciones. En primer lugar, permite descubrir vulnerabilidades antes de que sean explotadas, reduciendo significativamente el riesgo de incidentes de seguridad. Además, contribuye al cumplimiento de normativas y estándares internacionales, fortalece la confianza de clientes y socios, y facilita la capacitación del personal técnico en la gestión y respuesta ante incidentes.

El ejercicio del hacking ético se rige por principios fundamentales que garantizan su legitimidad y responsabilidad. Entre ellos se encuentran la necesidad de contar con autorización expresa antes de realizar cualquier prueba, la obligación de documentar de manera detallada todos los hallazgos, el mantenimiento de la confidencialidad de la información obtenida y la actuación bajo estrictos criterios de ética profesional.

Es importante distinguir claramente entre hacking ético y cibercrimen. Mientras que el primero se desarrolla en un marco autorizado y regulado, con el objetivo de mejorar la seguridad y colaborar con las organizaciones, el segundo se caracteriza por la ausencia de consentimiento, la búsqueda de beneficios ilícitos o el daño reputacional, así como el encubrimiento de las acciones realizadas.

En cuanto a las técnicas utilizadas, el hacking ético emplea diversos métodos para evaluar la seguridad de los sistemas. Entre los más comunes se encuentran el escaneo de puertos para identificar servicios abiertos, la explotación controlada de vulnerabilidades, las pruebas de fuerza bruta para evaluar contraseñas, la detección de fallos mediante inyección SQL en aplicaciones web y el análisis de tráfico mediante técnicas de sniffing.

El hacking ético tiene aplicaciones prácticas en múltiples sectores. Por ejemplo, en empresas de

telecomunicaciones se utilizan simulaciones de ataques de denegación de servicio (DDoS) para evaluar la resiliencia de la red; en instituciones financieras se realizan campañas controladas de phishing para medir la respuesta de los empleados; y en centros de datos se ejecutan evaluaciones tanto físicas como lógicas para detectar accesos no autorizados.

El perfil de un hacker ético profesional se caracteriza por un alto nivel de especialización técnica, que incluye el dominio de sistemas operativos como Linux, Windows y Unix, el conocimiento de protocolos de red como TCP/IP, HTTP, DNS y SMTP, y el manejo de herramientas de seguridad como Metasploit, Nmap, Wireshark y Burp Suite. A estas competencias se suma la capacidad de análisis crítico y la habilidad de pensar como un atacante, siempre dentro de un marco ético y legal.

Las fases del hacking ético constituyen una metodología sistemática que permite evaluar de manera organizada la seguridad de los sistemas de información. Este proceso se desarrolla en varias etapas secuenciales, cada una con objetivos específicos que contribuyen a la identificación y validación de vulnerabilidades.

En primer lugar, se encuentra la fase de definición del alcance, en la cual se establecen los sistemas, redes y procesos que serán objeto de evaluación. En esta etapa es fundamental contar con la autorización expresa del propietario de los sistemas, así como definir claramente los objetivos, los límites de la prueba y las condiciones bajo las cuales se llevará a cabo. Esta planificación inicial garantiza que las actividades se desarrollen dentro de un marco legal y controlado.

A continuación, se realiza el reconocimiento pasivo, que consiste en la recopilación de información sin interactuar directamente con los sistemas objetivo. Esta fase se basa en el uso de fuentes abiertas y técnicas de footprinting, que permiten obtener datos relevantes sobre la infraestructura tecnológica y organizacional. Entre la información recolectada se incluyen nombres de dominio, direcciones IP, servicios de red, características del sistema, así como datos públicos de la organización,

como portales web, información de empleados y políticas de seguridad. Para ello, se emplean herramientas como Netcraft, Maltego, Kali Linux y técnicas de Google Hacking.

Posteriormente, se lleva a cabo el reconocimiento activo, en el cual el evaluador interactúa directamente con los sistemas para obtener información más detallada. En esta fase se utilizan herramientas como Nmap para escanear puertos, identificar servicios activos y detectar sistemas operativos. Asimismo, se aplican técnicas de enumeración que permiten profundizar en el conocimiento de los servicios y aplicaciones disponibles, facilitando la identificación de posibles puntos de entrada.

La siguiente etapa es la explotación, donde se intenta aprovechar de manera controlada las vulnerabilidades identificadas en las fases anteriores. Este proceso puede implicar el uso de herramientas especializadas como Metasploit, así como la aplicación de técnicas de ingeniería social para obtener acceso a información confidencial. La explotación permite validar el impacto real de las vulnerabilidades y determinar el nivel de riesgo al que está expuesto el sistema.

En conjunto, estas fases permiten estructurar el proceso de hacking ético de manera lógica y eficiente, asegurando una evaluación integral de la seguridad y proporcionando información clave para la implementación de medidas correctivas (Figura 3.1).

Figura 3.1. Enumeración y explotación de servicios/vulnerabilidades.

Dentro del proceso de hacking ético, la fase de reconocimiento y explotación incluye diversas técnicas orientadas a la identificación de sistemas activos, la recopilación de información y la validación de vulnerabilidades. Estas técnicas permiten obtener una visión detallada de la infraestructura objetivo y constituyen la base para una evaluación efectiva de la seguridad.

Una de las primeras actividades es la identificación de hosts activos, comúnmente realizada mediante la técnica conocida como ICMP Sweep. Este procedimiento consiste en el envío de solicitudes ICMP (ICMP Echo Request) a múltiples direcciones IP dentro de una red, con el objetivo de determinar cuáles dispositivos se encuentran operativos. Los sistemas activos responden mediante mensajes ICMP Echo Reply, lo que permite al evaluador identificar los hosts disponibles. Herramientas como Nmap también pueden emplearse para este propósito, facilitando la detección de dispositivos y servicios en red.

Otra técnica relevante es el banner grabbing, utilizada para recopilar información sobre sistemas remotos. Esta consiste en obtener los encabezados o “banners” que los servicios envían durante el proceso de conexión, los cuales pueden contener datos como versiones de software, configuraciones y características del sistema operativo. Esta información resulta valiosa para identificar posibles vulnerabilidades asociadas a versiones específicas de software.

El escaneo de puertos TCP/UDP constituye una técnica fundamental para identificar servicios activos en un sistema. A través de este proceso se determinan los puertos abiertos, filtrados o cerrados, lo que permite inferir qué servicios están disponibles y potencialmente expuestos. Existen diferentes métodos de escaneo, como el TCP Connect Scan, que establece conexiones completas y es fácilmente detectable, y técnicas más sigilosas como el Xmas Scan o el Null Scan, diseñadas para evadir mecanismos de detección.

En este contexto, también se aplican técnicas de evasión de sistemas de detección de intrusos (IDS), cuyo objetivo es evitar ser detectado durante el proceso de reconocimiento. Estas técnicas incluyen la fragmentación de paquetes, la manipulación de direcciones IP o el envío de tráfico anómalo para dificultar la identificación de la actividad por parte de los sistemas de seguridad.

La enumeración representa una fase más avanzada de recopilación de información, en la cual se obtiene información detallada sobre servicios, usuarios y configuraciones del sistema. Entre las técnicas más comunes se encuentran la enumeración SMTP para identificar cuentas de correo, la enumeración de servicios y puertos para conocer los recursos disponibles, así como la enumeración de protocolos como SNMP, LDAP y NTP. Asimismo, en sistemas Linux/UNIX, esta fase permite obtener información específica sobre usuarios, procesos y configuraciones del sistema.

En la etapa de explotación, se emplean herramientas especializadas como Metasploit, una plataforma de código abierto diseñada para identificar, validar y explotar vulnerabilidades. Metasploit utiliza exploits, que son mecanismos para aprovechar fallos de seguridad, y payloads, que representan las acciones ejecutadas tras la explotación, como la obtención de acceso al sistema. Esta herramienta permite demostrar de manera controlada el impacto real de una vulnerabilidad, facilitando su posterior mitigación.

Finalmente, el proceso culmina con las fases de análisis y reporte. En la fase de análisis se evalúan los resultados obtenidos durante la prueba, identificando las debilidades y determinando su nivel de riesgo. Posteriormente, en la fase de reporte, se elaboran informes detallados que incluyen tanto un enfoque gerencial, orientado a la toma de decisiones, como un informe técnico que documenta los hallazgos, las evidencias, los métodos utilizados y las recomendaciones para mitigar las vulnerabilidades detectadas.

En este contexto, las pruebas de penetración o pentesting se consolidan como una de las herramientas

más relevantes dentro de la seguridad ofensiva. Estas consisten en la simulación controlada de ciberataques con el objetivo de evaluar el nivel de exposición de una organización, identificar vulnerabilidades y proponer medidas correctivas. A diferencia de los ataques maliciosos, el pentesting se realiza bajo autorización expresa, con un enfoque ético y orientado a la mejora continua de la seguridad.

Las pruebas de penetración, o pentesting, pueden clasificarse en función del nivel de información y acceso proporcionado a los evaluadores. Esta clasificación permite adaptar el enfoque de la evaluación según los objetivos de la organización y el grado de profundidad requerido. En este sentido, se distinguen tres tipos principales.

En primer lugar, las pruebas de caja negra (Black Box) se caracterizan por la ausencia total de información previa sobre los sistemas evaluados. Este enfoque simula de manera realista un ataque externo, en el que el evaluador actúa como un atacante sin conocimiento interno. Su principal ventaja radica en el alto nivel de realismo; sin embargo, puede presentar limitaciones, ya que algunas vulnerabilidades podrían no ser identificadas debido al desconocimiento inicial del entorno.

En segundo lugar, las pruebas de caja blanca (White Box) proporcionan acceso completo a la información del sistema, incluyendo código fuente, configuraciones y documentación técnica. Este tipo de evaluación permite un análisis exhaustivo y detallado de las vulnerabilidades, lo que incrementa la precisión de los resultados. No obstante, requiere mayor tiempo, recursos y un alto nivel de especialización por parte de los evaluadores.

Por último, las pruebas de caja gris (Grey Box) representan un enfoque intermedio, en el cual se brinda información parcial al evaluador. Este modelo simula el comportamiento de un usuario con privilegios limitados y combina el realismo del enfoque de caja negra con la profundidad técnica del enfoque de caja blanca. Por esta razón, es uno de los métodos más utilizados en entornos organizacionales.

El proceso de pentesting se desarrolla mediante una metodología estructurada que garantiza la validez y confiabilidad de los resultados. La primera fase corresponde a la planificación y definición del alcance, donde se establecen los objetivos, permisos, limitaciones y sistemas a evaluar. A continuación, se realiza el reconocimiento, que implica la recolección de información tanto pública como privada mediante técnicas como Open Source Intelligence (OSINT).

Posteriormente, se lleva a cabo la fase de escaneo, en la cual se identifican puertos abiertos, servicios activos y posibles vulnerabilidades, utilizando herramientas especializadas como Nmap o Nessus. Seguidamente, en la fase de explotación, se intenta aprovechar de manera controlada las vulnerabilidades detectadas, validando su impacto real en el sistema.

Una vez obtenido acceso, se procede a la escalada de privilegios, cuyo objetivo es determinar hasta qué nivel de control puede llegar un atacante dentro del sistema. Posteriormente, se simula el mantenimiento del acceso, evaluando la posibilidad de persistencia mediante mecanismos como la instalación controlada de accesos remotos. Finalmente, el proceso concluye con la fase de informe y recomendaciones, en la cual se documentan las vulnerabilidades identificadas, su impacto y las medidas correctivas necesarias, clasificando los riesgos en diferentes niveles de criticidad.

Para la ejecución de estas pruebas, los profesionales emplean diversas herramientas especializadas. Entre las más destacadas se encuentran Nmap, utilizada para el análisis de puertos y servicios; THC Hydra, orientada a la evaluación de mecanismos de autenticación; WPSCAN, enfocada en la seguridad de sitios WordPress; Nessus, para el escaneo de vulnerabilidades; y herramientas avanzadas como Hashcat, Aircrack-ng y BeEF, entre otras. Estas soluciones permiten realizar análisis profundos y automatizados sobre distintos componentes de la infraestructura tecnológica.

El pentesting ofrece múltiples beneficios para las organizaciones. Permite identificar vulnerabilidades desconocidas antes de que sean explotadas, evaluar la capacidad de respuesta ante incidentes, cumplir con normativas de seguridad como ISO/IEC 27001 y fortalecer la cultura de seguridad en el personal. Asimismo, contribuye a reducir riesgos económicos y daños reputacionales derivados de posibles brechas de seguridad.

No obstante, también presenta ciertas limitaciones. Los resultados obtenidos representan una evaluación puntual en el tiempo, ya que nuevas vulnerabilidades pueden surgir posteriormente. Además, su implementación implica costos elevados en términos de recursos humanos y tecnológicos, y, si no se gestiona adecuadamente, puede generar interrupciones en los sistemas evaluados. Por ello, el pentesting debe considerarse como un complemento dentro de una estrategia de seguridad continua.

Un ejemplo práctico de su aplicación se observa en el caso de una empresa de comercio electrónico que, tras realizar una prueba de penetración, identificó un puerto de base de datos expuesto sin mecanismos adecuados de autenticación. Mediante la explotación controlada, se evidenció el riesgo de acceso a datos sensibles de clientes. Como resultado, se implementaron medidas correctivas como el cierre de puertos innecesarios, la incorporación de autenticación multifactor y el cifrado de la base de datos, evitando así una posible brecha de seguridad.

Finalmente, es importante destacar que el pentesting debe desarrollarse bajo un marco legal y ético estricto. Toda prueba debe contar con autorización formal y contractual, ya que su ejecución sin consentimiento constituye un delito. Asimismo, los profesionales deben actuar con confidencialidad, transparencia y responsabilidad, garantizando que los resultados obtenidos se utilicen exclusivamente para mejorar la seguridad de la organización.

3.2. Estrategias de defensa y ataque en ciberseguridad: Red Team, Blue Team y SOC

La seguridad informática puede concebirse como un proceso dinámico caracterizado por la interacción constante entre estrategias de ataque y mecanismos de defensa. En este contexto, surgen dos enfoques fundamentales dentro de la ciberseguridad moderna, el Red Team, encargado de simular ataques ofensivos, y el Blue Team, responsable de la defensa y protección de los activos digitales. Ambos desempeñan un papel esencial en la construcción de entornos tecnológicos seguros y resilientes, ya que permiten evaluar de manera integral las capacidades de una organización frente a amenazas reales. Estudios recientes han demostrado que la interacción entre estos equipos no solo mejora la detección de vulnerabilidades, sino que también fortalece la capacidad de respuesta ante incidentes, promoviendo un aprendizaje continuo basado en escenarios prácticos (Chindrus y Caruntu, 2023).

En este sentido, la colaboración entre equipos ofensivos y defensivos ha evolucionado hacia modelos más estructurados de trabajo en equipo, donde la especialización de roles y la coordinación estratégica resultan factores determinantes para el éxito de las operaciones de ciberseguridad. Investigaciones sobre dinámicas de equipos en entornos competitivos indican que la asignación clara de funciones y la comunicación efectiva incrementan significativamente el rendimiento colectivo, permitiendo una defensa más eficiente frente a ataques complejos (Buchler et al., 2018). Esta sinergia ha dado lugar al concepto de integración entre Red Team y Blue Team, el cual se considera un elemento clave para mejorar la resiliencia organizacional, ya que facilita una comprensión más profunda de las amenazas y de las capacidades defensivas disponibles (Gaifulina, 2025).

Asimismo, enfoques más avanzados incorporan marcos de referencia como MITRE ATT and CK, los cuales permiten estructurar y estandarizar las simulaciones de ataques, proporcionando un lenguaje común para analizar tácticas, técnicas y procedimientos utilizados

por los adversarios. La integración de estos marcos con ejercicios de Red Teaming ha sido identificada como un cambio significativo en la evaluación de la ciberseguridad, ya que posibilita un análisis más sistemático y realista de las vulnerabilidades, así como la mejora continua de los mecanismos de defensa (Yulianto et al., 2025). Por otra parte, desde una perspectiva más amplia, se ha planteado la necesidad de alinear la práctica operativa con fundamentos teóricos sólidos, de modo que la ciberdefensa sea entendida como un sistema integral en el que convergen aspectos técnicos, organizacionales y estratégicos (De Nobrega et al., 2024).

Estos enfoques evidencian que la seguridad informática no puede depender exclusivamente de medidas defensivas aisladas, sino que requiere una interacción constante entre ataque simulado y defensa activa, donde la cooperación, la especialización y la integración metodológica permiten fortalecer de manera significativa la postura de seguridad de las organizaciones.

La dinámica Red vs Blue se ha consolidado como una metodología de referencia, ya que permite a las organizaciones evaluar sus capacidades de respuesta frente a escenarios de amenaza realistas (Figura 3.2).

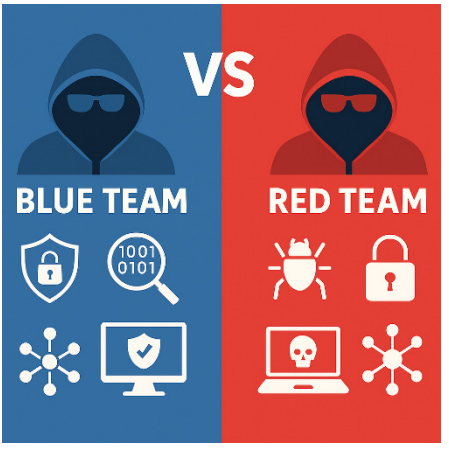


Figura 3.2. Blue team y Red team.

El Red Team asume el rol de un adversario controlado, cuyo objetivo es identificar debilidades en los sistemas mediante la simulación de ciberataques. Para ello, emplea técnicas ofensivas avanzadas, tales como campañas de

phishing, escaneo y explotación de vulnerabilidades, inyección de código malicioso y ataques de denegación de servicio. Su finalidad no es generar daño, sino evidenciar el nivel de exposición de una organización y proporcionar recomendaciones para mejorar su postura de seguridad.

Por su parte, el Blue Team se encarga de la defensa activa de la infraestructura tecnológica. Sus funciones incluyen la monitorización continua de sistemas, la configuración segura de dispositivos, la detección de intrusiones, el análisis de tráfico y la gestión de incidentes. Asimismo, utiliza herramientas especializadas como sistemas SIEM, IDS/IPS, soluciones antimalware y técnicas de análisis forense digital. Su labor no se limita a la respuesta ante ataques, sino que también abarca la prevención y la capacitación del personal en buenas prácticas de seguridad.

La interacción entre ambos equipos da lugar a la dinámica Red vs Blue, la cual promueve un proceso continuo de mejora. Mientras el Red Team evalúa la efectividad de las defensas mediante ataques simulados, el Blue Team fortalece sus capacidades de detección y respuesta al enfrentarse a situaciones que replican condiciones reales. Este proceso no busca establecer un ganador, sino generar retroalimentación que permita optimizar las políticas y mecanismos de seguridad.

En este escenario, surge el concepto de Purple Team, que actúa como un elemento integrador entre las funciones ofensivas y defensivas. Su principal objetivo es facilitar la comunicación entre ambos equipos, documentar los resultados obtenidos y asegurar que las lecciones aprendidas se traduzcan en mejoras concretas dentro de la estrategia de ciberseguridad.

La aplicación práctica de esta metodología se evidencia en diversos contextos organizacionales. Por ejemplo, en el sector financiero, pueden realizarse campañas de phishing simuladas para evaluar la respuesta de los empleados; en empresas de telecomunicaciones, se ejecutan pruebas controladas de denegación de servicio para medir la resiliencia de la red; y en centros de datos,

se desarrollan ejercicios de escalada de privilegios para identificar configuraciones inseguras.

En el ámbito educativo, la dinámica Red vs Blue posee un alto valor pedagógico, ya que permite a los estudiantes comprender tanto las técnicas de ataque como las estrategias de defensa en entornos controlados, promoviendo el desarrollo de habilidades como el pensamiento crítico, la resolución de problemas y el trabajo en equipo.

No obstante, tanto el Red Team como el Blue Team enfrentan desafíos significativos en la actualidad. El primero debe adaptarse a tecnologías emergentes como la computación en la nube, el Internet de las Cosas (IoT) y los sistemas industriales, mientras que el segundo debe gestionar grandes volúmenes de información, reducir los tiempos de detección de incidentes e incorporar tecnologías avanzadas como la inteligencia artificial.

En respuesta a la creciente complejidad de las amenazas, las organizaciones han implementado centros especializados conocidos como Security Operations Center (SOC). Estos centros constituyen entornos integrados donde convergen personas, procesos y tecnología con el objetivo de garantizar la detección temprana, el análisis y la respuesta ante incidentes de ciberseguridad. Un SOC opera de manera continua y centralizada, supervisando la infraestructura tecnológica, gestionando alertas, coordinando respuestas y generando informes estratégicos.

La estructura de un SOC se basa en tres componentes fundamentales: el recurso humano, conformado por analistas y especialistas; los procesos, que establecen metodologías estandarizadas; y la tecnología, que incluye herramientas de monitoreo, detección y respuesta. Asimismo, los SOC suelen organizarse en distintos niveles operativos, que abarcan desde la monitorización inicial hasta la gestión estratégica y la respuesta avanzada a incidentes.

En cuanto a las herramientas utilizadas, destacan soluciones como SIEM para la correlación de eventos, plataformas SOAR para la automatización de respuestas,

sistemas IDS/IPS para la detección y prevención de intrusiones, y tecnologías EDR/XDR para la protección de endpoints. Además, se emplean mecanismos como honeypots para analizar el comportamiento de los atacantes.

Finalmente, la implementación de un SOC puede adoptar diferentes modelos organizativos, tales como interno, tercerizado, híbrido o virtual, cada uno con ventajas y limitaciones en función de los recursos y necesidades de la organización. En conjunto, estos elementos consolidan un enfoque integral de ciberseguridad, basado en la coordinación entre capacidades ofensivas y defensivas, orientado a la protección efectiva de los activos digitales.

3.3. Ingeniería social: técnicas de manipulación, riesgos y estrategias de prevención

La ingeniería social constituye uno de los métodos de ataque más efectivos en el ámbito de la ciberseguridad, debido a que se enfoca en explotar la vulnerabilidad más impredecible de cualquier sistema: el factor humano. A diferencia de los ataques estrictamente técnicos, que dependen de fallos en software o hardware, la ingeniería social se basa en la manipulación psicológica de las personas con el objetivo de inducirlas a revelar información confidencial, facilitar accesos no autorizados o ejecutar acciones que comprometan la seguridad de los sistemas.

En este sentido, la ingeniería social no se dirige directamente a las infraestructuras tecnológicas, sino a los procesos cognitivos y conductuales de los individuos. Esta característica la convierte en una de las amenazas más complejas de mitigar, ya que la eficacia de las medidas de seguridad tecnológica resulta insuficiente si los usuarios no poseen el nivel adecuado de concienciación y formación en materia de riesgos digitales.

Desde una perspectiva conceptual, la ingeniería social puede definirse como el conjunto de técnicas orientadas a manipular el comportamiento humano mediante el uso de persuasión, engaño y explotación de emociones.

Los atacantes aprovechan factores psicológicos como la curiosidad, el miedo, la confianza, la urgencia o la autoridad, con el fin de influir en la toma de decisiones de las víctimas. A diferencia de otros tipos de ataques, no requiere el uso de exploits complejos ni técnicas de fuerza bruta, sino que se fundamenta en la interacción social y la construcción de escenarios creíbles. Además, su alcance es transversal, afectando tanto a grandes organizaciones como a usuarios individuales, especialmente en entornos digitales y redes sociales. En este sentido, diversos estudios han demostrado que la vulnerabilidad humana constituye uno de los eslabones más débiles en la ciberseguridad, ya que las decisiones de los usuarios suelen estar influenciadas por sesgos cognitivos y heurísticas que pueden ser explotadas por los atacantes (Greavu-Şerban et al., 2025; Montañez et al., 2020).

La literatura reciente enfatiza que la efectividad de la ingeniería social radica en la comprensión profunda del comportamiento humano y de los contextos sociales en los que se desarrollan las interacciones digitales. Factores como la falta de conciencia en seguridad, la sobrecarga de información y la confianza excesiva en entornos virtuales incrementan significativamente la susceptibilidad de los usuarios a este tipo de ataques (Alshammari et al., 2025). Asimismo, se ha señalado que las redes sociales representan un entorno particularmente propicio para la ejecución de estas técnicas, debido a la gran cantidad de información personal disponible y a la facilidad para construir identidades falsas o escenarios engañosos convincentes (Ukwadinachi, 2025). En consecuencia, los atacantes han perfeccionado estrategias como el phishing, el vishing y otras variantes más sofisticadas, adaptándolas a los avances tecnológicos y a los cambios en el comportamiento de los usuarios.

Por otra parte, investigaciones centradas en el impacto de la ingeniería social destacan que este tipo de ataques puede comprometer gravemente la seguridad de la información y la integridad de los sistemas, incluso en organizaciones con altos niveles de protección tecnológica (Alghamdi, 2022). Esto evidencia que la

seguridad no depende exclusivamente de herramientas técnicas, sino también de la preparación y concienciación de los usuarios. En este contexto, se han desarrollado enfoques más integrales para mitigar estos riesgos, como el uso de plataformas de automatización y respuesta ante incidentes, que permiten detectar y responder de manera más eficiente a amenazas basadas en ingeniería social (Waelchli y Walter, 2025). Asimismo, se subraya la importancia de incorporar programas de formación continua y estrategias centradas en el factor humano, con el fin de reducir la probabilidad de éxito de estos ataques (Nonum et al., 2025).

Estas aportaciones evidencian que la ingeniería social no solo representa una amenaza técnica, sino también un fenómeno complejo que involucra dimensiones psicológicas, sociales y tecnológicas. Por ello, su estudio y mitigación requieren un enfoque multidisciplinario que combine herramientas de ciberseguridad con el análisis del comportamiento humano, permitiendo así fortalecer de manera integral la resiliencia frente a este tipo de amenazas.

Entre las principales técnicas de ingeniería social se encuentran el phishing, que consiste en el envío masivo de comunicaciones fraudulentas que simulan ser legítimas; el spear phishing, una variante más sofisticada y personalizada dirigida a víctimas específicas; y el pretexting, donde el atacante construye un escenario ficticio para obtener información. Asimismo, destacan el baiting, que utiliza elementos atractivos como archivos o dispositivos infectados para inducir a la víctima a interactuar; el tailgating, que permite el acceso físico no autorizado mediante la explotación de la confianza; y el vishing y smishing, que trasladan estas prácticas a llamadas telefónicas y mensajes de texto, respectivamente.

La efectividad de estas técnicas se sustenta en principios psicológicos bien conocidos. Entre ellos, la autoridad, que impulsa a las personas a obedecer figuras de poder; la urgencia, que reduce la capacidad de análisis crítico; la reciprocidad, que genera compromiso al recibir un beneficio previo; el miedo, que induce a decisiones

impulsivas; y la confianza, que facilita la suplantación de identidades conocidas o legítimas.

Diversos casos reales evidencian el impacto de la ingeniería social en entornos digitales. Por ejemplo, ataques de spear phishing han permitido comprometer cuentas de alto perfil en plataformas digitales, mientras que estrategias como la distribución de dispositivos infectados han sido utilizadas para obtener acceso inicial a redes corporativas. Estos incidentes demuestran que, en muchos casos, el punto de entrada de un ciberataque no es una vulnerabilidad técnica, sino una acción humana aparentemente inofensiva.

El impacto de la ingeniería social en las organizaciones puede ser significativo, abarcando pérdidas económicas derivadas de fraudes, daño reputacional, filtración de información sensible y acceso no autorizado a sistemas críticos. De hecho, numerosos ataques complejos tienen su origen en una primera interacción basada en engaño o manipulación.

Para mitigar estos riesgos, es necesario adoptar un enfoque integral que combine medidas tecnológicas y estrategias orientadas al factor humano. Entre las medidas técnicas se incluyen la implementación de filtros antiphishing, el uso de autenticación multifactor y el monitoreo continuo de accesos. No obstante, las medidas humanas resultan igualmente fundamentales, destacando los programas de capacitación en ciberseguridad, la implementación de políticas de verificación de identidad y la realización de simulacros periódicos que permitan evaluar la preparación de los usuarios ante posibles ataques.

En la actualidad, las redes sociales representan una fuente significativa de información para los atacantes, quienes utilizan datos personales aparentemente inofensivos para diseñar ataques altamente personalizados. Fotografías, publicaciones y datos públicos pueden ser empleados para construir mensajes creíbles y aumentar la probabilidad de éxito de técnicas como el spear phishing.

Desde el punto de vista legal, la ingeniería social aplicada con fines maliciosos constituye un delito en la mayoría de los países, incluyendo prácticas como el fraude, el robo de identidad y el acceso indebido a sistemas. No obstante, también existe su aplicación en contextos éticos y controlados, como en auditorías de seguridad y pruebas de penetración, donde se simulan ataques con el objetivo de evaluar la resiliencia organizacional y fortalecer las medidas de protección.

3.4. Herramientas de ciberseguridad para la prevención y detección de amenazas

En la actualidad, ningún entorno tecnológico puede considerarse plenamente seguro si no dispone de mecanismos robustos orientados tanto a la prevención de ataques como a la detección temprana de anomalías. La creciente sofisticación de las amenazas, junto con la velocidad a la que evolucionan los ciberataques, obliga a las organizaciones a adoptar un enfoque proactivo en materia de ciberseguridad. En este contexto, resulta indispensable contar con un conjunto integrado de herramientas que permitan anticipar riesgos, bloquear actividades maliciosas y responder de manera eficaz ante incidentes de seguridad. La literatura reciente destaca que la inteligencia de amenazas cibernéticas se ha convertido en un componente fundamental para lograr este objetivo, ya que permite recopilar, analizar y correlacionar información sobre posibles ataques, facilitando la toma de decisiones estratégicas y la anticipación de riesgos (Santos et al., 2025). Asimismo, se ha evidenciado que el uso de herramientas de nueva generación, basadas en análisis predictivo y automatización, mejora significativamente la capacidad de detección y respuesta frente a incidentes complejos (Dalal, 2021).

En este sentido, el desarrollo de marcos de ciberseguridad orientados al futuro propone la integración de capacidades defensivas y ofensivas, así como la implementación de sistemas adaptativos capaces de responder dinámicamente a las amenazas emergentes (Safitra et al., 2023). Estas aproximaciones consideran no solo la protección de infraestructuras

tecnológicas, sino también la seguridad de la identidad digital, la cual se ha convertido en uno de los principales objetivos de los atacantes en entornos digitales (Alqaydi et al., 2024). A su vez, la incorporación de inteligencia artificial y aprendizaje automático ha permitido mejorar la detección de patrones anómalos y comportamientos sospechosos en grandes volúmenes de datos, lo que resulta especialmente útil en contextos como redes sociales y plataformas digitales, donde las amenazas evolucionan rápidamente (Alsodi et al., 2025). De igual manera, estas tecnologías han demostrado ser eficaces en la automatización de procesos de respuesta ante incidentes, reduciendo los tiempos de reacción y minimizando el impacto de los ataques (Rahman et al., 2025).

Las herramientas de prevención y detección no operan de forma aislada, sino que forman parte de una estrategia integral de ciberseguridad. Esta estrategia combina soluciones tecnológicas avanzadas, configuraciones adecuadas de los sistemas, monitoreo continuo de la infraestructura y, de manera fundamental, la capacitación del factor humano. La sinergia entre estos elementos permite construir un entorno resiliente capaz de enfrentar amenazas tanto conocidas como emergentes. En particular, se ha subrayado la importancia de abordar riesgos internos, como las amenazas provenientes de usuarios autorizados, mediante la implementación de controles, políticas de seguridad y mecanismos de monitoreo que reduzcan la probabilidad de incidentes (Alsowail y Al-Shehari, 2022). Asimismo, la integración de inteligencia de amenazas, automatización y análisis avanzado de datos contribuye a fortalecer la postura de seguridad organizacional, permitiendo una respuesta más coordinada y eficiente ante ataques sofisticados (Santos et al., 2025).

Estos estudios manifiestan que la ciberseguridad moderna requiere un enfoque holístico que combine tecnologías avanzadas, estrategias organizacionales y concienciación del usuario. Solo mediante la integración de estos elementos es posible desarrollar sistemas capaces de adaptarse a un panorama de amenazas en

constante evolución y garantizar la protección efectiva de los activos digitales.

Desde una perspectiva conceptual, la prevención en ciberseguridad se orienta a evitar que los ataques se materialicen o, en su defecto, a reducir significativamente la probabilidad de éxito de los mismos. Este enfoque se basa en el fortalecimiento de la infraestructura tecnológica mediante la reducción de la superficie de ataque. Entre las principales medidas preventivas se encuentran la configuración segura de servidores, routers y dispositivos de red, la implementación de controles de acceso basados en roles, la actualización constante de software y la aplicación oportuna de parches de seguridad. Asimismo, el uso de tecnologías como firewalls, mecanismos de cifrado y autenticación multifactor contribuye a establecer barreras efectivas frente a accesos no autorizados.

Por otro lado, la detección en ciberseguridad se centra en la identificación temprana de comportamientos anómalos o actividades sospechosas que puedan indicar la presencia de una amenaza o un compromiso del sistema. La detección eficaz permite activar mecanismos de respuesta antes de que el impacto del ataque sea significativo o irreversible. Para ello, se emplean técnicas como la monitorización continua del tráfico de red, el análisis de registros (logs) y la implementación de sistemas especializados como los IDS (Intrusion Detection Systems) y plataformas SIEM, capaces de correlacionar eventos provenientes de múltiples fuentes para identificar patrones de ataque.

En cuanto a las herramientas de prevención, destacan los firewalls, considerados la primera línea de defensa entre redes internas y externas, los cuales pueden ser de filtrado de paquetes, de inspección de estado o de próxima generación (NGFW). Asimismo, los sistemas antivirus y antimalware continúan siendo fundamentales, aunque actualmente se complementan con tecnologías basadas en inteligencia artificial que permiten detectar amenazas avanzadas y ataques de día cero (zero-day). Los sistemas de prevención de intrusos (IPS) representan una evolución de los IDS, ya que no solo detectan

actividades maliciosas, sino que también las bloquean en tiempo real. Por su parte, las soluciones de control de acceso a la red (NAC) garantizan que únicamente dispositivos que cumplen con las políticas de seguridad puedan conectarse a la infraestructura, reduciendo significativamente el riesgo de accesos indebidos.

En el ámbito de la detección, las herramientas más relevantes incluyen los IDS, que analizan el tráfico de red en busca de patrones sospechosos, y los sistemas SIEM, que centralizan y correlacionan grandes volúmenes de datos de seguridad para identificar incidentes complejos. Asimismo, las soluciones EDR (Endpoint Detection and Response) permiten monitorear y responder a amenazas en dispositivos finales, facilitando el aislamiento de equipos comprometidos. Adicionalmente, los honeypots actúan como sistemas señuelo diseñados para atraer a los atacantes, permitiendo estudiar sus técnicas y generar alertas tempranas sobre posibles intrusiones.

En los últimos años, han surgido diversas tendencias tecnológicas que están transformando el panorama de la prevención y detección. Entre ellas destaca la aplicación de técnicas de Machine Learning e inteligencia artificial, que permiten identificar patrones de comportamiento anómalos que pasarían desapercibidos para los sistemas tradicionales. Asimismo, la automatización mediante plataformas SOAR (Security Orchestration, Automation and Response) facilita la integración de procesos de detección, análisis y respuesta, reduciendo significativamente los tiempos de reacción. El modelo de seguridad Zero Trust, basado en el principio de no confiar en ningún acceso por defecto, y el desarrollo de soluciones específicas para entornos de computación en la nube, también representan avances clave en este ámbito.

La implementación de herramientas de prevención y detección aporta múltiples beneficios a las organizaciones. Entre ellos se encuentran la reducción del tiempo de exposición ante amenazas, la identificación de vulnerabilidades no detectadas previamente, el cumplimiento de normativas internacionales como ISO/IEC 27001, NIST o GDPR, y el fortalecimiento

de la confianza de clientes y socios. Además, estas herramientas contribuyen a optimizar el trabajo de los equipos de seguridad al proporcionar visibilidad y control sobre la infraestructura tecnológica.

No obstante, también existen limitaciones y retos asociados a su implementación. Entre los principales se encuentran la generación de falsos positivos, que puede provocar sobrecarga de alertas, los altos costos de adquisición y mantenimiento de las soluciones, la necesidad de contar con personal altamente capacitado y la complejidad de integrar múltiples herramientas dentro de un mismo entorno tecnológico.

Un ejemplo práctico de la aplicación de estas herramientas se observa en el ámbito académico y empresarial. Por ejemplo, una institución de educación superior que implementa firewalls de próxima generación junto con una plataforma SIEM puede detectar actividades sospechosas en su red. Ante intentos de acceso no autorizado mediante ataques de fuerza bruta, el sistema de detección genera alertas que permiten identificar comportamientos anómalos, mientras que las medidas preventivas bloquean el acceso indebido. De esta manera, la combinación de mecanismos de prevención y detección permite contener incidentes sin afectar la continuidad de las operaciones.

La integración de herramientas de prevención y detección constituye un pilar fundamental en la ciberseguridad moderna. Su correcta implementación, junto con una adecuada gestión y capacitación del personal, permite a las organizaciones enfrentar de manera efectiva un entorno de amenazas cada vez más complejo y dinámico.

3.5. El cibercrimen: definición, tipología y desafíos actuales

El cibercrimen y los ciberdelitos constituyen uno de los principales desafíos de la sociedad digital contemporánea. A medida que las tecnologías de la información y la comunicación se integran de manera creciente en ámbitos como las finanzas, la salud, la educación, la comunicación y la gestión gubernamental, también se amplían las oportunidades para la comisión

de actividades ilícitas en el ciberespacio. Estas amenazas abarcan desde fraudes relativamente simples hasta complejas operaciones criminales de carácter transnacional, ejecutadas por organizaciones altamente estructuradas. En este sentido, diversos estudios han señalado que el cibercrimen ha evolucionado hacia un fenómeno global interconectado, impulsado por factores como la digitalización acelerada, la disponibilidad de herramientas tecnológicas accesibles y la creciente interdependencia entre sistemas digitales (Chen et al., 2023). Asimismo, se reconoce que las dinámicas del cibercrimen están estrechamente vinculadas con el crimen organizado, adoptando estructuras similares en términos de coordinación, especialización y alcance internacional (Viano, 2017).

En este contexto, el cibercrimen no debe entenderse como un conjunto aislado de ataques informáticos, sino como un ecosistema globalizado en el que interactúan diversos actores, incluyendo individuos, grupos organizados e incluso entidades vinculadas a Estados. Esta complejidad implica que su análisis requiera una visión integral que contemple aspectos técnicos, económicos, sociales y legales. Investigaciones recientes han propuesto nuevas formas de conceptualizar el cibercrimen, destacando la necesidad de desarrollar tipologías y taxonomías que permitan clasificar de manera más precisa las distintas modalidades delictivas en el entorno digital (Phillips et al., 2022). Además, se ha planteado que el cibercrimen puede converger con otras formas de criminalidad, dando lugar a fenómenos híbridos que combinan elementos del mundo físico y digital, lo que incrementa su complejidad y dificulta su prevención y control (Zhou et al., 2024).

Desde una perspectiva criminológica, el estudio del cibercrimen ha evolucionado significativamente, incorporando teorías tradicionales del delito junto con enfoques adaptados al entorno digital. En este sentido, revisiones sistemáticas destacan que el análisis del comportamiento delictivo en el ciberespacio requiere considerar variables como la anonimidad, la accesibilidad tecnológica y las oportunidades del entorno virtual, las

cuales modifican las dinámicas tradicionales del crimen (Onwuadiamu, 2025). Asimismo, se ha subrayado la importancia de integrar la teoría con la práctica en el ámbito de la aplicación de la ley, ya que las fuerzas de seguridad enfrentan desafíos específicos al investigar delitos digitales, como la jurisdicción internacional, la recolección de evidencia digital y la rápida evolución de las técnicas utilizadas por los delincuentes (Curtis y Oxburgh, 2022). En esta línea, la investigación contemporánea enfatiza la necesidad de desarrollar nuevos enfoques metodológicos y marcos analíticos que permitan comprender mejor este fenómeno en constante transformación (Bossler y Berenblum, 2019).

Desde una perspectiva conceptual, el cibercrimen puede definirse como el conjunto de actividades ilícitas llevadas a cabo en entornos digitales, utilizando computadoras, redes o dispositivos electrónicos como medio principal para la ejecución del delito. Por su parte, el término ciberdelito hace referencia a la manifestación específica de dichas actividades, es decir, a las acciones concretas que materializan el crimen, como puede ser un ataque de ransomware o una estafa mediante phishing. En este sentido, la literatura especializada resalta la importancia de diferenciar entre categorías y niveles de análisis, lo que permite comprender mejor la diversidad de amenazas existentes y diseñar estrategias más efectivas para su mitigación (Phillips et al., 2022).

La tipología del cibercrimen es amplia y puede clasificarse en función de sus objetivos y métodos. En primer lugar, se encuentran los delitos contra la confidencialidad, integridad y disponibilidad de los sistemas, que incluyen el acceso no autorizado a redes, la interceptación de comunicaciones y la alteración o destrucción de datos. En segundo lugar, se identifican los delitos informáticos tradicionales, tales como fraudes electrónicos, suplantación de identidad y estafas en línea. Asimismo, existen los delitos de contenido, que abarcan la distribución de material ilegal o la difusión de información manipulada con fines políticos o sociales. Finalmente, destacan los delitos asociados al

ciberespacio, como el ciberterrorismo, el ciberespionaje y los ataques dirigidos a infraestructuras críticas.

Diversos factores han impulsado el crecimiento del cibercrimen en las últimas décadas. Entre ellos, el anonimato que ofrece el entorno digital facilita la ocultación de identidades, mientras que la globalización permite que los ataques se ejecuten desde cualquier parte del mundo con impacto transnacional. A esto se suma el bajo costo de acceso a herramientas de hacking, muchas de las cuales se encuentran disponibles en mercados clandestinos, así como la alta rentabilidad de actividades como el ransomware y el robo de datos. Además, el uso de tecnologías como redes privadas virtuales (VPN), criptomonedas y la denominada Dark Web dificulta considerablemente la trazabilidad y persecución legal de los delincuentes.

Entre los ciberdelitos más comunes se encuentran el phishing, orientado a la obtención fraudulenta de credenciales; el ransomware, que implica el secuestro de información mediante cifrado; el carding, relacionado con el uso indebido de datos de tarjetas de crédito; el Business Email Compromise (BEC), basado en la suplantación de identidades corporativas; y el uso de malware bancario para comprometer sistemas financieros. Estas prácticas evidencian la diversidad de estrategias utilizadas por los atacantes para obtener beneficios económicos o causar daño.

El cibercrimen contemporáneo se caracteriza, además, por un alto grado de organización. Lejos de tratarse de actividades individuales, muchas de estas operaciones son ejecutadas por redes criminales que funcionan de manera similar a empresas, con estructuras jerárquicas, roles definidos y división de tareas. En este contexto, la Dark Web desempeña un papel relevante como mercado clandestino, donde se comercializan bases de datos robadas, herramientas de malware, servicios de hacking por encargo y documentos falsificados, lo que contribuye a la profesionalización de estas actividades ilícitas.

Desde el punto de vista legal, la lucha contra el cibercrimen requiere marcos normativos sólidos y una estrecha cooperación internacional. Instrumentos como el Convenio de Budapest representan esfuerzos importantes para armonizar legislaciones y facilitar la colaboración entre países. Sin embargo, uno de los principales retos radica en que el desarrollo tecnológico avanza a un ritmo más rápido que la capacidad de adaptación del derecho, generando vacíos legales que pueden ser aprovechados por los delincuentes.

La prevención y mitigación del cibercrimen no dependen únicamente de la legislación, sino también de la implementación de estrategias tecnológicas y educativas. En el ámbito tecnológico, destacan medidas como el uso de firewalls avanzados, sistemas de prevención de intrusos y mecanismos de cifrado de datos. Desde la perspectiva humana, resulta fundamental promover programas de concienciación en ciberseguridad, capacitar a los usuarios en la detección de fraudes y establecer políticas estrictas de gestión de accesos. Asimismo, la cooperación internacional, mediante el intercambio de información y la coordinación entre agencias, constituye un elemento clave para enfrentar amenazas de carácter global.

Finalmente, diversos casos emblemáticos ponen de manifiesto la magnitud del cibercrimen en la actualidad. Incidentes como el ataque de ransomware WannaCry en 2017, que afectó a cientos de miles de sistemas a nivel mundial, la filtración masiva de datos de Yahoo o el ataque a la infraestructura de Colonial Pipeline evidencian que el cibercrimen no es una amenaza futura, sino una realidad presente que impacta de manera directa en la estabilidad económica, social y tecnológica de las organizaciones y los Estados.

El concepto de Kill Chain nació en el ámbito militar como un modelo para describir las fases de un ataque enemigo, desde la identificación del objetivo hasta la neutralización. Este enfoque fue adoptado posteriormente en la ciberseguridad por Lockheed Martin, que lo transformó en una metodología para analizar cómo se desarrollan

los ciberataques y, más importante aún, en qué punto pueden ser interrumpidos.

La Cyber Kill Chain se ha consolidado como uno de los modelos más utilizados en el ámbito de la ciberseguridad para comprender la secuencia lógica de un ataque digital. Su importancia radica no solo en describir las etapas que sigue un atacante, sino también en proporcionar a los defensores una estructura clara que permita identificar puntos críticos de intervención. De este modo, el modelo facilita el diseño de estrategias de mitigación orientadas a interrumpir el ataque en sus primeras fases, reduciendo significativamente su impacto.

Desde una perspectiva conceptual, la Kill Chain en ciberseguridad puede definirse como un modelo estructurado que describe las distintas fases de un ataque cibernético, desde la recopilación inicial de información hasta la consecución de los objetivos del atacante. Este enfoque permite descomponer el ataque en etapas identificables, lo que facilita su monitoreo, detección y contención. Su principal valor radica en el cambio de paradigma que propone: en lugar de adoptar una postura reactiva ante incidentes ya materializados, promueve una defensa proactiva orientada a detener el ataque en sus etapas iniciales.

El modelo clásico desarrollado por Lockheed Martin establece siete fases principales que describen el ciclo completo de un ciberataque. La primera fase es el reconocimiento (reconnaissance), en la cual el atacante recopila información sobre la víctima, incluyendo direcciones IP, dominios, tecnologías utilizadas y datos de empleados. Esta información puede obtenerse mediante fuentes abiertas, como redes sociales o sitios web corporativos, constituyendo la base para las etapas posteriores.

La segunda fase, conocida como armamento (weaponization), implica la preparación del vector de ataque. En esta etapa, el atacante desarrolla o adapta herramientas maliciosas, como malware o exploits, que serán utilizados para comprometer el sistema objetivo.

Por ejemplo, puede crearse un archivo aparentemente legítimo que contenga código malicioso oculto.

A continuación, se encuentra la fase de entrega (delivery), donde el atacante envía el vector de ataque al objetivo. Los medios más comunes incluyen correos electrónicos fraudulentos, sitios web comprometidos o dispositivos físicos infectados. Esta fase es crítica, ya que representa el punto de contacto inicial con la víctima.

La cuarta fase es la explotación (exploitation), en la cual se aprovecha una vulnerabilidad del sistema o un error humano para ejecutar el código malicioso. Esto puede ocurrir, por ejemplo, cuando un usuario abre un archivo infectado o interactúa con un enlace malicioso.

Posteriormente, en la fase de instalación (installation), el atacante establece un mecanismo de persistencia dentro del sistema comprometido, generalmente mediante la instalación de malware que le permite mantener el acceso a largo plazo.

La sexta fase corresponde al comando y control (command and control, C2), donde el sistema comprometido establece comunicación con servidores externos controlados por el atacante. A través de esta conexión, el atacante puede enviar instrucciones, recibir información y mantener el control del entorno afectado.

Finalmente, la última fase es la de acciones sobre el objetivo (actions on objectives), en la cual se ejecutan los objetivos finales del ataque. Estos pueden incluir el robo de información, el sabotaje de sistemas, la exfiltración de datos o el cifrado de archivos con fines de extorsión, como ocurre en los ataques de ransomware.

Para ilustrar este modelo, puede considerarse el caso de un ataque de ransomware dirigido a una organización del sector salud. En la fase de reconocimiento, el atacante identifica debilidades en los sistemas de correo electrónico de la institución. Posteriormente, en la fase de armamento, prepara un documento malicioso con macros. Dicho archivo es enviado durante la fase de entrega, simulando provenir de un proveedor confiable. Al ser abierto por un empleado, se activa la explotación, lo que permite la instalación del malware. A continuación,

el sistema comprometido establece comunicación con el servidor de comando y control, desde donde recibe instrucciones. Finalmente, en la fase de acciones sobre el objetivo, los archivos de la organización son cifrados y se exige un rescate para su recuperación.

De esta manera, la Cyber Kill Chain proporciona un marco analítico fundamental para comprender el comportamiento de los atacantes y fortalecer las capacidades defensivas. Su aplicación permite a las organizaciones adoptar un enfoque estratégico, orientado a la detección temprana y la interrupción de amenazas antes de que alcancen sus objetivos.

04.

Análisis forense digital y gestión de evidencia electrónica

4.1. Cómputo forense digital: fundamentos, evidencia digital y proceso metodológico

El cómputo forense es una rama especializada de la ciberseguridad que integra conocimientos de informática, derecho y criminalística con el objetivo de identificar, preservar, analizar y presentar evidencias digitales en el contexto de una investigación. A diferencia de otras áreas de la seguridad informática, orientadas a la prevención o a la respuesta inmediata ante incidentes, el cómputo forense se enfoca en la reconstrucción de los hechos una vez ocurrido un evento, lo que permite determinar cómo sucedió, quién lo ejecutó y cuáles fueron sus consecuencias.

El auge del cómputo forense está directamente relacionado con el incremento de delitos en entornos digitales, los cuales abarcan desde fraudes

electrónicos hasta actividades de ciberespionaje. En la actualidad, esta disciplina no es exclusiva de organismos policiales o judiciales, sino que también es utilizada por empresas privadas, instituciones financieras, aseguradoras y diversas organizaciones que requieren investigar incidentes de seguridad o resolver disputas legales vinculadas al uso de tecnologías.

El análisis forense digital se define como la disciplina encargada de adquirir, preservar, examinar y presentar evidencia digital proveniente de dispositivos informáticos y entornos tecnológicos. Su finalidad es esclarecer incidentes cibernéticos, reconstruir eventos, identificar responsables y proporcionar pruebas confiables que puedan ser utilizadas en procesos legales. A diferencia de la seguridad informática preventiva, su propósito no es evitar ataques, sino comprender lo sucedido, aportar claridad técnica y respaldar jurídicamente los hechos investigados.

En la actualidad, la vida cotidiana y las actividades empresariales dependen en gran medida de la tecnología, lo que implica que numerosos delitos dejan huellas digitales que pueden ser analizadas. En este contexto, el análisis forense digital se convierte en una herramienta fundamental para recuperar datos eliminados o dañados en dispositivos como computadoras, teléfonos móviles, discos duros y memorias externas. Asimismo, permite investigar delitos cibernéticos como fraudes, accesos no autorizados, ataques de malware y robo de identidad. De igual manera, contribuye al esclarecimiento de delitos tradicionales en los que la evidencia digital actúa como complemento, como homicidios, secuestros o robos. Además, desempeña un papel clave en la protección de información confidencial dentro de organizaciones y en la aportación de pruebas técnicas que respalden procesos judiciales, garantizando justicia para las víctimas. Su relevancia radica en su capacidad para determinar quién, cómo, cuándo y dónde ocurrió un incidente o filtración de datos.

Diversos estudios destacan que el análisis forense digital ha evolucionado hacia un enfoque metodológico estructurado que permite garantizar la validez y

confiabilidad de los resultados obtenidos en las investigaciones (Patel y Patel, 2015). En este sentido, la aplicación de metodologías avanzadas y marcos normativos, como los estándares ISO/IEC, contribuye a fortalecer la precisión del análisis y la admisibilidad de la evidencia en entornos legales (Morić et al., 2026). Asimismo, el uso de herramientas especializadas, incluyendo soluciones de código abierto, ha ampliado las capacidades de los investigadores para examinar grandes volúmenes de datos de manera eficiente, permitiendo identificar patrones, reconstruir eventos y detectar actividades ilícitas con mayor precisión (Abeysekera, 2026).

Por otra parte, la literatura resalta que el cómputo forense no solo tiene un impacto técnico, sino también social y jurídico, ya que se ha convertido en un elemento clave para la resolución de delitos en la era digital. La evidencia digital, considerada en muchos casos como "invisible", permite revelar información crítica que de otro modo sería imposible de detectar, facilitando la identificación de responsables y la reconstrucción detallada de los hechos (Klasén et al., 2024). Además, se ha señalado la necesidad de fortalecer la formación y las capacidades institucionales en este ámbito, con el fin de responder adecuadamente a los desafíos que plantea el incremento de delitos tecnológicos y la complejidad de las investigaciones digitales (Aleisa, 2026). En este contexto, el análisis forense digital se consolida como un pilar esencial dentro de la ciberseguridad moderna, integrando conocimientos técnicos, metodológicos y legales para abordar incidentes de manera integral (Onome, 2023).

La evidencia digital presenta características particulares que la diferencian de otros tipos de evidencia pericial. En primer lugar, destaca su fragilidad y volatilidad, ya que puede perderse o modificarse con facilidad si no se maneja de manera adecuada. En segundo lugar, posee la capacidad de reproducirse mediante copias forenses que no alteran el original, lo que permite preservar su

integridad y facilitar su análisis sin comprometer la fuente original. Asimismo, es altamente escalable, dado que puede aplicarse a diferentes dispositivos y a grandes volúmenes de información, lo que la convierte en un recurso versátil dentro de las investigaciones. Finalmente, su valor probatorio depende del cumplimiento riguroso de la cadena de custodia y de la documentación detallada de cada procedimiento realizado, lo cual garantiza su validez en contextos judiciales y refuerza la credibilidad de los hallazgos obtenidos.

El cómputo forense sigue una metodología estructurada compuesta por varias fases que garantizan la validez técnica y jurídica de los resultados obtenidos. En una primera etapa se realiza un estudio preliminar en el que se recopila información inicial sobre el incidente y los dispositivos involucrados, incluyendo datos del equipo afectado, sistema operativo, usuarios y horarios de uso, así como entrevistas con las personas responsables. Esta fase permite establecer una primera línea temporal de los hechos. Posteriormente, se lleva a cabo la adquisición y recolección de evidencia, que consiste en la creación de copias exactas de los medios de almacenamiento mediante técnicas especializadas, respetando el orden de volatilidad de los datos. Para ello, se emplean herramientas de hardware y software que garantizan la integridad de la información.

La preservación de la evidencia constituye una etapa crítica del proceso, en la que los datos se almacenan en condiciones controladas para evitar su alteración o contaminación. En esta fase se utilizan mecanismos como el cálculo de valores hash para verificar que las copias no han sido modificadas, así como la documentación detallada de la cadena de custodia. Posteriormente, se realiza el análisis de la información mediante herramientas especializadas que permiten examinar archivos visibles, sectores ocultos, datos eliminados y patrones de comportamiento. Este análisis puede ser lógico, físico o estadístico, y tiene como objetivo identificar actividades sospechosas, malware, metadatos y eventos relevantes.

A lo largo de todo el proceso, se lleva a cabo una documentación exhaustiva que registra cada paso realizado, incluyendo procedimientos, herramientas utilizadas y resultados obtenidos, lo que garantiza la transparencia y reproducibilidad del análisis. Finalmente, los resultados se presentan mediante informes periciales que pueden ser técnicos o ejecutivos, dependiendo del público al que se dirijan. Estos informes deben exponer de manera clara y estructurada los hallazgos, estableciendo una línea temporal que permita reconstruir los hechos y facilitar la comprensión de la investigación.

Para el desarrollo de estas actividades, existen diversas herramientas forenses adaptadas a distintos sistemas operativos, que permiten realizar tareas como la creación de imágenes forenses, el análisis de sistemas de archivos, la recuperación de datos eliminados y el estudio de la memoria volátil. Estas herramientas resultan fundamentales para llevar a cabo investigaciones precisas y obtener evidencia relevante en entornos digitales.

No obstante, el análisis forense digital enfrenta importantes desafíos en la actualidad. Entre ellos destaca el manejo de grandes volúmenes de datos, que requiere capacidades avanzadas de procesamiento y análisis. Asimismo, el uso creciente de técnicas de cifrado dificulta el acceso a la información y su posterior análisis. Otro reto relevante es la detección de malware sofisticado diseñado para evadir mecanismos de análisis. Finalmente, el cumplimiento de normativas relacionadas con la protección de datos y la privacidad, como GDPR o HIPAA, impone restricciones adicionales que obligan a equilibrar la investigación con el respeto a los derechos de los individuos. Estos desafíos evidencian la necesidad de continuar desarrollando metodologías y herramientas que fortalezcan esta disciplina en un entorno digital en constante evolución.

4.2. Cadena de custodia y adquisición de evidencias digitales

En el ámbito del análisis forense digital, no es suficiente con identificar o recuperar información de un dispositivo comprometido; resulta imprescindible garantizar que dicha información pueda ser utilizada como evidencia válida en procesos judiciales o disciplinarios. En este contexto, adquiere especial relevancia la cadena de custodia, un mecanismo que asegura la trazabilidad de la evidencia desde su recolección hasta su presentación final, ya sea en un tribunal o en un informe pericial. La literatura reciente destaca que, en la era digital, la cadena de custodia ha evolucionado significativamente debido a la complejidad y naturaleza intangible de los datos, lo que exige procedimientos más rigurosos y adaptados a entornos tecnológicos avanzados (D'Anna et al., 2023).

De manera complementaria, el proceso de adquisición de evidencias constituye una etapa crítica dentro de la investigación forense, ya que establece las condiciones bajo las cuales la información es recolectada, copiada y preservada. La correcta aplicación de estos procedimientos resulta determinante para evitar la alteración, contaminación o invalidez de la evidencia, marcando la diferencia entre una prueba admisible y una que pueda ser rechazada. En este sentido, organismos especializados han señalado la importancia de aplicar buenas prácticas estandarizadas durante la manipulación de evidencia digital, incluyendo la preservación de la integridad de los datos, el uso de herramientas confiables y la documentación detallada de cada acción realizada (Guttman et al., 2022). Estas prácticas no solo garantizan la validez técnica de la evidencia, sino también su aceptación en contextos legales.

Desde una perspectiva conceptual, la cadena de custodia puede definirse como un procedimiento documentado que garantiza que la evidencia digital ha

sido recolectada, manipulada, almacenada y transferida de manera controlada, preservando su autenticidad, integridad y validez legal. En términos prácticos, constituye el historial completo de la evidencia, en el cual se registran todas las personas que han tenido acceso a ella, así como las acciones realizadas durante el proceso investigativo. Sin embargo, en entornos digitales modernos, este proceso enfrenta nuevos desafíos derivados del volumen, la volatilidad y la distribución de los datos, lo que ha impulsado el desarrollo de soluciones innovadoras basadas en tecnologías emergentes como blockchain, las cuales permiten mejorar la transparencia, inmutabilidad y trazabilidad de la evidencia digital (Haji et al., 2026: Sakshi et al., 2023).

Asimismo, investigaciones recientes subrayan que la gestión de la cadena de custodia debe adaptarse a nuevas realidades tecnológicas, como la computación en la nube, el Internet de las cosas y los entornos distribuidos, donde la evidencia puede encontrarse fragmentada en múltiples ubicaciones y dispositivos (Nath et al., 2024). Este escenario incrementa la complejidad de las investigaciones forenses, requiriendo metodologías más robustas que permitan garantizar la autenticidad de los datos en todo momento. En consecuencia, la cadena de custodia no solo se mantiene como un elemento fundamental del análisis forense digital, sino que también se posiciona como un componente estratégico para asegurar la legitimidad de las pruebas en un contexto donde la evidencia digital adquiere un papel cada vez más determinante en la resolución de conflictos legales y en la persecución de delitos informáticos.

La importancia de la cadena de custodia radica en múltiples aspectos fundamentales. En primer lugar, garantiza la validez legal de la evidencia, ya que la ausencia de trazabilidad puede derivar en su rechazo en instancias judiciales. Asimismo, protege la integridad técnica de los datos, asegurando que no han sido modificados, ya sea de forma intencional o accidental. Además, promueve la transparencia del proceso, permitiendo la verificación por parte de terceros, y

asegura la reproducibilidad de los resultados, facilitando que otros analistas puedan replicar el procedimiento. Finalmente, establece un mecanismo de responsabilidad, ya que cada individuo que manipula la evidencia queda debidamente registrado.

Para cumplir con estos objetivos, la cadena de custodia debe incluir ciertos elementos esenciales, tales como la identificación detallada de la evidencia (tipo de dispositivo, modelo, número de serie), los datos de la persona responsable de su recolección, el lugar de hallazgo, las condiciones de almacenamiento y los mecanismos de protección física. Asimismo, es fundamental registrar valores hash (como MD5, SHA-1 o SHA-256), que permiten verificar la integridad de la evidencia, así como documentar cada transferencia entre custodios, incluyendo fecha, hora y motivo.

En cuanto a las buenas prácticas, se recomienda el uso de formularios estandarizados, el sellado físico de los dispositivos mediante etiquetas inviolables, la preservación del original mediante el uso de copias forenses para el análisis, la restricción de acceso a personal autorizado y la documentación exhaustiva de cada acción realizada.

Por su parte, la adquisición de evidencias consiste en el proceso mediante el cual se extrae información de un sistema o dispositivo digital, generando una copia exacta que será utilizada para el análisis, mientras que el original se conserva intacto. Este principio responde a una de las reglas fundamentales de la informática forense: nunca trabajar directamente sobre la evidencia original.

Existen diversos métodos de adquisición, cuya selección depende del tipo de dispositivo, las condiciones del entorno y los objetivos de la investigación. La adquisición física permite realizar una copia bit a bit de todo el dispositivo, incluyendo información eliminada y sectores ocultos, lo que posibilita un análisis exhaustivo. En contraste, la adquisición lógica se limita a los archivos visibles para el sistema operativo, siendo más rápida

pero menos completa. La adquisición en vivo se realiza sobre sistemas en funcionamiento, permitiendo capturar información volátil como procesos en memoria o conexiones activas. Finalmente, la adquisición remota se aplica en entornos distribuidos o en la nube, requiriendo protocolos seguros de transmisión de datos.

Para llevar a cabo estos procesos, se emplean herramientas especializadas como FTK Imager, EnCase Forensic, el comando dd en sistemas Linux, Autopsy/ Sleuth Kit y Volatility, entre otras. Estas herramientas permiten generar copias forenses, verificar la integridad de los datos y realizar análisis detallados de los sistemas comprometidos.

Asimismo, existen principios fundamentales que deben regir la adquisición de evidencias. Entre ellos destacan la prohibición de trabajar sobre el original, la generación y verificación de valores hash antes y después de la copia, la documentación detallada del proceso, el uso de dispositivos write-blocker para evitar modificaciones accidentales y el mantenimiento de un orden cronológico riguroso en todas las acciones realizadas.

Un ejemplo práctico de aplicación se observa en el análisis de un equipo implicado en un fraude bancario. En este caso, el procedimiento inicia con la documentación detallada del dispositivo, seguido de su aseguramiento físico. Posteriormente, en un entorno controlado, se realiza una imagen forense utilizando herramientas especializadas, registrando los valores hash correspondientes. La copia obtenida es analizada para identificar evidencias relevantes, mientras que todo el proceso es documentado en la cadena de custodia, garantizando la validez legal de los hallazgos (Tabla 4.1 y 4.2).

La adecuada gestión de la cadena de custodia y la correcta adquisición de evidencias constituyen pilares fundamentales del análisis forense digital, asegurando no solo la integridad técnica de la información, sino también su validez en el ámbito legal.

Tabla 4.1. Ejemplo de formato de cadena de custodia.

N.° de evidencia	Descripción del objeto	Número de serie / Identificación	Persona que recolecta	Fecha y hora de recolección	Lugar de recolección	Método de adquisición	Hash (MD5/ SHA-1/ SHA-256)
001	Disco duro externo 1TB, color negro, marca Seagate	SN: 9XY12345	Ing. Juan Pérez	12/04/2025 – 10:30 AM	Oficina principal – Servidor 2	Imagen forense con FTK Imager	MD5: f4a3c5... SHA-256: 9d8a4f...

Tabla 4.2. Registro de transferencias.

Fecha y hora	Persona que entrega	Firma	Persona que recibe	Firma	Motivo de la transferencia
12/04/2025 – 12:00 PM	Ing. Juan Pérez (Recolector)	_____	Lic. María Gómez (Analista Forense)	_____	Inicio del análisis forense
14/04/2025 – 09:15 AM	Lic. María Gómez (Analista Forense)	_____	Abg. Carlos Ruiz (Custodio Legal)	_____	Entrega para proceso judicial

En el contexto actual de la informática forense, la adquisición y custodia de evidencias digitales enfrentan múltiples desafíos derivados de la evolución tecnológica, la complejidad de los entornos digitales y la globalización de los sistemas de información. Estos retos no solo impactan los aspectos técnicos del proceso, sino también las dimensiones legales y operativas de las investigaciones.

Uno de los principales desafíos está relacionado con las evidencias en entornos de computación en la nube. A diferencia de los sistemas tradicionales, donde los dispositivos físicos pueden ser incautados y analizados directamente, en la nube los datos se encuentran distribuidos en infraestructuras remotas, muchas veces

ubicadas en diferentes países. Esto dificulta el acceso directo a la evidencia y exige la existencia de acuerdos legales internacionales, así como la colaboración con proveedores de servicios, lo que puede ralentizar significativamente el proceso investigativo.

Otro reto importante es la volatilidad de los datos, especialmente en lo que respecta a la información almacenada en memoria RAM, procesos activos o conexiones de red. Este tipo de evidencia es altamente efímera y puede perderse fácilmente si no se actúa con rapidez y precisión. Por ello, los investigadores deben aplicar técnicas de adquisición en vivo que permitan capturar estos datos antes de que desaparezcan, lo que incrementa la complejidad del procedimiento.

Asimismo, el volumen masivo de información representa un desafío significativo. Los dispositivos actuales pueden almacenar grandes cantidades de datos, que van desde terabytes hasta petabytes en entornos corporativos. Este crecimiento exponencial dificulta tanto la adquisición como el análisis de la evidencia, incrementando los tiempos de procesamiento y requiriendo herramientas avanzadas capaces de gestionar grandes volúmenes de información de manera eficiente.

El uso de mecanismos de cifrado avanzado constituye otro obstáculo relevante. La implementación de tecnologías de cifrado en discos duros, dispositivos móviles y sistemas de almacenamiento en la nube, si bien fortalece la seguridad de la información, también dificulta el acceso a los datos durante una investigación forense. En muchos casos, la imposibilidad de descifrar la información puede limitar el alcance del análisis o requerir el uso de técnicas especializadas que implican mayores recursos y tiempo.

Finalmente, el marco legal internacional representa un desafío complejo en la gestión de evidencias digitales. La diversidad de legislaciones entre países en materia de privacidad, protección de datos y procedimientos de recolección de evidencia genera conflictos jurisdiccionales que pueden afectar la validez de las pruebas. La falta de armonización normativa obliga

a los investigadores a actuar con especial cuidado, asegurando el cumplimiento de las leyes aplicables en cada jurisdicción involucrada.

Estos retos evidencian la necesidad de desarrollar metodologías cada vez más sofisticadas, así como de fortalecer la cooperación internacional y la capacitación especializada en informática forense, con el fin de garantizar la correcta adquisición, preservación y validez de las evidencias digitales en un entorno tecnológico en constante evolución.

4.3. Análisis forense de ataques cibernéticos y malware

El análisis de ataques y malware constituye una de las fases más complejas, especializadas y críticas dentro del ámbito del cómputo forense digital. Mientras que las etapas previas del proceso forense se centran principalmente en la identificación, recolección y preservación de evidencias, esta fase tiene como objetivo profundizar en la comprensión del incidente de seguridad, determinando cómo se llevó a cabo el ataque, qué vulnerabilidades fueron explotadas y cuál fue el impacto real sobre los sistemas y la información comprometida. En este sentido, diversos marcos metodológicos han sido desarrollados para estructurar la investigación de incidentes, como el modelo D4I, que propone un enfoque sistemático para analizar ataques cibernéticos mediante la correlación de evidencias, la reconstrucción de eventos y la identificación de patrones de comportamiento malicioso (Dimitriadis et al., 2020). Este tipo de enfoques permite no solo comprender el incidente de manera integral, sino también mejorar la capacidad de respuesta ante futuros ataques.

En el contexto actual, caracterizado por la constante evolución de las amenazas cibernéticas, el análisis de malware se ha convertido en una disciplina altamente técnica que requiere conocimientos avanzados en sistemas operativos, redes, programación y seguridad informática. La proliferación de códigos maliciosos cada vez más sofisticados, como ransomware, troyanos avanzados, rootkits y spyware, ha obligado

a los analistas forenses a desarrollar metodologías y herramientas especializadas que permitan no solo identificar la presencia del malware, sino también comprender su comportamiento, estructura y propósito. Estudios recientes destacan que los ataques modernos suelen desarrollarse en múltiples etapas, lo que incrementa su complejidad y dificulta su detección, requiriendo enfoques forenses capaces de analizar cada fase del ataque de manera integrada (Nisioti et al., 2023). Asimismo, investigaciones sobre ciberespionaje han evidenciado que ciertos tipos de malware están diseñados para permanecer ocultos durante largos periodos, recopilando información de forma sigilosa, lo que hace imprescindible el uso de técnicas avanzadas de análisis para su identificación (Kara, 2021).

El objetivo del análisis no se limita a la eliminación del software malicioso, sino que busca reconstruir la secuencia completa del ataque, identificar posibles vectores de entrada, determinar el alcance del daño y generar inteligencia de amenazas que permita prevenir incidentes futuros. En este contexto, el uso de tecnologías emergentes como el aprendizaje profundo ha demostrado ser altamente efectivo para la detección y clasificación de malware, especialmente en entornos complejos como el Internet de las cosas, donde la diversidad de dispositivos y la cantidad de datos dificultan el análisis tradicional (Qureshi et al., 2024). De igual manera, los enfoques basados en inteligencia artificial permiten realizar análisis predictivos y proactivos, facilitando la identificación temprana de amenazas y mejorando la capacidad de respuesta de las organizaciones (Abirami et al., 2024).

Por otra parte, revisiones sistemáticas recientes han resaltado la importancia de combinar técnicas tradicionales con métodos automatizados para mejorar la precisión en la detección y clasificación de malware, lo que contribuye a optimizar los procesos de análisis forense y a reducir los tiempos de respuesta ante incidentes (Berrios et al., 2025). En este sentido, el análisis de ataques y malware no solo constituye una fase investigativa, sino también un componente

estratégico dentro de la ciberseguridad, ya que permite transformar la información obtenida en conocimiento útil para fortalecer las defensas, mejorar las políticas de seguridad y anticiparse a nuevas amenazas. En consecuencia, esta disciplina desempeña un papel fundamental en la protección de los sistemas digitales y en la evolución de las estrategias de defensa en un entorno tecnológico cada vez más dinámico y complejo.

Desde una perspectiva conceptual, un ataque cibernético puede definirse como cualquier acción deliberada llevada a cabo por un actor malicioso con el propósito de comprometer la confidencialidad, integridad o disponibilidad de los sistemas de información. Estos ataques pueden tener diversas motivaciones, tales como beneficios económicos, espionaje, sabotaje o activismo político, y suelen apoyarse en el uso de herramientas automatizadas o software malicioso.

Por su parte, el término malware (acrónimo de malicious software) hace referencia a cualquier programa diseñado con fines maliciosos, como dañar sistemas, robar información o interrumpir el funcionamiento normal de los servicios. A diferencia del software legítimo, el malware actúa de forma encubierta y está orientado a beneficiar al atacante, ya sea mediante la obtención de ganancias económicas o el acceso a información estratégica.

El análisis forense requiere una adecuada clasificación del malware, ya que cada tipo presenta características específicas y deja diferentes rastros en el sistema. Entre las principales categorías se encuentran los virus, que infectan archivos legítimos; los gusanos, que se propagan automáticamente a través de redes; los troyanos, que se presentan como software confiable; el spyware, orientado al espionaje; el ransomware, que cifra información con fines de extorsión; los rootkits, diseñados para ocultar la presencia del atacante; el adware, que despliega publicidad no autorizada; y los keyloggers, que registran las pulsaciones del teclado para capturar credenciales.

Para llevar a cabo el análisis de malware, se emplean diversas metodologías que permiten estudiar tanto

la estructura del código como su comportamiento en ejecución. El análisis estático consiste en examinar el malware sin ejecutarlo, analizando su código, estructura interna, cabeceras y cadenas de texto (strings), lo que permite identificar funciones sospechosas, como rutinas de cifrado o conexiones a redes externas. Por otro lado, el análisis dinámico implica la ejecución del malware en entornos controlados o sandbox, donde se observan sus acciones en tiempo real, como la creación de archivos, modificaciones en el sistema o conexiones a servidores remotos. Finalmente, el análisis híbrido combina ambos enfoques, proporcionando una visión más completa del comportamiento del malware.

Durante este proceso, los analistas identifican los denominados Indicadores de Compromiso (IoC), que son evidencias técnicas que permiten detectar la presencia de malware en sistemas o redes. Estos indicadores pueden incluir direcciones IP sospechosas, dominios asociados a servidores de comando y control, hashes de archivos maliciosos, procesos inusuales, modificaciones en el registro del sistema y patrones de tráfico anómalos. La recopilación de estos indicadores es fundamental para alimentar sistemas de inteligencia de amenazas y fortalecer las capacidades de detección.

El análisis forense también busca reconstruir las fases de un ataque típico basado en malware. Este proceso generalmente comienza con una infección inicial, que puede ocurrir mediante técnicas como phishing, descargas maliciosas o dispositivos externos. Posteriormente, el malware se ejecuta en el sistema, establece mecanismos de persistencia para mantenerse activo y se comunica con servidores de comando y control (C2). Finalmente, se llevan a cabo las acciones maliciosas, como el robo de información o el cifrado de archivos, mientras el malware implementa técnicas de evasión para evitar su detección.

Un caso práctico ilustra la aplicación de estas técnicas: ante un incidente de ransomware en una organización, el equipo forense procede a obtener una imagen del sistema afectado, identificar los archivos cifrados y analizar el malware en un entorno controlado. A través

del cálculo de hashes y su comparación en bases de datos de amenazas, así como el monitoreo de su comportamiento, se identifican los indicadores de compromiso, los cuales son posteriormente utilizados para bloquear la propagación del ataque en la red corporativa.

No obstante, el análisis de malware enfrenta importantes desafíos en la actualidad. Entre ellos destaca el malware polimórfico, capaz de modificar su código para evadir mecanismos de detección basados en firmas; el malware sin archivos (fileless), que opera directamente en la memoria sin dejar rastros en el disco; el uso de técnicas avanzadas de cifrado y empaquetamiento que dificultan la ingeniería inversa; y los ataques persistentes avanzados (APT), que emplean malware altamente personalizado y diseñado para objetivos específicos.

4.4. Recuperación forense de datos y generación de imágenes digitales

En el ámbito del cómputo forense digital, uno de los desafíos más relevantes y recurrentes es la recuperación de datos eliminados y la correcta generación de imágenes forenses que permitan su análisis sin comprometer la integridad de la evidencia original. En la mayoría de las investigaciones digitales, resulta fundamental acceder a información que ha sido intencionalmente borrada por atacantes, usuarios internos o incluso administradores de sistemas con el propósito de ocultar actividades ilícitas. No obstante, desde una perspectiva técnica, el acto de eliminar un archivo en un sistema informático no implica su desaparición inmediata del medio de almacenamiento, lo que abre la posibilidad de reconstruir evidencias que pueden resultar determinantes en procesos judiciales o disciplinarios. Diversos estudios han demostrado que una cantidad significativa de dispositivos de almacenamiento contiene datos residuales recuperables incluso después de haber sido formateados o vendidos, lo que evidencia la persistencia de información digital y la importancia de aplicar técnicas forenses adecuadas para su recuperación (Shah et al., 2022).

En este contexto, el desarrollo de herramientas y metodologías especializadas ha permitido mejorar significativamente la eficacia de los procesos de recuperación de datos. Investigaciones recientes destacan la incorporación de técnicas avanzadas en laboratorios de análisis forense digital, las cuales facilitan la reconstrucción de archivos eliminados, la recuperación de estructuras de sistemas de archivos y la identificación de fragmentos de información dispersa en medios de almacenamiento (Cruz, 2024). Asimismo, en escenarios reales de incidentes de seguridad, el análisis forense ha demostrado ser clave para restaurar información crítica y comprender las causas de las brechas de seguridad, permitiendo a las organizaciones fortalecer sus mecanismos de protección y respuesta (Yusuf et al., 2025).

De manera complementaria, la creación de imágenes forenses constituye un pilar metodológico dentro de la informática forense, ya que garantiza la preservación íntegra de los datos originales. Este procedimiento permite a los analistas trabajar sobre copias exactas y verificadas, evitando cualquier alteración accidental o intencional de la evidencia, y asegurando así su validez desde el punto de vista legal y técnico. En este sentido, la literatura especializada reconoce el análisis de imágenes forenses como una herramienta fundamental en las investigaciones modernas, ya que posibilita la adquisición de datos de forma controlada, reproducible y verificable (Zhang, 2022). Además, avances recientes en el campo de la imagen forense han ampliado su aplicación a entornos más complejos, como el análisis de imágenes médicas digitales, donde la detección de manipulaciones requiere el estudio detallado de metadatos y artefactos específicos (Oh et al., 2026).

Por otra parte, estudios contemporáneos subrayan que la evolución de las tecnologías de almacenamiento y procesamiento de datos plantea nuevos retos para la informática forense, especialmente en lo relacionado con la integridad, autenticidad y confiabilidad de las imágenes forenses generadas (Dedouit et al., 2025). Estos desafíos exigen el uso de estándares rigurosos y herramientas

avanzadas que permitan garantizar la fidelidad de las copias obtenidas y su correcta interpretación durante el análisis. En consecuencia, tanto la recuperación de datos como la generación de imágenes forenses no solo constituyen procesos técnicos esenciales, sino también elementos críticos para asegurar la validez de la evidencia digital en investigaciones contemporáneas, contribuyendo de manera significativa al esclarecimiento de incidentes y a la administración de justicia en el entorno digital.

Desde una perspectiva conceptual, la recuperación de datos borrados puede definirse como el conjunto de técnicas, procedimientos y herramientas orientadas a restaurar información eliminada, dañada o inaccesible en diversos dispositivos de almacenamiento, tales como discos duros, memorias USB, tarjetas SD, dispositivos móviles o sistemas en la nube. Este proceso es posible debido a la forma en que los sistemas operativos gestionan la eliminación de archivos: en lugar de borrar inmediatamente los datos, el sistema marca el espacio ocupado como disponible para ser sobrescrito. Mientras dicho espacio no sea reutilizado, la información puede ser recuperada mediante herramientas especializadas.

La importancia de la recuperación de datos en el análisis forense radica en su capacidad para revelar evidencias que, en apariencia, han sido eliminadas de forma definitiva. En muchos casos, los archivos borrados contienen información crítica que permite reconstruir actividades sospechosas, como comunicaciones, transacciones o accesos no autorizados. Asimismo, estos datos aportan contexto a la investigación, facilitando la construcción de líneas de tiempo y la identificación de patrones de comportamiento. Desde el punto de vista legal, la recuperación de información eliminada ha sido determinante en la resolución de casos de fraude, espionaje industrial, delitos informáticos y acoso digital.

Para llevar a cabo este proceso, se emplean diversas técnicas de recuperación de datos. La recuperación a nivel de sistema de archivos se basa en el análisis de estructuras internas como tablas FAT, la Master File Table (MFT) en sistemas NTFS o los inodos en sistemas

EXT, lo que permite reconstruir nombres de archivos, rutas y marcas de tiempo. Por otro lado, la recuperación a nivel físico accede directamente a los sectores del disco, permitiendo recuperar fragmentos de información incluso cuando el sistema de archivos ha sido dañado o parcialmente sobrescrito. Asimismo, la recuperación de memoria volátil se centra en la obtención de datos almacenados en la memoria RAM, como credenciales, procesos activos o conexiones de red, los cuales son altamente efímeros. Finalmente, la recuperación en dispositivos móviles emplea técnicas especializadas de extracción lógica o física, adaptadas a la arquitectura de los sistemas operativos móviles.

En cuanto a las herramientas utilizadas en este ámbito, destacan soluciones como Autopsy/Sleuth Kit, que permiten el análisis de discos y la recuperación de archivos eliminados; Foremost y PhotoRec, que utilizan firmas digitales para reconstruir archivos; R-Studio Forensic, orientado a la recuperación avanzada de datos; y EnCase Forensic, ampliamente utilizado en investigaciones judiciales. Estas herramientas proporcionan capacidades avanzadas para el análisis de sistemas complejos y la reconstrucción de información crítica.

Por otra parte, el concepto de imagen forense se refiere a la creación de una copia exacta, bit a bit, de un dispositivo de almacenamiento o de memoria. Esta copia incluye no solo los archivos visibles, sino también los datos eliminados, el espacio no asignado y los sectores ocultos, lo que garantiza una representación completa del estado del sistema en el momento de la adquisición. A diferencia de una copia convencional, la imagen forense preserva la integridad de los datos mediante el uso de funciones hash, lo que permite verificar que no se han producido modificaciones durante el proceso.

Existen diferentes tipos de imágenes forenses, entre las que se incluyen la imagen completa o bit a bit, que copia la totalidad del dispositivo; la imagen lógica, que se limita a los archivos visibles; la imagen de memoria, utilizada para capturar el estado de la RAM; y la imagen en la

nube, que permite extraer datos de entornos distribuidos mediante mecanismos seguros.

El procedimiento para la creación de una imagen forense sigue una serie de pasos rigurosos que garantizan la integridad de la evidencia. En primer lugar, se prepara el entorno utilizando dispositivos write-blocker para evitar modificaciones accidentales. Posteriormente, se identifica y documenta el dispositivo, registrando sus características físicas. A continuación, se generan valores hash del contenido original, se realiza la copia bit a bit mediante herramientas especializadas y se verifica la integridad de la copia comparando los valores hash obtenidos. Finalmente, la imagen forense es almacenada en un entorno seguro junto con la documentación del proceso.

Un ejemplo práctico ilustra la importancia de estos procedimientos: en una investigación por fraude financiero, un analista forense recibe un disco duro sospechoso. Tras documentar el dispositivo y asegurar su integridad, se crea una imagen forense utilizando herramientas especializadas. Posteriormente, el análisis de la imagen permite recuperar documentos eliminados que evidencian transacciones ilícitas. Estos hallazgos, junto con la documentación del proceso y los valores hash, son presentados como evidencia válida en el proceso judicial.

La recuperación de datos borrados y la creación de imágenes forenses constituyen componentes esenciales dentro del análisis forense digital. Su correcta aplicación no solo permite acceder a información crítica que ha sido ocultada, sino que también garantiza la integridad, autenticidad y validez legal de la evidencia, contribuyendo de manera significativa al éxito de las investigaciones en entornos digitales.

4.5. Gestión, documentación y presentación de evidencias digitales en el análisis forense

En la informática forense, el proceso investigativo no concluye con la simple identificación, adquisición y análisis de los datos digitales. Si bien estas etapas son esenciales para localizar y preservar la

evidencia, el verdadero valor probatorio y técnico de una investigación forense se materializa cuando los hallazgos son debidamente documentados y presentados de forma clara, estructurada, verificable y legalmente válida. En este sentido, la forma en que el analista forense redacta sus conclusiones y expone sus resultados puede determinar si la evidencia digital es aceptada o rechazada en un proceso judicial, una auditoría interna o una investigación disciplinaria. Diversos estudios han señalado que la documentación constituye un componente crítico dentro de los modelos de investigación forense, ya que permite garantizar la coherencia, reproducibilidad y transparencia de los procedimientos realizados, fortaleciendo la credibilidad de los resultados obtenidos (Abbas, 2015).

La documentación de evidencias cumple una función fundamental, ya que garantiza la trazabilidad de todas las actuaciones realizadas durante el análisis técnico. Cada procedimiento, herramienta utilizada, resultado obtenido y decisión tomada debe quedar registrado de manera precisa, con el fin de demostrar que el trabajo fue desarrollado bajo criterios metodológicos rigurosos y respetando los principios de integridad y autenticidad de la evidencia. En este sentido, la literatura especializada destaca que la transformación de los datos digitales en evidencia válida implica superar diversos desafíos conceptuales y metodológicos, dado que la naturaleza intangible y fácilmente manipulable de la información digital exige procesos estrictos de validación y verificación (Biedermann y Kotsoglou, 2020). Asimismo, se han propuesto enfoques innovadores para reforzar la preservación y documentación de la evidencia, como el uso de tecnologías basadas en blockchain, que permiten garantizar la inmutabilidad y trazabilidad de los registros forenses (AlKhanafseh y Surakhi, 2024).

Paralelamente, la presentación de resultados permite traducir los hallazgos técnicos a un lenguaje comprensible para actores que, en muchas ocasiones, no poseen formación especializada en tecnologías de la información, como jueces, fiscales, abogados, directivos o comités disciplinarios. Este proceso implica no solo

la correcta estructuración de informes, sino también la capacidad de contextualizar la evidencia dentro del marco del caso investigado. Investigaciones recientes han demostrado que el análisis de datos provenientes de entornos digitales complejos, como redes sociales o plataformas en la nube, requiere metodologías específicas para su interpretación y presentación, debido a la gran cantidad de información y a la diversidad de formatos involucrados (Arshad et al., 2025; Karagiannis y Vergidis, 2021).

Además, estudios basados en casos reales han evidenciado que la transición de una traza digital a una evidencia admisible en juicio no es un proceso automático, sino que implica una cuidadosa validación de los datos, una adecuada contextualización y una presentación clara que permita a las autoridades comprender su relevancia dentro del caso (Bérubé et al., 2025). En este sentido, la fase de documentación y presentación no solo cumple una función técnica, sino también estratégica, ya que influye directamente en la interpretación de la evidencia y en la toma de decisiones legales. Por ello, esta etapa exige no solo conocimientos técnicos avanzados, sino también competencias en redacción profesional, comunicación efectiva, síntesis analítica y objetividad pericial, elementos fundamentales para garantizar la correcta valoración de la evidencia digital en cualquier contexto investigativo.

Desde una perspectiva metodológica, la documentación forense puede definirse como el registro sistemático, cronológico y detallado de cada una de las actividades desarrolladas durante una investigación digital. Para que dicha documentación sea válida y útil, debe ajustarse a una serie de principios fundamentales. En primer lugar, debe regirse por la exactitud, es decir, la información registrada debe reflejar fielmente lo observado, sin alteraciones ni interpretaciones subjetivas. En segundo lugar, debe garantizar la integridad, evitando omitir datos relevantes, incluso si estos no respaldan la hipótesis inicial del investigador. Asimismo, debe cumplir con el principio de claridad, de modo que el contenido resulte comprensible para lectores no especializados. A ello

se suma la trazabilidad, que implica que otro perito pueda reproducir el procedimiento y obtener resultados equivalentes, y la objetividad, que exige centrar el informe en hechos verificables y no en opiniones personales.

En la práctica, la documentación de evidencias suele estructurarse a través de un informe forense, el cual integra diversos componentes esenciales. Entre ellos se encuentran los datos de identificación del caso, como número de expediente, fecha, institución solicitante y responsables del análisis. También se incluye una descripción detallada de la evidencia, especificando el tipo de dispositivo, sus características técnicas, número de serie, estado físico y condiciones de hallazgo. Otro apartado indispensable corresponde a la metodología aplicada, donde se detallan las herramientas empleadas, los procedimientos ejecutados, la generación de valores hash y las medidas de preservación adoptadas. Posteriormente, se describe el análisis realizado, incorporando hallazgos técnicos, reconstrucción de eventos, evidencias recuperadas y correlación entre los distintos elementos observados. Asimismo, deben documentarse las limitaciones encontradas durante la investigación, tales como cifrado, corrupción de datos o restricciones de acceso, ya que estas pueden influir en el alcance de las conclusiones. Finalmente, el informe debe contener conclusiones técnicas claras y anexos documentales, como capturas de pantalla, registros de logs, tablas comparativas, imágenes forenses y copias de verificación hash.

En función de la audiencia a la que se dirija, en informática forense pueden elaborarse distintos tipos de informes. El informe técnico está orientado a especialistas, otros peritos o equipos de seguridad, por lo que incluye un alto nivel de detalle, terminología especializada, comandos utilizados, rutas de acceso, resultados de herramientas y explicaciones metodológicas profundas. El informe ejecutivo, en cambio, se elabora con un enfoque más sintético y accesible, utilizando un lenguaje menos técnico y apoyándose en tablas, gráficos, diagramas o indicadores porcentuales para facilitar la toma de decisiones por parte de directivos,

gerentes o autoridades institucionales. Por su parte, el informe judicial responde a las exigencias formales y legales establecidas por el sistema normativo de cada país, por lo que debe redactarse con máxima precisión, imparcialidad y respaldo documental, considerando además que puede ser objeto de revisión, impugnación o contrainterrogatorio en sede judicial.

La presentación de resultados no se limita al informe escrito. En numerosos casos, el analista o perito debe defender oralmente sus hallazgos ante distintas audiencias. En procesos judiciales, esto puede implicar una declaración pericial en la que se explique, de manera clara y ordenada, cómo fue recolectada, preservada y analizada la evidencia digital. En entornos corporativos, puede requerirse la exposición de riesgos, vulnerabilidades y hechos detectados ante la alta dirección o los responsables de cumplimiento normativo. También son frecuentes las presentaciones en auditorías internas o en reuniones de contraste con equipos de seguridad, donde los resultados deben ser discutidos con fines preventivos o correctivos.

Para que estas exposiciones resulten efectivas, es recomendable utilizar recursos visuales como gráficos, diagramas de flujo, cronologías o infografías, ya que facilitan la comprensión de procesos complejos. Del mismo modo, el perito debe prepararse para responder preguntas críticas relacionadas con la integridad de la evidencia, la validez del procedimiento seguido o la confiabilidad de las herramientas empleadas. En este sentido, una exposición sólida no depende únicamente del conocimiento técnico, sino también de la capacidad de explicar con precisión conceptos complejos en un lenguaje moderado y accesible para audiencias no expertas.

A pesar de su relevancia, la fase de documentación y presentación de resultados suele verse afectada por errores que comprometen la calidad y validez del trabajo forense. Entre los errores más comunes se encuentra la omisión del cálculo o registro de valores hash, la descripción ambigua de procedimientos sin detallar herramientas ni comandos empleados, la

emisión de conclusiones exageradas o no respaldadas técnicamente, la falta de mención de limitaciones relevantes y el uso de un lenguaje excesivamente técnico cuando la audiencia no posee conocimientos especializados. Estas deficiencias pueden debilitar la credibilidad del informe y poner en riesgo la admisibilidad de la evidencia.

Un ejemplo práctico permite ilustrar la importancia de esta etapa. En una investigación por acoso laboral digital, se incauta un teléfono móvil del cual se sospecha que contiene mensajes eliminados relacionados con el caso. En primer lugar, se realiza una imagen forense del dispositivo mediante una herramienta especializada, preservando la integridad de la evidencia. Posteriormente, se recuperan mensajes borrados que constituyen prueba directa de los hechos denunciados. A partir de estos hallazgos, el perito elabora un informe técnico detallado, en el que incorpora capturas de pantalla, rutas de archivos, valores hash y la descripción precisa del procedimiento utilizado. Paralelamente, se genera un informe ejecutivo dirigido al equipo legal, en el que se resumen únicamente los mensajes relevantes en un lenguaje claro y orientado a la toma de decisiones. Finalmente, durante el juicio, el perito presenta oralmente sus conclusiones, explicando paso a paso cómo se garantizó que la evidencia no fue alterada y cómo los resultados obtenidos sustentan objetivamente las conclusiones del caso.

La documentación de evidencias y la adecuada presentación de resultados constituyen etapas esenciales de la informática forense, ya que convierten el análisis técnico en prueba útil, comprensible y jurídicamente sostenible. Su correcta elaboración fortalece la credibilidad del peritaje, garantiza la transparencia del proceso investigativo y permite que los hallazgos digitales cumplan efectivamente su función probatoria en contextos judiciales, disciplinarios y organizacionales.

REFERENCIAS

BIBLIOGRÁFICAS

Abbas, T. M. J. (2015). *Studying the documentation process in digital forensic investigation frameworks/models. Journal of Al-Nahrain University – Science, 18*(4), 153–162. https://doi.org/10.22401/JNUS.18.4.21

Abdullah, M., Nawaz, M. M., Saleem, B., Zahra, M., Ashfaq, E. b., & Muhammad, Z. (2025). Evolution Cybercrime—Key Trends, Cybersecurity Threats, and Mitigation Strategies from Historical Data. *Analytics, 4*(3), 25. https://doi.org/10.3390/analytics4030025

Abeysekera I. (2026). Open source tools: an evaluation for digital forensic investigations. *Frontiers in research metrics and analytics, 11,* 1790333. https://doi.org/10.3389/frma.2026.1790333

Abirami, A., Lakshmanaprakash, S., Priya, R. L., Hirlekar, V., & Dalal, B. (2024). *Proactive analysis and detection of cyber-attacks using deep learning techniques. Indian Journal of Science and Technology, 17*(15), 1596–1605. https://doi.org/10.17485/IJST/v17i15.3044

Admass, W. S., Munaye, Y. Y., & Diro, A. A. (2024). Cyber security: State of the art, challenges and future directions. *Cyber Security and Applications*, 2, 100031. https://doi.org/10.1016/j.csa.2023.100031

Alanazi, M., Mahmood, A., & Chowdhury, M. J. M. (2023). SCADA vulnerabilities and attacks: A review of the state-of-the-art and open issues. *Computers & Security, 125*, 103028. https://doi.org/10.1016/j.cose.2022.103028

Aleisa, N. (2026). The Study of Digital Forensics in KSA: Education, and Prosecution Capabilities: A Needs-Based Analysis. *Electronics*, *15*(2), 316. https://doi.org/10.3390/electronics15020316

Alghamdi, A. (2022). *The role of social engineering in cybersecurity and its impact. Journal of Information Security, 13*, 363–379. https://doi.org/10.4236/jis.2022.134020

Alhamed, M., & Rahman, M. M. H. (2023). A Systematic Literature Review on Penetration Testing in Networks: Future Research Directions. *Applied Sciences*, *13*(12), 6986. https://doi.org/10.3390/app13126986

AlKhanafseh, M., & Surakhi, O. (2024). Evidence Preservation in Digital Forensics: An Approach Using Blockchain and LSTM-Based Steganography. *Electronics*, *13*(18), 3729. https://doi.org/10.3390/electronics13183729

Almaiah, M. A., Saqr, L. M., Al-Rawwash, L. A., Altellawi, L. A., Al-Ali, R., & Almomani, O. (2024). Classification of cybersecurity threats, vulnerabilities and countermeasures in database systems. *Computers, Materials & Continua, 81*(2), 3189–3220. https://doi.org/10.32604/cmc.2024.057673

Alqaydi, S., Othman, A., & Abuljadayel, L. H. (2024). *The role of AI in cyber security: Safeguarding digital identity. Journal of Information Security, 15*, 245–278. https://doi.org/10.4236/jis.2024.152015

Alshammari, S. S., Soh, B., & Li, A. (2025). Understanding Social Engineering Victimisation on Social Networking Sites: A Comprehensive Review of Factors Influencing User Susceptibility to Cyber-Attacks. *Information*, *16*(2), 153. https://doi.org/10.3390/info16020153

Alsodi, O., Zhou, X., Gururajan, R., Shrestha, A., & Btoush, E. (2025). From Tweets to Threats: A Survey of Cybersecurity Threat Detection Challenges, AI-Based Solutions and Potential Opportunities in X. *Applied Sciences*, *15*(7), 3898. https://doi.org/10.3390/app15073898

Alsowail, R. A., & Al-Shehari, T. (2022). Techniques and countermeasures for preventing insider threats. *PeerJ. Computer science*, *8*, e938. https://doi.org/10.7717/peerj-cs.938

Amine, A. M., Chakir, E. M., Issam, T., & Idrissi Khamlichi, Y. (2023). A review of cybersecurity management standards applied in higher education institutions. *International Journal of Safety and Security Engineering, 13*(6), 1109–1116. https://www.iieta.org/download/file/fid/115034

Antariksa, M. D. S., Perangin Angin, M., & Widodo, A. P. (2025). COBIT 2019 framework in IT governance: A systematic literature review of implementation challenges and benefits across various industry sectors. *Journal of Renewable Energy Electrical and Computer Engineering*, *5*(1), 99–105. https://doi.org/10.29103/jreece.v5i1.19501

Antunes, M., Maximiano, M., & Gomes, R. (2022). A Client-Centered Information Security and Cybersecurity Auditing Framework. *Applied Sciences*, *12*(9), 4102. https://doi.org/10.3390/app12094102

Arshad, M., Ahmad, A., Onn, C. W., & Sam, E. A. (2025). *Investigating methods for forensic analysis of social media data to support criminal investigations. Frontiers in Computer Science*, 7, Article 1566513. https://doi.org/10.3389/fcomp.2025.1566513

Aslan, Ö., Aktuğ, S. S., Ozkan-Okay, M., Yilmaz, A. A., & Akin, E. (2023). A comprehensive review of cyber security vulnerabilities, threats, attacks, and solutions. Electronics, *12*(6), 1333. https://doi.org/10.3390/electronics12061333

Awan, M., & Alam, A. (2025). Cybersecurity Threats and Defensive Strategies for Small and Medium Firms: A Systematic Mapping Study. *Administrative Sciences, 15*(12), 481. https://doi.org/10.3390/admsci15120481

Baltuttis, D., Teubner, T., & Adam, M. T. P. (2024). A typology of cybersecurity behavior among knowledge workers. *Computers & Security, 140*, 103741. https://doi.org/10.1016/j.cose.2024.103741

Berrios, S., Leiva, D., Olivares, B., Allende-Cid, H., & Hermosilla, P. (2025). Systematic Review: Malware Detection and Classification in Cybersecurity. *Applied Sciences, 15*(14), 7747. https://doi.org/10.3390/app15147747

Bérubé, M., Beaulieu, L.-A., Allard, S., & Denault, V. (2025). *From digital trace to evidence: Challenges and insights from a trial case study. Science & Justice, 65*(5), Article 101306. https://doi.org/10.1016/j.scijus.2025.101306

Biedermann, A., & Kotsoglou, K. N. (2020). Digital evidence exceptionalism? A review and discussion of conceptual hurdles in digital evidence transformation. *Forensic science international. Synergy, 2*, 262–274. https://doi.org/10.1016/j.fsisyn.2020.08.004

Bossler, A. M., & Berenblum, T. (2019). *Introduction: New directions in cybercrime research. Journal of Crime and Justice, 42*(5), 495–499. https://doi.org/10.1080/0735648X.2019.1692426

Bouramdane, A.-A. (2023). Cyberattacks in Smart Grids: Challenges and Solving the Multi-Criteria Decision-Making for Cybersecurity Options, Including Ones That Incorporate Artificial Intelligence, Using an Analytical Hierarchy Process. *Journal of Cybersecurity and Privacy, 3*(4), 662-705. https://doi.org/10.3390/jcp3040031

Buchler, N., La Fleur, C. G., Hoffman, B., Rajivan, P., Marusich, L., & Lightner, L. (2018). Cyber Teaming and Role Specialization in a Cyber Security Defense Competition. *Frontiers in psychology*, *9*, 2133. https://doi.org/10.3389/fpsyg.2018.02133

Chen, S., Hao, M., Ding, F., Jiang, D., Dong, J., Zhang, S., Guo, Q., & Gao, C. (2023). Exploring the global geography of cybercrime and its driving forces. *Humanities & social sciences communications, 10*(1), 71. https://doi.org/10.1057/s41599-023-01560-x

Chindrus, C., & Caruntu, C.-F. (2023). Securing the Network: A Red and Blue Cybersecurity Competition Case Study. *Information, 14*(11), 587. https://doi.org/10.3390/info14110587

Chu, W.-L., Lin, C.-J., & Chang, K.-N. (2019). Detection and Classification of Advanced Persistent Threats and Attacks Using the Support Vector Machine. *Applied Sciences*, *9*(21), 4579. https://doi.org/10.3390/app9214579

Cremer, F., Sheehan, B., Fortmann, M., Kia, A. N., Mullins, M., Murphy, F., & Materne, S. (2022). Cyber risk and cybersecurity: A systematic review of data availability. *The Geneva Papers on Risk and Insurance: Issues and Practice, 47*(3), 698–736. https://doi.org/10.1057/s41288-022-00266-6

Cruz, C. (2024). Innovative Learning in a Digital Forensics Laboratory: Tools and Techniques for Data Recovery. *Applied Sciences,* 14(23), 11095. https://doi.org/10.3390/app142311095

Curtis, J., & Oxburgh, G. (2022). *Understanding cybercrime in 'real world' policing and law enforcement. The Police Journal: Theory, Practice and Principles, 96*(4). https://doi.org/10.1177/0032258X221107584

Dalal, A. (2021). *Exploring next-generation cybersecurity tools for advanced threat detection and incident response. SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.5424096

D'Anna, T., Puntarello, M., Cannella, G., Scalzo, G., Buscemi, R., Zerbo, S., & Argo, A. (2023). The Chain of Custody in the Era of Modern Forensics: From the Classic Procedures for Gathering Evidence to the New Challenges Related to Digital Data. *Healthcare (Basel, Switzerland), 11*(5), 634. https://doi.org/10.3390/healthcare11050634

Darem, A., Al-Hashmi, A., Alkhaldi, T. M., Alashjaee, A. M., Alanazi, S., & Ebad, S. (2023). Cyber threats classifications and countermeasures in banking and financial sector. *IEEE Access*. https://doi.org/10.1109/ACCESS.2023.3327016

Darmawan, R. K., & Cahyono, A. D. (2025). Implementation of Zero-Knowledge Encryption in a Web-Based Password Manager. I*nternational Journal Software Engineering and Computer Science (IJSECS), 5*(2), 633-644. https://doi.org/10.35870/ijsecs.v5i2.4207

De Nobrega, K. M., Rutkowski, A.-F., & Saunders, C. (2024). *The whole of cyber defense: Syncing practice and theory. The Journal of Strategic Information Systems, 33*(4), 101861. https://doi.org/10.1016/j.jsis.2024.101861

Dedouit, F., Ducloyer, M., Elifritz, J., Adolphi, N. L., Yi-Li, G. W., Decker, S., Ford, J., Kolev, Y., & Thali, M. (2025). The current state of forensic imaging - perspectives. *International journal of legal medicine, 139*(6), 2819–2827. https://doi.org/10.1007/s00414-025-03466-6

Dhanaraj, A. (2025). The evolution of cyber threats: From traditional attacks to AI-powered challenges. *European Journal of Computer Science and Information Technology, 13*(36), 50–61. https://doi.org/10.37745/ejcsit.2013/vol13n365061

Dimitriadis, A., Ivezic, N., Kulvatunyou, B., & Mavridis, I. (2020). D4I - Digital forensics framework for reviewing and investigating cyber attacks. *Array (New York, N.Y.)*, *5*, 10.1016/j.array.2019.100015. https://doi.org/10.1016/j.array.2019.100015

Dimitrov, G. (2020). *A brief history of cyber intelligence. American Intelligence Journal, 37*(1), 107–114.https://www.jstor.org/stable/27087688

Dizon, M. A. C., & Meehan, A. (2024). Technical principles and protocols of encryption and their significance and effects on technology regulation. *Information & Communications Technology Law, 34*(2), 79–105. https://doi.org/10.1080/13600834.2024.2404280

Ewoh, P., & Vartiainen, T. (2024). Vulnerability to cyberattacks and sociotechnical solutions for health care systems: Systematic review. *Journal of Medical Internet Research, 26*, e46904. https://doi.org/10.2196/46904

Fehis, S., Nouali, O., & Kechadi, T. (2021). Secure encryption key management as a SecaaS based on Chinese wall security policy. *Journal of Information Security and Applications, 63*, 102975. https://doi.org/10.1016/j.jisa.2021.102975

Ferrer-Oliva, M., Medina-Merodio, J. A., Martínez-Herraiz, J. J., & Cilleruelo-Rodríguez, C. (2025). Relational Framework of Cyberattacks: Empirical Evidence from Multistage Incidents. Sensors *(Basel, Switzerland)*, *25*(23), 7124. https://doi.org/10.3390/s25237124

Gaifulina, A. (2025). *The concept of red team and blue team synergy as a factor in enhancing an organization's resilience to cyberattacks. The American Journal of Applied Sciences, 7*(11), 55–60. https://doi.org/10.37547/tajas/Volume07Issue11-06

Gilbert, C., Gilbert, M. A., & Jnr, M. D. (2025). Detection and response strategies for advanced persistent threats (APTs). *International Journal of Scientific Research and Modern Technology, 4*(4), 5–21. https://doi.org/10.38124/ijsrmt.v4i4.367

Greavu-Şerban, V., Constantin, F., & Necula, S.-C. (2025). Exploring Heuristics and Biases in Cybersecurity: A Factor Analysis of Social Engineering Vulnerabilities. *Systems, 13*(4), 280. https://doi.org/10.3390/systems13040280

Grigaliūnas, Š., Brūzgienė, R., & Venčkauskas, A. (2023). The Method for Identifying the Scope of Cyberattack Stages in Relation to Their Impact on Cyber-Sustainability Control over a System. *Electronics*, *12*(3), 591. https://doi.org/10.3390/electronics12030591

Guttman, B., White, D. R., & Walraven, T. (2022). *Digital evidence preservation: Considerations for evidence handlers*. National Institute of Standards and Technology. https://doi.org/10.6028/NIST.IR.8387

Haji, I. A., Mohammed, S. D., & Mousa, K. M. (2026). *Blockchain based chain of custody and digital evidence legality in post conflict prosecutions. Frontiers in Blockchain, 9*, Article 1801364. https://doi.org/10.3389/fbloc.2026.1801364

Hatfield, J. M. (2019*). Virtuous human hacking: The ethics of social engineering in penetration-testing. Computers & Security, 83*, 354–366. https://doi.org/10.1016/j.cose.2019.02.012

International Organization for Standardization, & International Electrotechnical Commission. (2012). *ISO/IEC 27032:2012 Information technology—Security techniques—Guidelines for cybersecurity*. https://itgeu.blob.core.windows.net/files/download/3911-isoiec-27032%7Bed1.0%7Den.pdf

International Organization for Standardization, & International Electrotechnical Commission. (2022). *ISO/IEC 27001:2022 Information security, cybersecurity and privacy protection—Information security management systems—Requirements* (3rd ed.). https://www.exactls.com/wp-content/uploads/2025/02/ISO_IEC-270012022-ed.3.pdf

Jones, N., Whaiduzzaman, M., Jan, T., Adel, A., Alazab, A., & Alkreisat, A. (2025). A CIA Triad-Based Taxonomy of Prompt Attacks on Large Language Models. *Future Internet, 17*(3), 113. https://doi.org/10.3390/fi17030113

Jouini, M., Ben Arfa Rabai, L., & Ben Aissa, A. (2014). Classification of security threats in information systems. *Procedia Computer Science, 32*, 489–496. https://doi.org/10.1016/j.procs.2014.05.452

Kara, I. (2021). *Cyber-espionage malware attacks detection and analysis: A case study. Journal of Computer Information Systems, 62*(6), 1253–1270. https://doi.org/10.1080/08874417.2021.2004566

Karagiannis, C., & Vergidis, K. (2021). Digital Evidence and Cloud Forensics: Contemporary Legal Challenges and the Power of Disposal. *Information, 12*(5), 181. https://doi.org/10.3390/info12050181

Kitsios, F., Chatzidimitriou, E., & Kamariotou, M. (2023). The ISO/IEC 27001 Information Security Management Standard: How to Extract Value from Data in the IT Sector. *Sustainability, 15*(7), 5828. https://doi.org/10.3390/su15075828

Klasén, L., Fock, N., & Forchheimer, R. (2024). *The invisible evidence: Digital forensics as key to solving crimes in the digital age. Forensic Science International, 362*, 112133. https://doi.org/10.1016/j.forsciint.2024.112133

Ksibi, S., Jaidi, F., & Bouhoula, A. (2022). A Comprehensive Study of Security and Cyber-Security Risk Management within e-Health Systems: Synthesis, Analysis and a Novel Quantified Approach. *Mobile networks and applications: MONET*, 1–21. Advance online publication. https://doi.org/10.1007/s11036-022-02042-1

Kumar, P. R., & Goel, S. (2025). A secure and efficient encryption system based on adaptive and machine learning for securing data in fog computing. *Scientific reports, 15*(1), 11654. https://doi.org/10.1038/s41598-025-92245-9

Kuzminykh, L., Ghita, B., & Shiaeles, S. (2021). Comparative analysis of cryptographic key management systems. *20th International Conference on Next Generation Teletraffic and Wired/Wireless Advanced Networks and Systems (NEW2AN 2020) and 13th Conference on the Internet of Things and Smart Spaces (ruSMART 2020)*. St. Petersburg, Russia.

Lallie, H. S., Thompson, A., Titis, E., & Stephens, P. (2025). Analysing Cyber Attacks and Cyber Security Vulnerabilities in the University Sector. *Computers, 14*(2), 49. https://doi.org/10.3390/computers14020049

Li, Y., & Liu, Q. (2021). A comprehensive review study of cyber-attacks and cyber security: Emerging trends and recent developments. *Energy Reports, 7*, 8176–8186. https://doi.org/10.1016/j.egyr.2021.08.126

Liu, X., Ahmad, S. F., Anser, M. K., Ke, J., Irshad, M., Ul-Haq, J., & Abbas, S. (2022). Cyber security threats: A never-ending challenge for e-commerce. *Frontiers in psychology, 13*, 927398. https://doi.org/10.3389/fpsyg.2022.927398

Lorenzini, G., Shaw, D. M., & Elger, B. S. (2022). It takes a pirate to know one: ethical hackers for healthcare cybersecurity. *BMC medical ethics, 23*(1), 131. https://doi.org/10.1186/s12910-022-00872-y

Luidold, C., & Jungbauer, C. (2024). Cybersecurity policy framework requirements for the establishment of highly interoperable and interconnected health data spaces. *Frontiers in medicine, 11*, 1379852. https://doi.org/10.3389/fmed.2024.1379852

Mandru, S. (2020). *Ethical hacking and penetration testing: Analyzing the role of ethical hacking and penetration testing in identifying vulnerabilities and improving overall cybersecurity defenses. International Journal of Core Engineering & Management, 6*(7). https://ijcem.in/wp-content/uploads/2024/08/ETHICAL-HACKING-AND-PENETRATION-TESTING-ANALYZING-THE-ROLE-OF-ETHICAL-HACKING-AND-PENETRATION-TESTING-IN-IDENTIFYING-VULNERABILITIES-AND-IMPROVING-OVERALL-CYBERSECURITY-DEFENSES.pdf

McCumber, C. J. R. (1991). *Information systems security: A comprehensive model. 14th National Computer Security Conference*, Washington, D.C., Estados Unidos.

McIntosh, T. R., Susnjak, T., Liu, T., Watters, P., Xu, D., Liu, D., Nowrozy, R., & Halgamuge, M. N. (2024). From COBIT to ISO 42001: Evaluating cybersecurity frameworks for opportunities, risks, and regulatory compliance in commercializing large language models. *Computers & Security*, *144*, 103964. https://doi.org/10.1016/j.cose.2024.103964

Mohamed, A. M. E., Alnor, N. H. A., Mohammed, O. A. A., Al-Matari, E. M., Alhebri, A., & Ah, A. (2024). The impact of information technology governance according to the COBIT on performance. *International Journal of Advanced and Applied Sciences, 11*(3), 127–136. https://www.science-gate.com/IJAAS/Articles/2024/2024-11-03/1021833ijaas202403014.pdf

Montañez, R., Golob, E., & Xu, S. (2020). Human Cognition Through the Lens of Social Engineering Cyberattacks. *Frontiers in psychology*, 11, 1755. https://doi.org/10.3389/fpsyg.2020.01755

Morić, Z., Dakić, V., & Ogrizek Biškupić, I. (2026). An Empirical Assessment of Digital Forensic Process Reliability Using Integrated ISO/IEC 27037 and 27041 Standards. *Journal of Cybersecurity and Privacy, 6*(2), 57. https://doi.org/10.3390/jcp6020057

Nachrowi, E., Nurhadryani, Y., & Sukoco, H. (2020). Evaluation of governance and management of information technology services using COBIT 2019 and ITIL 4. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), 4*(4), 764–774. https://doi.org/10.29207/resti.v4i4.2265

Nath, S., Summers, K., Baek, J., & Ahn, G.-J. (2024). Digital evidence chain of custody: Navigating new realities of digital forensics. *International Conference on Trust, Privacy and Security in Intelligent Systems, and Applications*. Washington, D.C., United States.

Nisioti, A., Loukas, G., Mylonas, A., & Panaousis, E. (2023). *Forensics for multi-stage cyber incidents: Survey and future directions. Forensic Science International: Digital Investigation, 44*, Article 301480. https://doi.org/10.1016/j.fsidi.2022.301480

Nonum, E. O., Avwokuruaye, O., & Umar, A. M. (2025). *Social engineering: Understanding human factors in cyber security. International Journal of Convergent and Informatics Science Research, 7*(9). https://doi.org/10.70382/hijcisr.v07i9.032

Oh, S., Lee, E., Lee, S., Park, J., Park, S., & Kim, G. (2026). Forensic detection of medical image manipulation using PACS and DICOM artifacts. *Journal of forensic sciences, 71*(1), 371–387. https://doi.org/10.1111/1556-4029.70191

Onome, O. A. (2023). *Computer forensics and advanced methodology. International Journal of Emerging Science and Engineering (IJESE), 11*(7), 1–15. https://doi.org/10.35940/ijese.G2552.0611723

Onwuadiamu, G. (2025). *Cybercrime in criminology: A systematic review of criminological theories, methods, and concepts. Journal of Economic Criminology, 8*, 100136. https://doi.org/10.1016/j.jeconc.2025.100136

Oppenheimer, H. (2024). How the process of discovering cyberattacks biases our understanding of cybersecurity. *Journal of Peace Research, 61*(1). https://doi.org/10.1177/00223433231217687

Patel, H. G., & Patel, J. (2015). *An analysis upon various process and model of digital forensic investigation: Developing a structured approach for digital forensic investigation. International Journal of Information Technology and Management, 8*(12). https://ignited.in/index.php/ijitm/article/view/594/1045

Phillips, K., Davidson, J. C., Farr, R. R., Burkhardt, C., Caneppele, S., & Aiken, M. P. (2022). Conceptualizing Cybercrime: Definitions, Typologies and Taxonomies. *Forensic Sciences*, *2*(2), 379-398. https://doi.org/10.3390/forensicsci2020028

Qureshi, M. B., Qureshi, M. S., Tahir, S., Anwar, A., Hussain, S., Uddin, M., & Chen, C.-L. (2022). Encryption Techniques for Smart Systems Data Security Offloaded to the Cloud. *Symmetry, 14*(4), 695. https://doi.org/10.3390/sym14040695

Qureshi, S. U., He, J., Tunio, S., Zhu, N., Nazir, A., Wajahat, A., Ullah, F., & Wadud, A. (2024). *Systematic review of deep learning solutions for malware detection and forensic analysis in IoT. Journal of King Saud University – Computer and Information Sciences, 36*(8), Article 102164. https://doi.org/10.1016/j.jksuci.2024.102164

Qusef, A., & Alkilani, H. (2022). The effect of ISO/IEC 27001 standard over open-source intelligence. *PeerJ. Computer science,* 8, e810. https://doi.org/10.7717/peerj-cs.810

Raghuwanshi, K. P., Dongare, A. M., Nimkande, S. A., & Thakare, D. V. (2025). The history and evolution of cyber attacks: A comprehensive study. *International Journal of Advanced Research in Computer and Communication Engineering,* 14(11). https://doi.org/10.17148/IJARCCE.2025.141160

Rahman, M. M., Dhakal, K., Gony, N., Shuvra, M. K., & Rahman, M. (2025). *AI integration in cybersecurity software: Threat detection and response. International Journal of Innovative Research and Scientific Studies, 8*(3). https://doi.org/10.53894/ijirss.v8i3.7403

Rana, S., Khoda Parast, F., Kelly, B., Wang, Y., & Kent, K. B. (2023). A comprehensive survey of cryptography key management systems. *Journal of Information Security and Applications, 78*, 103607. https://doi.org/10.1016/j.jisa.2023.103607

Razaque, A., Hariri, S., Alajlan, A. M., & Yoo, J. (2025). A comprehensive review of cybersecurity vulnerabilities, threats, and solutions for the Internet of Things at the network-cum-application layer. *Computer Science Review, 58*, 100789. https://doi.org/10.1016/j.cosrev.2025.100789

Reuben-Owoh, B., & Haig, E. (2025). A systematic review of voluntary cybersecurity standards and frameworks. *International Journal of Information Security, 24*, 206. https://doi.org/10.1007/s10207-025-01121-0

Roman, A.-S. (2023). Evaluating the Privacy and Utility of Time-Series Data Perturbation Algorithms. *Mathematics, 11*(5), 1260. https://doi.org/10.3390/math11051260

Safitra, M. F., Lubis, M., & Fakhrurroja, H. (2023). Counterattacking Cyber Threats: A Framework for the Future of Cybersecurity. *Sustainability, 15*(18), 13369. https://doi.org/10.3390/su151813369

Sakshi, Malik, A., & Sharma, A. K. (2023). *Blockchain-based digital chain of custody multimedia evidence preservation framework for internet-of-things. Journal of Information Security and Applications, 77*, 103579. https://doi.org/10.1016/j.jisa.2023.103579

Salas Riega, J. L., Riega, Y., Ninaquispe Soto, M. E., & Salas-Riega, J. M. (2025). Cybersecurity and the NIST framework: A systematic review of its implementation and effectiveness against cyber threats. *International Journal of Advanced Computer Science and Applications, 16*(6). https://doi.org/10.14569/IJACSA.2025.0160672

Salim, D. T., Singh, M. M., & Keikhosrokiani, P. (2023). A systematic literature review for APT detection and Effective Cyber Situational Awareness (ECSA) conceptual model. *Heliyon*, *9*(7), e17156. https://doi.org/10.1016/j.heliyon.2023.e17156

Santos, P., Abreu, R., Reis, M. J. C. S., Serôdio, C., & Branco, F. (2025). A Systematic Review of Cyber Threat Intelligence: The Effectiveness of Technologies, Strategies, and Collaborations in Combating Modern Threats. *Sensors (Basel, Switzerland), 25*(14), 4272. https://doi.org/10.3390/s25144272

Shah, Z., Kyaw, A., Truong, H. P., Ullah, I., & Levula, A. (2022). Forensic Investigation of Remnant Data on USB Storage Devices Sold in New Zealand. *Applied Sciences,* 12(12), 5928. https://doi.org/10.3390/app12125928

Shoufan, A., & Damiani, E. (2017). On inter-rater reliability of information security experts. *Journal of Information Security and Applications, 37*, 101–111. https://doi.org/10.1016/j.jisa.2017.10.006

Smith, K. J., Dhillon, G., & Carter, L. (2021). User values and the development of a cybersecurity public policy for the IoT. *International Journal of Information Management, 56*, 102123. https://doi.org/10.1016/j.ijinfomgt.2020.102123

Taherdoost, H. (2022). Understanding Cybersecurity Frameworks and Information Security Standards—A Review and Comprehensive Overview. *Electronics*, *11*(14), 2181. https://doi.org/10.3390/electronics11142181

Tariq, U., Ahmed, I., Bashir, A. K., & Shaukat, K. (2023). A Critical Cybersecurity Analysis and Future Research Directions for the Internet of Things: A Comprehensive Review. *Sensors (Basel, Switzerland),* 23(8), 4117. https://doi.org/10.3390/s23084117

Toivanen, J., & Luoma-aho, V. (2026). The life cycle approach to effective crisis communications in mitigating cyber threats and attacks. *Proceedings of the 21st International Conference on Cyber Warfare and Security, 21*(1), 494–504. https://doi.org/10.34190/iccws.21.1.4509

Tzavara, V., & Vassiliadis, S. (2024). Tracing the evolution of cyber resilience: A historical and conceptual review. *International Journal of Information Security, 23*, 1695–1719. https://doi.org/10.1007/s10207-023-00811-x

Ukwadinachi, N. (2025). *From Mitnick to modern database threats: Evolution of social engineering tactics and their impact on data integrity. Journal of Technology Studies, 50*(1), 14–29. https://doi.org/10.21061/jts.438

Urooj, S., Lata, S., Ahmad, S., Mehfuz, S., & Kalathil, S. (2023). Cryptographic data security for reliable wireless sensor network. *Alexandria Engineering Journal, 72*, 37–50. https://doi.org/10.1016/j.aej.2023.03.061

Vaya-Arboledas, Á., Ferrer-Oliva, M., & Medina-Merodio, J. A. (2025). Evolution and Perspectives in IT Governance: A Systematic Literature Review. *Computers, 14*(12), 520. https://doi.org/10.3390/computers14120520

Viano, E. C. (2017). *Cybercrime, organized crime, and societal responses*. Springer.

Viriyatama Lim, M., & Indah Fianty, M. (2023). Enhancing Information Technology Governance: A Comprehensive Evaluation Of The 2019 COBIT Framework In The Retail Industry. *International Journal of Science, Technology & Management,* 4(5), 1389–1395. https://doi.org/10.46729/ijstm.v4i5.955

Waelchli, S., & Walter, Y. (2025). *Reducing the risk of social engineering attacks using SOAR measures in a real world environment: A case study. Computers & Security, 148*, 104137. https://doi.org/10.1016/j.cose.2024.104137

Yulianto, S., Soewito, B., Lumban Gaol, F., & Kurniawan, A. (2025). *Enhancing cybersecurity resilience through advanced red-teaming exercises and MITRE ATT&CK framework integration: A paradigm shift in cybersecurity assessment. Cyber Security and Applications, 3*, 100077. https://doi.org/10.1016/j.csa.2024.100077

Yusuf, A. M., Sari, D. M., & Musawwir. (2025). *Digital forensics in open journal systems: Case study on security breach and data recovery. Journal of Embedded Systems, Security and Intelligent Systems, 6*(3), 522–557. https://doi.org/10.59562/jessi.v6i3.9778

Zaid, T., & Garai, S. (2024). Emerging trends in cybersecurity: A holistic view on current threats, assessing solutions, and pioneering new frontiers. *Blockchain in Healthcare Today, 7*, Article 302. https://doi.org/10.30953/bhty.v7.302

Zhang M. (2022). Forensic imaging: a powerful tool in modern forensic investigation. Forensic sciences research, 7(3), 385–392. https://doi.org/10.1080/20961790.2021.2008705

Zhang, W., Xing, J., & Li, X. (2026). *Penetration testing for system security: Methods and practical approaches* (arXiv:2505.19174v2). arXiv. https://arxiv.org/abs/2505.19174

Zhou, Y., Tiwari, M., Bernot, A., & Lin, K. (2024). *Metacrime and cybercrime: Exploring the convergence and divergence in digital criminality. Asian Journal of Criminology, 19*, 419–439. https://doi.org/10.1007/s11417-024-09436-y

Luis Eduardo Carrizo García

Es un profesional en Telecomunicaciones con Maestría en Ciberseguridad, con sólida formación académica y amplia experiencia en el ámbito de las tecnologías de la información, la docencia universitaria y la gestión académica. Su formación abarca desde la ingeniería en telecomunicaciones hasta estudios de posgrado en ciberseguridad, complementada con múltiples certificaciones en redes, seguridad informática, tecnologías educativas y competencias docentes. Es docente universitario, investigador y gestor académico, con experiencia en el diseño curricular, coordinación de procesos de titulación y planificación estratégica institucional. Ha liderado iniciativas orientadas a la mejora de la calidad educativa, elaboración de planes estratégicos (PEDI), planificación operativa anual (POA) y procesos de rendición de cuentas en instituciones de educación superior. Cuenta con experiencia en la enseñanza de áreas como redes, seguridad informática, ciberseguridad, servidores, informática, investigación y tecnologías aplicadas, integrando metodologías activas e innovadoras para fortalecer el proceso de enseñanza-aprendizaje. Ha participado en investigaciones relacionados con ciberseguridad, machine learning y detección de intrusos, así como en la publicación de artículos científicos y ponencias en

congresos académicos. Posee habilidades en liderazgo, trabajo en equipo, comunicación efectiva e innovación educativa, con formación continua en metodologías como aprendizaje basado en proyectos, aula invertida y neurodidáctica. Se caracteriza por su compromiso con la excelencia académica, la actualización profesional constante y la aplicación del conocimiento para generar soluciones tecnológicas y educativas con impacto positivo en la sociedad.

ORCID: https://orcid.org/0000-0002-9656-9917

Diana Carolina Decimavilla Alarcón

Cuenta con 14 años de trayectoria profesional en los campos de las telecomunicaciones, las redes de datos y la educación superior tecnológica, de los cuales once han sido dedicados a la docencia universitaria. Su formación académica incluye los títulos de Ingeniera en Telemática y Magíster en Telecomunicaciones por la Escuela Superior Politécnica del Litoral en Ecuador; lo que sustenta una sólida base técnica y científica. Ha fortalecido su perfil con certificaciones internacionales en computación en la nube y acreditación como formadora de formadores, además de una destacada formación continua que supera las setecientas horas en áreas como metodología científica, redacción académica, neuroeducación, Internet de las Cosas, inteligencia artificial aplicada e infraestructura tecnológica en la nube. En la actualidad

se desempeña como docente en la carrera de software de la Universidad de Guayaquil, donde participa activamente en la formación de profesionales en el ámbito del desarrollo tecnológico y las ciencias de la computación, promoviendo el pensamiento crítico, la innovación y la investigación aplicada. De manera paralela, ejerce como coordinadora de la carrera de tecnología superior en diseño y mantenimiento de redes en el Instituto Superior Tecnológico Vicente Rocafuerte, rol en el que lidera procesos de gestión curricular, aseguramiento de la calidad educativa ante los organismos de control, vinculación con la sociedad y supervisión académica. Su producción científica incluye artículos publicados en revistas arbitradas e indexadas, abordando temáticas de alto impacto como arquitecturas de microservicios para Internet de las Cosas en entornos de nube, sistemas de almacenamiento distribuido basados en tecnología blockchain, seguridad multinivel en infraestructuras tecnológicas, inteligencia artificial aplicada a redes inalámbricas y el impacto de las tecnologías de la información en el desarrollo cognitivo y rendimiento académico. Domina el idioma español como lengua materna, cuenta con nivel avanzado de inglés y conocimientos básicos de portugués, lo que facilita su participación en redes académicas y científicas de alcance internacional.

ORCID: https://orcid.org/0000-0002-0375-0216

José Ottón Pinela Tigua

Cuenta con 19 años de trayectoria en el ámbito de la tecnología de la información, la docencia en educación superior y la capacitación profesional. Es Ingeniero en Sistemas Computacionales y Magíster en Educación Informática por la Universidad de Guayaquil, lo que respalda una sólida formación técnica y pedagógica. Ha fortalecido su perfil con diversas certificaciones, entre ellas la acreditación en formación de formadores, así como con una constante actualización en redes y tecnologías, destacando su formación en redes bajo estándares internacionales impartida por la Escuela Superior Politécnica del Litoral. Actualmente se desempeña como docente en la Facultad de Ciencias Matemáticas y Físicas de la Universidad de Guayaquil y en el Instituto Superior Tecnológico Vicente Rocafuerte, donde ha consolidado una destacada trayectoria académica y de gestión. A lo largo de su carrera ha ocupado roles estratégicos como coevaluador de área, coordinador y subcoordinador de carrera, coordinador de vinculación con la sociedad y gestor principal de prácticas preprofesionales. Asimismo, ha sido miembro activo en procesos de investigación, desarrollo tecnológico e innovación, contribuyendo significativamente al fortalecimiento académico e institucional. Su experiencia integra la docencia, la gestión educativa y la aplicación de tecnologías, posicionándolo como un profesional

comprometido con la calidad, la formación integral y el desarrollo del talento humano en el ámbito tecnológico.

ORCID: https://orcid.org/0000-0003-1713-8973

Luis Arturo Caisaguano Caisaguano

Profesional con más de 25 años de trayectoria en ingeniería electrónica, automatización industrial y docencia universitaria. Es Ingeniero en Electricidad con especialización en Electrónica y Automatización Industrial, graduado de la Escuela Superior Politécnica del Litoral, y Máster en Sistemas de Energías Renovables por la Universidad Internacional de La Rioja (España). Su actividad principal se desarrolla en la Universidad de Guayaquil, donde se desempeña como Docente Investigador y Profesor de la Carrera de Ciencia de Datos, contribuyendo activamente al diseño curricular, la investigación aplicada y la formación de profesionales en tecnologías emergentes. Ha liderado procesos de creación y rediseño de carreras en la Facultad de Ciencias Matemáticas y Físicas, consolidando su rol como referente académico en el área de ingeniería y ciencias exactas. De manera complementaria, ejerce la docencia en el Instituto Superior Tecnológico Vicente Rocafuerte, donde también forma parte del departamento Administrativo Financiero y ha participado en iniciativas de diseño curricular institucional. Cuenta

con certificaciones vigentes en Aprendizaje Basado en Proyectos (ABP), Riesgos Eléctricos y actualización continua en metodología de investigación. Su producción científica incluye artículos publicados en revistas indexadas sobre sistemas fotovoltaicos y formación universitaria, áreas en las que ha coordinado proyectos de automatización industrial y generación de energía renovable para sectores residencial e industrial. Su perfil integra sólido conocimiento técnico, liderazgo académico y compromiso con una ingeniería electrónica de enfoque sostenible e inclusivo.

ORCID: https://orcid.org/0000-0003-2992-2714

Ivette Auxiliadora Mateo Washbrum

Cuenta con más de 25 años de trayectoria en el ámbito de las telecomunicaciones y las soluciones en tecnologías de la información y la comunicación, de los cuales doce han sido dedicados a la docencia universitaria. Su formación académica integra una base técnica y estratégica, con estudios en Ingeniería en Electricidad con especialización en Electrónica y Telecomunicaciones, complementados por una doble titulación de posgrado: Magíster en Administración de Empresas con mención en Negocios Internacionales por la Universidad de Guayaquil y Máster Universitario en Ingeniería de Software y Sistemas Informáticos por la Universidad Internacional de La Rioja, España. Ha

fortalecido su perfil con certificaciones en formación de formadores y en metodologías activas de enseñanza, destacando el aprendizaje basado en proyectos, además de mantener una actualización permanente en metodología de la investigación científica. Su trayectoria académica se complementa con una participación constante en congresos de tecnología y la publicación de artículos en revistas indexadas, abordando temáticas relacionadas con tecnologías de la información, innovación y transformación digital. Actualmente se desempeña como docente investigadora y profesora de la carrera de diseño y mantenimiento de redes en el Instituto Superior Tecnológico Vicente Rocafuerte, donde contribuye a la formación de profesionales altamente capacitados, promoviendo el pensamiento crítico, la innovación y la aplicación práctica del conocimiento. Su experiencia profesional abarca la gestión de redes de comunicación, el diseño e implementación de soluciones tecnológicas y la consultoría en tecnologías de la información y la comunicación, consolidando un perfil integral que combina expertise técnico, visión estratégica y liderazgo. Su enfoque actual incorpora el desarrollo y aplicación de soluciones basadas en inteligencia artificial, orientadas a la optimización de procesos, la innovación tecnológica y la transformación digital en distintos entornos organizacionales.

ORCID: https://orcid.org/0000-0002-7523-7219

Zoila Amada Pineda Calle

Cuenta con 15 años de experiencia en el desarrollo de software y 10 años en la docencia en educación superior. Su formación académica se sustenta en una sólida base en tecnologías de la información, siendo Analista de Sistemas y Licenciada en Sistemas de Información por la Escuela Superior Politécnica del Litoral. Complementa su perfil con una Maestría en Administración de Empresas con mención en Control de Calidad y Productividad por la Universidad de Guayaquil, y una Maestría en Tecnologías de la Información con mención en Transformación Digital e Innovación por la Universidad Estatal de Milagro. En el ámbito profesional ha desempeñado roles estratégicos como líder de proyectos y jefa de sistemas, participando en el diseño, desarrollo e implementación de soluciones tecnológicas orientadas a la gestión empresarial. Entre los productos desarrollados destacan sistemas de administración contable, control de inventarios y plataformas web para la gestión de producción e inventarios, evidenciando una sólida capacidad para integrar soluciones tecnológicas en entornos organizacionales. En el campo académico, ha contribuido con la generación de conocimiento a través de la publicación de artículos científicos en temas de innovación educativa y tecnologías emergentes, entre los que se destacan: la transformación del aprendizaje

mediante metodologías ágiles como Lean Startup, la optimización de la seguridad en redes inalámbricas mediante inteligencia artificial para la prevención de ataques cibernéticos, y el análisis del uso ético y responsable de herramientas de inteligencia artificial en procesos educativos. Asimismo, ha participado activamente en procesos de vinculación con la sociedad, liderando y ejecutando proyectos orientados al desarrollo comunitario, generando impacto social y fortaleciendo la relación entre la academia y su entorno.

ORCID: https://orcid.org/0009-0004-0593-714X

✨ Conocimiento que transforma sociedades
✨ Tú inspírate, nosotros publicamos

En un entorno digital caracterizado por la complejidad, la interconexión y la constante transformación tecnológica, la ciberseguridad se posiciona como un eje estratégico para la protección de la información y la continuidad de los sistemas. Este libro ofrece una visión integral que trasciende los enfoques tradicionales, articulando de manera coherente los fundamentos conceptuales, los modelos de gobernanza, las dinámicas de ataque y defensa, y los procesos de análisis forense digital que permiten comprender y enfrentar los desafíos actuales del ciberespacio. A través de un enfoque estructurado y aplicado, la obra explora cómo las amenazas evolucionan en paralelo con la tecnología, evidenciando la necesidad de adoptar estrategias adaptativas que integren aspectos técnicos, organizacionales y humanos. Asimismo, se examinan marcos normativos y metodologías que orientan la gestión del riesgo, destacando su papel en la toma de decisiones y en la construcción de entornos resilientes. De igual manera, se analizan las prácticas ofensivas y defensivas que configuran el escenario de la seguridad digital, así como la influencia del comportamiento humano en la generación de vulnerabilidades. Finalmente, se profundiza en el análisis forense como mecanismo clave para la reconstrucción de incidentes, la generación de evidencia y el fortalecimiento de las capacidades de respuesta. Esta obra constituye una herramienta académica y profesional que permite comprender la ciberseguridad como un sistema dinámico, donde la anticipación, la adaptación y el aprendizaje continuo son esenciales para proteger el valor más crítico de la era digital: la información.

SOPHIA EDITIONS

ISBN: 978-1-968794-45-3

9 781968 794453 >