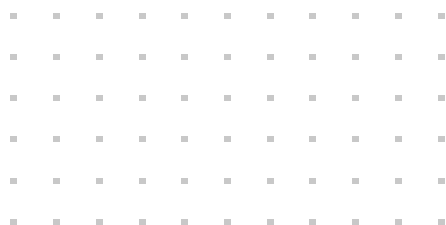


WHITE PAPER

# Uso de la inteligencia artificial para combatir las ciberamenazas

Ayude al equipo de seguridad y desarme a los atacantes



## Resumen

Hace solo unos años, todo lo que tenía que ver con la inteligencia artificial (IA) era poco más que habladurías; pero, en la actualidad, la velocidad de la innovación en IA y las consecuencias positivas y negativas que puede tener en las empresas cambian a una escala sin precedentes. En el campo de la ciberseguridad, la aparición de herramientas de IA eficaces y disruptivas y su uso por parte de los ciberdelincuentes acentúan tanto la complejidad como la necesidad apremiante de proteger las empresas y sus infraestructuras digitales de las nuevas ciberamenazas basadas en la IA. La buena noticia es que los proveedores de ciberseguridad llevan ya años aplicando diferentes tecnologías de IA. Sin embargo, teniendo en cuenta que en el futuro habrá infinidad de ciberdelincuentes que utilizarán tácticas basadas en la IA, resulta fundamental que los responsables de IT y de seguridad y sus equipos hagan evolucionar sus estrategias de seguridad para hacer frente a estas nuevas amenazas sofisticadas basadas en la IA.

## Innovar para fortalecer

Las empresas siguen poniendo en marcha iniciativas de digitalización, y esto hace que la expansión de las superficies de ataque sea inevitable. La adopción de la nube, la desaparición del «air gap» o aislamiento físico entre la tecnología de la información (IT) y la tecnología operativa (OT), la proliferación de dispositivos del Internet de las cosas (IoT) que se conectan a la red o la adaptación a las características del trabajo híbrido son solo algunos ejemplos de las iniciativas que están emprendiendo muchas empresas y que están llevando al límite los recursos de sus equipos de IT y de seguridad. El hecho de que, ahora, los ciberdelincuentes también recurran a herramientas de IA está empeorando una situación ya de por sí complicada y dinámica.



Hace poco, un grupo de ciberdelincuentes utilizó deepfakes del CFO de una empresa y de otros empleados durante una videollamada para convencer a un trabajador de que hiciera una transferencia de 25,6 millones de dólares.<sup>1</sup>

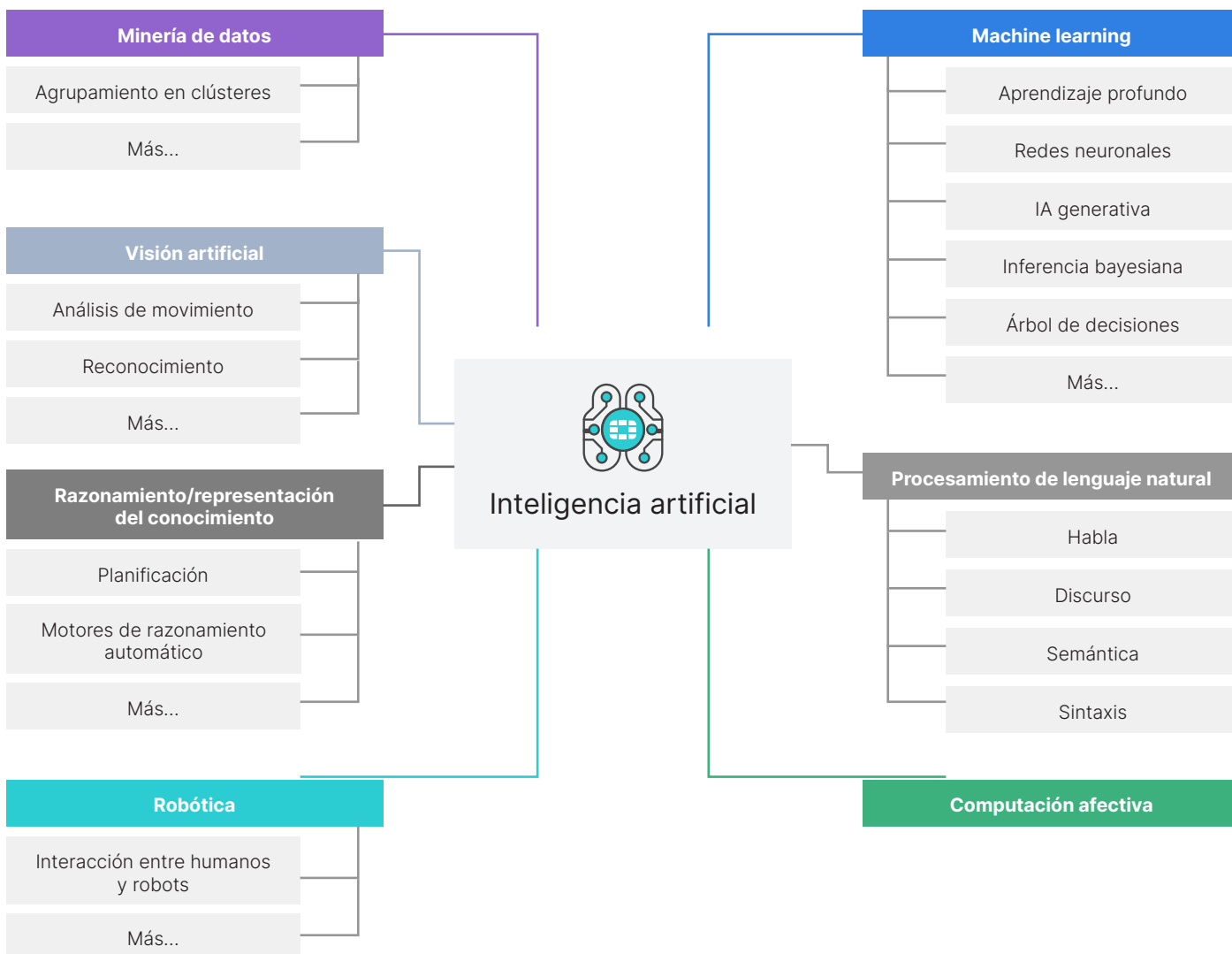


Figura 1: Subcampos o ámbitos de aplicación de la inteligencia artificial

## Cómo utilizan la IA «los malos»

Los ciberdelincuentes no han querido desaprovechar las posibilidades que les ofrece la IA para desarrollar y utilizar amenazas nuevas y más eficaces y poderosas, incluidas las de día cero. La IA permite lanzar ataques contra objetivos muy específicos más rápido que nunca, además de tener otras muchas ventajas para los atacantes:

- Las tecnologías de IA, como los transformadores generativos preentrenados (GPT) o la IA generativa (GenAI) están allanando el camino para que nuevos ciberdelincuentes se inicien en el mundillo. Gracias a esta tecnología, hoy en día cualquiera que no hable español puede redactar correos electrónicos de phishing muy verosímiles y crear ataques de ingeniería social utilizando la sintaxis que emplearía un español nativo.
- La IA puede utilizarse para escribir código malicioso nuevo y simplificar drásticamente el desarrollo de malware novedoso.
- El uso de la tecnología «deepfake» por parte de los ciberdelincuentes ya ha incendiado la clase política y al electorado más de una vez, y ha hecho que cometer ciberdelitos a gran escala sea más que viable.
- La IA también se puede usar para detectar y explotar vulnerabilidades en las aplicaciones más rápidamente, lo que aumenta el riesgo de sufrir ataques a la cadena de suministro al que se exponen las empresas en cualquier parte del mundo.
- La IA puede utilizarse para crear variantes de malware adaptativas y para lanzar ataques multivector y tipo enjambre coordinados.

Las tácticas de IA maliciosas de hoy en día abarcan el ciclo de vida completo de los ataques, representado en el marco MITRE ATT&CK. MITRE ha desarrollado una base de conocimientos denominada ATLAS (Adversarial Threat Landscape for Artificial-Intelligence Systems) que detalla las tácticas y técnicas con IA utilizadas por los adversarios.<sup>2</sup>

## Aumento de los desafíos

Los desafíos que plantean las amenazas modernas en constante evolución se ven exacerbados por el hecho de que ahora los ciberdelincuentes utilicen la IA, ya que esto pone bajo aún más presión a unos equipos de IT y de seguridad que ya están desbordados. Proteger un entorno de red y una superficie de ataque cada vez más grandes de estas amenazas nuevas es más complicado que nunca, y su dificultad radica en los siguientes factores:

- La visibilidad fragmentada de los distintos entornos.
- La falta de una aplicación de políticas centralizada y coordinada.
- El uso de numerosas herramientas y consolas de seguridad distintas, que alarga innecesariamente los procesos de supervisión, clasificación de alertas e investigación y respuesta a incidentes.
- Dificultades a la hora de contratar a expertos en seguridad y conservarlos, que no parece que vayan a desaparecer en el corto plazo.

Para tener éxito con la IA, las empresas deberán reducir la complejidad y la fricción y optimizar las operaciones.

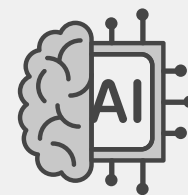
## Cómo utilizan la IA «los buenos»

La convergencia de la IA y la ciberseguridad es mucho más que un avance tecnológico. Es un paso evolutivo muy necesario y que cada vez urge más dar para ayudar a las empresas a reforzar sus mecanismos de defensa frente a las amenazas emergentes que afectan a sus superficies de ataque modernas. Muchos proveedores de ciberseguridad llevan ya años aplicando varias tecnologías de inteligencia artificial. Sin ir más lejos, Fortinet lleva más de una década investigando y utilizando tecnologías de IA y continúa adaptándose y respondiendo a los desafíos que plantea la superficie de ataque moderna con sus mecanismos de defensa basada en la IA.

## Inteligencia de amenazas basada en la IA

De los diferentes usos de la IA en el ámbito de la ciberseguridad, el más importante es el de detectar las amenazas y ofrecer protección frente a ellas, para lo cual resulta fundamental generar inteligencia de amenazas y mejorarla continuamente. El uso aplicado de las tecnologías de IA resulta fundamental para la recopilación, el análisis y la correlación de datos y, en última instancia, para la conversión de esos datos a inteligencia práctica. Este tipo de inteligencia de amenazas basada en la IA puede compartirse mediante integraciones para abordar una gran variedad de vectores de ataque y diversas clases de amenazas, estén o no basados en la IA. De ahí que la forma en la que usen la IA los proveedores, los datos que manejen y la variedad de fuentes de datos que utilicen sean factores importantes que tener en cuenta. Cuanta más visibilidad tenga un proveedor de sus datos, más podrán aprender los modelos de IA.

Entender la naturaleza de la inteligencia de amenazas en la que se basa la infraestructura de seguridad elemental de su empresa es un buen punto de partida para entender también cómo utilizan la tecnología de IA los proveedores con los que trabaja. Uno de los componentes principales es su infraestructura de firewall, ya que representa la primera línea de defensa de su empresa.



Un grupo de informáticos de la Universidad de Illinois en Urbana-Champaign probaron a utilizar ChatGPT-4 de OpenAI conjuntamente con LangChain y con Playwright como agente malicioso para analizar webs en busca de vulnerabilidades y para atacarlas sin intervención humana. Sorprendentemente, los autores del experimento aseguran que la herramienta fue capaz de ejecutar un proceso de 38 pasos asociado a un ataque UNION de inyección SQL.<sup>3</sup>



Figura 2: Funciones mejoradas con IA y nativas en la IA como parte de la formulación de la inteligencia de amenazas

Entender la naturaleza de la inteligencia de amenazas en la que se basa la infraestructura de seguridad elemental de su empresa es un buen punto de partida para entender también cómo utilizan la tecnología de IA los proveedores con los que trabaja. Uno de los componentes principales es su infraestructura de firewall, ya que representa la primera línea de defensa de su empresa.

Los firewalls de nueva generación (NGFW) actuales incluyen una combinación de funciones que van más allá de los firewalls tradicionales. Por ejemplo, su firewall podría integrar funciones de prevención de intrusiones, de protección antimalware (como antivirus y sandboxing) y de seguridad web (como el filtrado de DNS y de URL). Como el firewall integra una gran variedad de funciones y cada una de ellas es importante, pregunte al proveedor cómo aplica la IA para mejorarlas.

Si el proveedor no es capaz de explicarle cómo aplica la IA, es hora de pasar al siguiente candidato y empezar a buscar proveedores que sí utilicen las tecnologías más vanguardistas para mejorar la eficacia de sus soluciones.

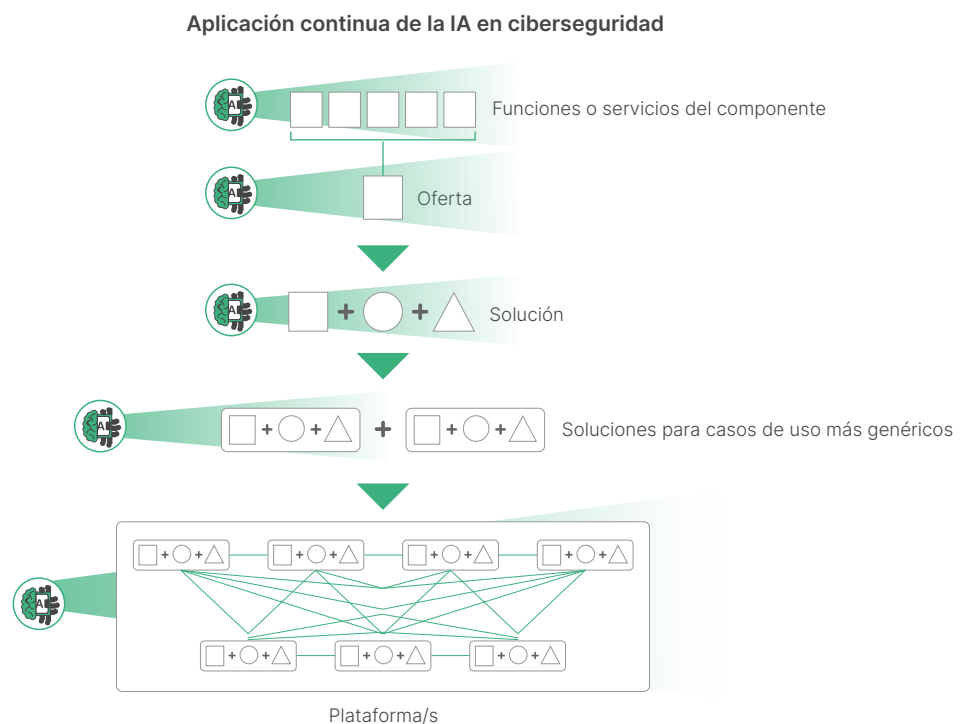


Figura 3: Aplicación de la IA y la inteligencia de amenazas desde el componente hasta la plataforma

Hoy en día, las soluciones optimizadas con IA ayudan a mejorar los resultados y reportan ventajas tanto a los proveedores como a los clientes:

- **Firewalls:** los NGFW incluyen funciones de seguridad que suelen estar basadas en varios modelos de IA ejecutados en segundo plano. Es probable que ciertas funciones como la prevención de intrusiones, los sistemas antivirus, la seguridad web y el sandboxing integrado recurran a las tecnologías de IA para mejorar las funciones individuales integradas en el firewall. Combinar firewalls de malla híbrida con NGFW beneficia a las empresas por partida doble: por un lado, disfrutan de una protección contra amenazas basada en la IA y, por el otro, obtienen mejoras en la visibilidad global del firewall y una gestión centralizada de las políticas y los firewalls.
- **Análisis de aplicaciones:** aunque los ciberdelincuentes pueden utilizar la IA para crear agentes maliciosos, las soluciones de análisis de aplicaciones y los responsables de realizar las pruebas de pentesting pueden utilizar las mismas funciones para encontrar y corregir más rápidamente las vulnerabilidades que puedan existir tanto en la fase de desarrollo como en la de producción.
- **Detección y respuesta en el endpoint (EDR):** las soluciones EDR utilizan redes neuronales para reconocer patrones y dar sentido a todos esos datos de eventos que se capturan en los endpoints, lo que incluye datos sobre las actividades que llevan a cabo, los procesos que ejecutan, las modificaciones que se realizan en los registros y los accesos a la memoria.
- **Sistema de información de seguridad y gestión de eventos (SIEM):** los SIEM utilizan modelos de machine learning (ML) supervisados y no supervisados para llevar a cabo una regresión lineal sofisticada que incluye la regresión de vectores de soporte, la regresión de procesos gaussianos y la regresión con árboles de decisión. También utilizan el ML para ejecutar distintos algoritmos de agrupamiento en clústeres. Este análisis ayuda a la solución SIEM a identificar con precisión las amenazas y vulnerabilidades y a reducir al mínimo los falsos positivos. Las soluciones SIEM también utilizan la tecnología GPT y el procesamiento de lenguaje natural (NLP) para ofrecer al personal del centro de operaciones de seguridad una experiencia más guiada y más basada en la información. Los analistas pueden consultar directamente al motor de IA y recibir información práctica sobre las amenazas, además de recomendaciones para responder a los incidentes correctamente.
- **Análisis de imágenes:** las tecnologías de visión artificial, reconocimiento de imágenes y redes neuronales se combinan para llevar a cabo el análisis de imágenes. También se pueden utilizar algoritmos del vecino más próximo. Las imágenes entrantes incrustadas en un correo electrónico o que se descargan de Internet se pueden analizar para determinar si contienen algún riesgo o exposición. Estas imágenes pueden incluir códigos QR, contenido visual pornográfico, violento o extremista o imágenes en las que aparezcan armas, alcohol o drogas.
- **Pentesting:** ChatGPT-4 de OpenAI se puede utilizar para realizar pruebas de pentesting más avanzadas, y en Internet hay videos que muestran cómo utilizar el modelo de lenguaje de gran tamaño de ChatGPT-4 para escribir scripts en Python y Bash en cuestión de minutos para después utilizarlos en pruebas de penetración.

El potencial de la inteligencia artificial no acaba con la formulación y la aplicación de la inteligencia de amenazas basada en la IA. En Fortinet, por ejemplo, utilizamos tecnologías de IA para mejorar nuestra plataforma Fortinet Security Fabric con el objetivo de hacerla aún más proactiva, unificada e inteligente.

### Ciberseguridad nativa en la IA

La aparición de soluciones de ciberseguridad que utilizan funciones de IA como punto de partida suele conocerse como «ciberseguridad nativa en la IA». Aunque en el sector no existe una definición «oficial» para el término, las herramientas de ciberseguridad nativas en la IA operan a velocidad de máquina. Por ejemplo, cuando se analiza una posible amenaza, el veredicto se da y las acciones correspondientes se toman a velocidad de máquina. Llevar a cabo estas acciones más rápido tiene una serie de ventajas muy positivas tanto desde el punto de vista de la ciberseguridad como desde el punto de vista del negocio. Por lo general, las herramientas de ciberseguridad nativas en la IA tienen en común estas características:

- Utilizan modelos de IA diseñados con fines específicos.
- Integran la IA como elemento central o como base.
- Nunca dejan de aprender y adaptarse a las amenazas nuevas.
- Realizan acciones a velocidad de máquina.
- Funcionan en tiempo real

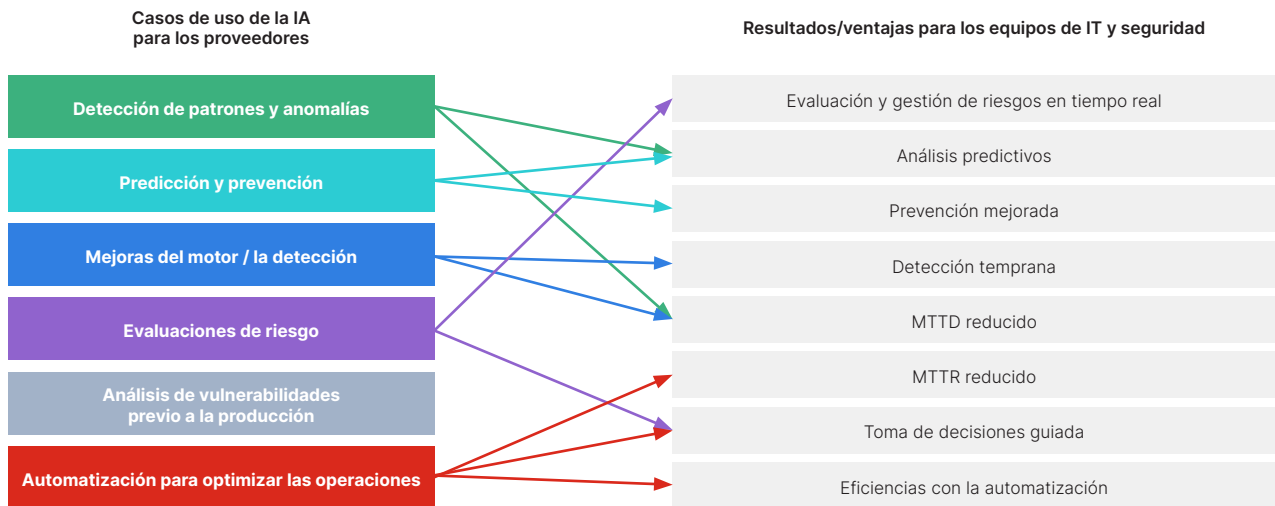


Figura 4: Factores que empujan al proveedor a usar la IA y ventajas para los clientes

### Casos de uso y recomendaciones

Es importante que los equipos de seguridad y de IT que no trabajan en el sector de la ciberseguridad sean conscientes de que los proveedores de ciberseguridad están aplicando la IA. También deben saber qué tipo concreto de IA utilizan, cómo la aplican y, lo más importante, cuáles son las ventajas directas de esa IA para la empresa.

La Figura 4 representa las formas en las que los proveedores de ciberseguridad podrían aplicar las tecnologías de IA a sus soluciones, los procesos que lo hacen posible y las ventajas que se obtienen. La lista de esta figura le ayudará a saber qué preguntar a los proveedores para averiguar qué hace su IA para mejorar la seguridad de sus clientes, sobre todo frente a las amenazas basadas en la IA.

Le recomendamos que tenga en cuenta los siguientes consejos a la hora de incorporar la IA a sus estrategias de seguridad.

### Dé prioridad a la IA

Conciencie a su equipo de la importancia de la inteligencia artificial. Determine su grado de familiarización con las tecnologías y los principios de la IA. Establezca la adopción de la IA como un objetivo estratégico en las principales áreas de su infraestructura de seguridad y de IT y priorice la evaluación de esas áreas. Como los proveedores utilizan la IA en mayor o menor medida en sus soluciones, prepare un cuestionario y establezca un proceso para evaluarlos en función de su experiencia con la IA y el grado de integración que ofrecen.

### Haga los deberes

Los responsables de IT y de seguridad deberían hacer lo necesario para informarse bien acerca de la IA y de cómo la pueden utilizar para mejorar tanto sus iniciativas como las soluciones que adquieren y, a continuación, explicárselo a sus equipos. Hacer esto ayudará también a entender mejor las cosas cuando la empresa esté experimentando activamente con las tecnologías de IA o utilizándolas para lo que se han concebido. Además, los equipos estarán mejor preparados para hacer las preguntas oportunas en relación con los distintos casos de uso. En Internet encontrará numerosos recursos gratuitos con los que empezar a familiarizarse con la IA. Una vez que los responsables y sus equipos cuenten con un nivel de conocimientos adecuado, quizás convendría plantearse adquirir recursos de formación en línea de pago o incluso asistir a cursos formativos impartidos por una empresa de ciberseguridad de renombre, como SANS.

### Nunca deje de informarse

Haga todo lo posible por estar siempre al tanto de los últimos avances de la IA en materia de ciberseguridad. El ritmo de la innovación es frenético, por lo que no es difícil quedarse desactualizado.

### Evalúe su infraestructura de seguridad

Piense en cómo puede utilizar la tecnología de IA en su infraestructura de seguridad. Empiece por analizar las ventajas que tendrían las tecnologías de IA para la principal línea de defensa de su empresa y, a continuación, haga lo propio con otros controles que pueda tener. En muchos sentidos, el avance de la evaluación puede seguir el orden de la priorización de riesgos que haya establecido para su entorno en general (de modo que se evalúen primero los aspectos de mayor riesgo).

## Haga las preguntas necesarias

Pregunte a los proveedores de ciberseguridad con los que trabaja cómo utilizan la IA para entender las tecnologías que aplican, cómo las aplican y, lo más importante, qué ventajas reporta la IA a los clientes. Puede empezar con preguntas como estas:

- ¿Qué nivel de visibilidad tienen de las amenazas y qué fuentes de datos utilizan para formular la inteligencia de amenazas en la que se basa su producto, su servicio o su solución?
- ¿Qué papel desempeña la IA en la formulación de esa inteligencia de amenazas?
- ¿Qué experiencia tiene su empresa con el uso de tecnologías de IA en sus productos, servicios y soluciones?
- ¿Qué tecnología o tecnologías de IA específicas se utilizan en este producto, servicio o solución en concreto y cómo se aplican?
- ¿Qué fuentes de datos utiliza el producto, el servicio o la solución para alimentar la tecnología de IA?
- ¿Cómo entrenan y reentrenan sus modelos de IA?
- ¿Nosotros podremos interactuar con la IA directamente de algún modo?
- ¿Qué hacen para evitar el envenenamiento de datos por parte de los ciberdelincuentes?
- ¿Podría explicarme cómo aplican la IA para...?
  - Reducir los riesgos
  - Mejorar la prevención
  - Reducir el tiempo medio de detección
  - Reducir los falsos positivos
  - Contribuir a la clasificación de alertas y la investigación de incidentes
  - Reducir el tiempo medio de corrección
  - Ayudar a los analistas de operaciones de seguridad en su trabajo diario

Esta no pretende ser una lista exhaustiva de preguntas sobre la IA, sino más bien servir de guía. Modifique estas preguntas o añada las que considere oportuno en función de las necesidades particulares de su empresa.

## Incluya la IA entre los criterios de evaluación

Asegúrese de incluir el uso de la IA en sus documentos de solicitud de propuestas. Utilice las preguntas anteriores, y otras que considere oportunas, para entender cómo utilizan la IA los distintos proveedores de ciberseguridad. Así, no solo tendrá claro cómo utiliza las tecnologías de IA el proveedor en cuestión para ayudar a sus clientes, sino que también podrá comparar respuestas para determinar si el uso que hace de la IA cierto proveedor se traducirá en ventajas reales para su empresa.

## Conclusión

Los responsables y profesionales de IT y de seguridad deben familiarizarse mejor con las distintas tecnologías categorizadas con la etiqueta de «IA». Es importante mostrarse proactivo a la hora de aprender sobre IA y abrir la puerta a todas las posibilidades que ofrece en materia de ciberseguridad. Muchas de las soluciones de seguridad disponibles son de proveedores que ya utilizan la IA, por lo que debería hacer todo lo posible por entender sus distintas aplicaciones y casos de uso antes de incorporar herramientas con IA de forma generalizada en su empresa.

Antes de decidirse por una solución u otra, asegúrese de hacer a los proveedores preguntas específicas sobre cómo utilizan la IA en sus productos. Saber cómo incorporan la IA en las operaciones de IT y de seguridad le ayudará a hacer frente a las sofisticadas amenazas basadas en la IA de hoy en día.

<sup>1</sup> Heather Chen y Kathleen Magramo, «[Finance worker pays out \\$25 million after video call with deepfake 'chief financial officer'](#)» (en inglés), CNN, 4 de febrero de 2024.

<sup>2</sup> [MITRE ATLAS](#) Adversarial Threat Landscape for Artificial-Intelligence Systems.

<sup>3</sup> Richard Fang, et al., «[LLM Agents can Autonomously Hack Websites](#)» (en inglés), 6 de febrero de 2024.