El doble filo de la IA en el mercado de la ciberseguridad

El uso de la inteligencia artificial, de manera masiva, en la ciberseguridad ha abierto un nuevo capítulo en este complejo mercado. En la misma ecuación, el ecosistema tecnológico debe saber despejar la oportunidad que supone su uso frente al riesgo de las amenazas alimentadas con ella. Los mayores desafíos que tiene que enfrentar no son solo tecnológicos: es esencial incorporarla, pero hay que gobernarla con criterio, transparencia, ética y de acuerdo a criterios normativos. Fabricantes como Check Point Software, ESET, Fortinet, Microsoft, Sophos, SonicWall, Trend Micro y WatchGuard analizan este complejo panorama.

Marilés de Pedro

Amenazas *versus* defensas

El uso de la inteligencia artificial está permitiendo a las empresas pasar de un modelo reactivo a uno claramente predictivo. "La inteligencia artificial está convirtiendo la ciberseguridad en un proceso mucho más proactivo, capaz de detectar anomalías y ataques complejos en tiempo casi real analizando volúmenes masivos de datos que los equipos humanos no son capaces de revisar en un periodo de tiempo aceptable para proporcionar una respuesta eficaz", explica Josep Albors, director de investigación y concienciación de ESET España. Un salto que no es solo analítico, también operativo. "La IA permite automatizar una gran parte de la respuesta a incidentes, lo que reduce de horas a minutos el tiempo necesario para contener una amenaza y libera a los analistas para tareas de mayor importancia". Sin embargo, esta ventaja no es exclusiva del lado defensivo. "También está potenciando a los atacantes, incluidos aquellos con pocos conocimientos técnicos, con *malware* más evasivo,

> "La inteligencia artificial está convirtiendo la ciberseguridad en un proceso mucho más proactivo" **(Josep Albors)**

campañas automatizadas y *deepfakes*, generando una nueva carrera armamentística en la que la clave es cómo gobernar y explotar la IA mejor y más rápido que el adversario".

Una visión "dual" compartida por Eusebio Nieva, director técnico de Check Point Software para España y Portugal, que reconoce que es un arma de doble filo: "La IA representa una ventaja defensiva significativa cuando se implementa de manera estratégica. Maximiza la defensa si se utiliza correctamente, pero también puede potenciar la ofensiva de los ciberdelincuentes si cae en manos equivocadas".

Rafe Pilling, director de Threat Intelligence en Sophos X-Ops, ahonda en esta dicotomía pero resalta la evolución, destacada, de las capacidades defensivas. "La IA agéntica representa un salto cualitativo al observar entornos en tiempo real, aprender continuamente de comportamientos normales y anómalos y tomar decisiones autónomas para responder a amenazas críticas, pasando del análisis a la acción inmediata". En el lado de los "malos", aunque asegura que su uso aún no ha generado grandes avances disruptivos, sí está produciendo

mejoras que están sentando las bases de ataques más creíbles y escalables. "Además del uso malicioso de la IA, existe otra superficie de ataque potencial que se introduce cuando las empresas integran la IA en aplicaciones y conjuntos de tecnologías existentes. Y ésta también debe protegerse", alerta.

IA "real" *versus* automatización avanzada

En un mercado donde muchas soluciones se presentan como "inteligentes", distinguir entre automatización avanzada e IA auténtica es clave. "La diferencia clave radica en la capacidad de aprendizaje y de respuesta autónoma: la IA se adapta, mientras que la automatización sigue instrucciones", señala Eusebio Nieva.

También en quién reconoce los patrones que diagnostican el contenido malicioso: en las soluciones tradicionales están definidos por los analistas, programadores e ingenieros de seguridad mientras que en las soluciones basadas en IA los algoritmos crean sus propios modelos. "Aprenden y evolucionan a partir de los datos que analizan de forma continua, lo que permite contar con una capacidad predictiva y adaptativa que anticipa los ataques que se alejan de los patrones conocidos, algo que la automatización, basada en reglas definidas por expertos, no puede hacer", puntualiza Nieva.

> "La IA representa una ventaja defensiva significativa cuando se implementa de manera estratégica"
> (Eusebio Nieva)

Para distinguir las soluciones basadas en IA, Rafe Pilling apela a varios aspectos técnicos. El primero, un aprendizaje continuo, ajustando sus modelos en función de nuevos datos. "La IA agéntica puede formular hipótesis de ataque, establecer correlaciones complejas entre eventos aparentemente desconectados y ejecutar medidas de remediación sin intervención humana, siempre bajo una gobernanza definida". Una capacidad que ante la escasez de talento (se calcula que faltan cuatro millones de expertos en ciberseguridad en todo el mundo) es esencial.

Iratxe Vázquez, *senior product marketing manager* de ciberseguridad de WatchGuard Technologies, recuerda que la IA y el *machine learning* llevan décadas utilizándose para automatizar la detección y respuesta frente a amenazas; pero que ahora actúan como un multiplicador de la eficacia. "Ayudan a convertir señales dispersas en decisiones accionables y a priorizar lo que realmente importa", explica. Ahora bien, insiste en que, para automatizar con confianza, "hay que basarse en hechos comprobables y funcionar con límites claros y supervisión humana. El objetivo no es tener más IA, sino una mejor seguridad con menos ruido, acelerando la detección y la respuesta sin perder control".

"El uso de IA generativa puede llegar a triplicar la efectividad del phishing, permitiendo ataques más personalizados y difíciles de detectar" **(Elena García)**

¿Dónde marca la diferencia?

Si hay un ámbito donde la inteligencia artificial está demostrando un impacto tangible es en la detección avanzada de amenazas y en la capacidad de respuesta operativa. La combinación de análisis de comportamiento, correlación de eventos y automatización está redefiniendo la forma en la que los equipos de seguridad identifican y gestionan incidentes.

"Su mayor valor se encuentra en la detección de amenazas desconocidas o sin firma", explica Sergio Martínez, director general de Sonicwall en Iberia. "Permite identificar patrones anómalos que pasarían desapercibidos si se aplicara un enfoque tradicional, reduciendo significativamente el tiempo de detección". Además, se optimiza la gestión

posterior del incidente "ayudando a los equipos de seguridad a centrarse en las amenazas realmente críticas".

Josep Albors insiste en su valor en estos procesos de detección, tanto estática como dinámica, y en la clasificación de todo tipo de amenazas, permitiendo su identificación en tiempo real conforme se detectan comportamientos anómalos en los sistemas. "Ofrece una ayuda muy relevante a los equipos de respuesta ante incidentes, utilizando la automatización para la orquestación de tareas, reduciendo los tiempos de respuesta y la carga de trabajo".

La ventaja comparativa de la IA frente a los métodos tradicionales se hace especialmente evidente en la detección de amenazas Zero Day. "Al utilizar análisis basados en el comportamiento y en la búsqueda de anomalías, puede analizar patrones de tráfico, procesos y usuarios de forma rápida para detectar acciones sospechosas en tiempo real", explica Albors. Este cambio de paradigma permite identificar ataques que no habían sido vistos antes, "con tasas de éxito elevadas y con un número de falsos positivos reducido".

En definitiva, la IA no solo acelera la detección, sino que amplía el espectro de amenazas detectables. Frente a un modelo estático basado

en lo ya conocido, introduce una defensa dinámica capaz de adaptarse a comportamientos nuevos y potencialmente maliciosos.

Otra de las mejoras es la reducción radical de los tiempos de detección y contención, pasando de horas o días a minutos o, incluso, segundos. "En un entorno donde los atacantes pueden comprometer sistemas y exfiltrar datos con gran rapidez, esa diferencia temporal es decisiva", apunta Eusebio Nieva.

También en la ejecución de los servicios MDR. "La IA agéntica automatiza la clasificación inicial de alertas, construye correlaciones entre eventos, formula hipótesis de ataque y ejecuta las primeras medidas de remediación", relata Rafe Pilling. Esta colaboración humano-máquina "permite a los analistas concentrarse en incidentes complejos mientras se reduce el tiempo total de respuesta".

La IA como motor del cibercrimen

Junto al uso que hacen de ella los fabricantes, los maleantes virtuales también la están utilizando. El uso de IA generativa, por ejemplo, puede llegar a triplicar la efectividad del *phishing*, permitiendo ataques más personalizados y difíciles de detectar. En el primer semestre de 2025, según los datos de Microsoft, España fue el decimocuarto país a nivel mundial cuyos clientes se vieron más afectados por la actividad cibernética maliciosa. En Europa concentró, aproximadamente, el 5,4 % de los clientes afectados por estas actividades, lo que colocó a España en quinto lugar. "Los patrones de ataque han cambiado de forma significativa", alerta Elena García, *chief security advisor* en Microsoft. Ahora los atacantes rara vez irrumpen directamente en los sistemas: en la mayoría de los casos acceden mediante credenciales legítimas. Más del 97 % de los ataques contra identidades se apoyan en contraseñas y la multinacional estima que hasta el 99 % de estos intentos podría bloquearse mediante el uso de autenticación multifactor resistente al *phishing*. "La aplicación de la inteligencia artificial nos impulsa a revisar nuestras estrategias hacia modelos Zero Trust", recomienda la directiva.

> "El mayor valor de la IA se encuentra en la detección de amenazas desconocidas o sin firma" **(Sergio Martínez)**

Sergio Martínez añade que, además de para automatizar y escalar ataques, especialmente en campañas de *phishing* y suplantación de identidad, los malos la emplean "para optimizar *malware*, encadenar ataques y ajustar dinámicamente su comportamiento para evitar ser detectados".

También es especialmente preocupante el m*alware* polimórfico y la orquestación automática de ataques complejos. "Los ataques de *phishing* generados por IA pueden adap-

tarse al estilo de comunicación de cada víctima, aumentando la probabilidad del éxito. La IA permite la creación de *malware* capaz de evadir detección mediante aprendizaje continuo de los mecanismos de defensa", alerta Eusebio Nieva, que suma el abuso de modelos generativos para la desinformación, el fraude financiero y la manipulación de datos empresariales.

Según las predicciones de Sophos hay tres categorías de amenazas que generan especial preocupación. En primer lugar, el fraude mediante *deepfakes* de voz y vídeo alcanzará una escala empresarial, permitiendo eludir controles de identidad en procesos críticos como las aprobaciones financieras, el restablecimiento de contraseñas o la incorporación de proveedores. "La evolución del fraude del CEO marcará un nuevo punto de inflexión, con la combinación de IA generativa y modelos agénticos creando campañas altamente personalizadas", señala Rafe Pilling. También potenciará los ataques de *ransomware* como

> "La integración de la IA debe ser segura, ética y operativa"
> **(José María Vigueras)**

servicio (RaaS) al automatizar la explotación de vulnerabilidades, el *phishing* adaptativo, la selección de víctimas e incluso la posible negociación autónoma de extorsiones. "Esta automatización criminal ha democratizado el acceso a capacidades de ataque sofisticadas con mínimas habilidades técnicas". Grupos como Scattered Spider están intensificando campañas que comienzan con el robo de credenciales y de identidad, un vector particularmente preocupante considerando que el 71 % de los puntos de entrada iniciales en los ciberataques se producen a través de dispositivos periféricos como VPN y *firewalls*. Por último, Pilling alerta de los ataques perpetrados por la inyección de *prompts* contra aplicaciones de IA corporativas. "El riesgo se verá amplificado por los errores cometidos por trabajadores que utilizan herramientas de IA sin el control adecuado, exponiendo información sensible a través de integraciones no controladas, fugas de *prompts* o conectores mal configurados".

En relación a los ataques totalmente automatizados y adaptativos, más fáciles de diseñar gracias a la IA, Iratxe Vázquez asegura que se está mejor preparado pero no de manera completa y que se verán ataques *end to end*

orquestados por IA autónoma que automatizan desde el reconocimiento hasta la exfiltración y la extorsión. Frente a este escenario, "la detección reactiva por sí sola no es suficiente", alerta. No se trata de si la IA atacará o defenderá mejor, sino de contar con plataformas unificadas de ciberseguridad que integren, de manera nativa la IA, para "combinar la prevención, la correlación, la automatización y los servicios MDR y mantenerse por delante del atacante".

Una idea en la que insiste Elena García. "Las organizaciones mejor preparadas son aquellas que han incorporado mecanismos defensivos igualmente automatizados, apoyados en inteligencia artificial, y que sitúan la protección continua de la identidad en el centro de su estrategia de seguridad".

El riesgo de la "confianza ciega"

A pesar de las fortalezas tecnológicas que ofrece la IA en el diseño de una buena postura de ciberseguridad, los fabricantes coinciden en la necesidad de practicar la prudencia. "El principal riesgo es asumir que la IA es infalible", alerta Raúl Guillén, *cybersecurity strategy evangelist* en Trend Micro. Si se utiliza de forma aislada puede generar una falsa sensación de seguridad. "Los modelos pueden ser engañados, sufrir sesgos o degradarse si no se entrenan y supervisan correctamente", continúa. La IA potencia a los equipos humanos "pero no sustituye la visibilidad, el contexto ni la toma de decisiones estratégicas".

> "El riesgo se verá amplificado por los errores cometidos por trabajadores que utilizan herramientas de IA sin el control adecuado" **(Rafe Pilling)**

Desde Fortinet refuerzan esa idea. "La inteligencia artificial es una poderosa aliada, pero debe complementarse con supervisión humana, buenas prácticas y una estrategia sólida. La confianza ciega puede generar una falsa sensación de seguridad frente a ataques novedosos o manipulaciones de los propios modelos", señala José María Vigueras, *specialist systems engineer public cloud* del fabricante.

Regulación y privacidad: innovación con responsabilidad

Para un uso correcto de la IA es necesario definir una regulación sólida y garantizar la privacidad de los datos que procesa. Una normativa equilibrada ayuda a generar confianza, fomenta la adopción segura de la tecnología y establece responsabilidades claras sobre cómo se entrenan y aplican los modelos de IA. Pero, esta obligatoriedad de regular su uso, ¿supone un freno a la innovación? José María Vigueras cree que el equilibrio es posible. "Una regulación excesiva podría ra-

> "El principal riesgo es asumir que la IA es infalible"
> **(Raúl Guillén)**

lentizar la innovación, pero una regulación equilibrada, basada en la transparencia y la responsabilidad, puede generar confianza y acelerar la adopción segura de la IA".

Elena García también cree que no existe ninguna contradicción entre la innovación y la regulación responsable. "Son procesos complementarios que deben avanzar de forma coordinada. Una regulación clara contribuye a generar la confianza necesaria para una transformación digital sostenible, sin comprometer la innovación".

Al mismo tiempo, proteger la privacidad de los datos es clave para evitar fugas, sesgos o usos indebidos de información sensible. Diseñar la IA con principios de minimización de datos, trazabilidad y soberanía digital no solo permite cumplir con las normativas, sino que asegura que la tecnología sea transparente y confiable. "El equilibrio pasa por un diseño responsable desde el origen", explica Raúl Guillén. "Apostamos por arquitecturas que procesan datos de forma segura, con control sobre dónde se almacenan y cómo se utilizan, alineadas con regulaciones como GDPR o la futura AI Act".

Además de corroborar la estrategia de integrar, desde el diseño, la IA y la privacidad, Elena García suma la incorporación de capacidades avanzadas de seguridad, como el cifrado, la clasificación de la información, la prevención de la fuga de datos y los mecanismos de control de accesos, junto con compromisos explícitos en materia de soberanía del dato, como garantes de la privacidad.

Grandes desafíos

Pintado el panorama, los fabricantes se enfrentan a desafíos importantes en los que se

combinan la innovación tecnológica, la gobernanza ética y la protección operacional.

Para Iratxe Vázquez el desafío principal consiste en gestionar simultáneamente los riesgos "ofensivos" y "defensivos". "Hay que mitigar ambos frentes simultáneamente, usando la IA para reducir el ruido, aumentar la precisión y priorizar una prevención proactiva frente a una detección reactiva, combinando IA local y en *cloud*, y aplicando *human in the loop* en entornos SOC y MDR".

Raúl Guillén insiste en la complejidad de equilibrar ambos frentes. "Debemos proteger un entorno cada vez más complejo, donde la IA es utilizada tanto por defensores como por atacantes. Los fabricantes tienen que innovar al mismo ritmo que evoluciona la amenaza".

Una "batalla" en la que el mayor desafío, para Josep Albors, es mantener una ventaja defensiva frente a los atacantes que generan *malware* altamente evasivo y adaptable. "Esto exige innovación continua en sistemas QQ, cumplimiento regulatorio efectivo y soluciones integrales".

Sergio Martínez apunta a la confianza y coherencia operacional como elementos críticos. "La IA debe integrarse de forma coherente en plataformas unificadas sin añadir complejidad operativa. El éxito depende de ofrecer protección inteligente, adaptable y fácil de gestionar, capaz de evolucionar al mismo ritmo que los ataques impulsados por IA".

El reto no es solo técnico también ético y regulatorio. "La adopción de IA plantea desafíos éticos y regulatorios: transparencia en los algoritmos, explicación de decisiones automatizadas y cumplimiento de normativas de protección de datos. El equilibrio entre innovación tecnológica y responsabilidad es clave", señala Eusebio Nieva.

> "Hay que contar con plataformas unificadas de ciberseguridad que integren de manera nativa la IA"
> (Iratxe Vázquez)

Apela también a la ética José María Vigueras. "La integración de la IA debe ser segura, ética y operativa: el reto es proteger tanto los sistemas tradicionales como los nuevos entornos impulsados por IA, asegurando que la tecnología sea comprensible, operable y realmente útil para los equipos de seguridad. La combinación de IA, automatización y experiencia humana es la base".

En Microsoft recuerdan que la IA también debe protegerse de ataques específicos, lo que comienza por garantizar la observabilidad y el gobierno. "Hay que garantizar que los sistemas sean seguros, confiables y estén adecuadamente gobernados y cuenten con una protección adecuada", señala Elena García.

Rafe Pilling insiste en que los retos van más allá de la tecnología, alcanzando conceptos vinculados con la gobernanza, la transparencia y el liderazgo estratégico. "La adopción de IA autónoma no puede hacerse sin una reflexión ética rigurosa ya que, incluso bien entrenada, no es infalible y puede generar falsos positivos, reaccionar excesivamente o estar expuesta a datos sesgados", alerta.

No olvida referirse a la vinculación entre los servicios gestionados y las herramientas im-

pulsadas por IA que obliga a los proveedores a demostrar dónde interviene el juicio humano y quién asume la responsabilidad durante un incidente. "Los fabricantes deberán implementar medidas de gobernanza que incluyan la compartimentalización de modelos (diseñando agentes especializados con perímetros de intervención bien definidos), trazabilidad completa para que cada decisión pueda ser explicada y auditada, y protección estricta de la privacidad mediante separación de entornos entre funciones públicas y sistemas críticos".

Unos servicios que permiten a los fabricantes democratizar el acceso a la ciberseguridad. "Muchas de las disrupciones más graves no resultarán de técnicas sofisticadas, sino de fallos básicos de higiene de seguridad completamente prevenibles. La realidad es que tener un CISO en una empresa es un lujo actualmente, lo que subraya la magnitud de la escasez de talento especializado", concluye el directivo de Sophos.

Escasez de profesionales

Uno de los grandes retos que tiene que afrontar el sector es la "creación" y formación del talento. En 2025 había aproximadamente 5,5 millones de profesionales de ciberseguridad trabajando en todo el mundo pero aún existía una brecha global de alrededor de 4 a 4,8 millones de profesionales sin cubrir. Esto significa que el sector necesita casi duplicar el tamaño de la fuerza laboral actual para satisfacer la demanda real de talento especializado. En Europa el déficit se calcula en 347.000 profesionales de ciberseguridad. Incluso algunas fuentes apuntan que si se considera toda la demanda (incluyendo necesidades futuras y perfiles especializados adicionales), el número de profesionales necesarios podría acercarse o superar el medio millón. En España organizaciones como la Fundación CEOE ha señalado que se necesitarían más de 40.000 especialistas centrados en ciberseguridad para poder hacer frente a la creciente amenaza de ataques.

Ante este panorama, ¿la IA podría reducir la necesidad de contar con expertos en ciberseguridad? Sergio Martínez cree que lo que hará será transformar su rol. "El valor del experto humano será cada vez mayor en la interpretación del contexto, la validación de decisiones automatizadas y la definición de políticas de seguridad, trabajando en conjunto con la IA como un multiplicador de capacidades, no como un sustituto".

De la misma opinión es José María Vigueras. "La IA automatizará tareas repetitivas y permitirá que los profesionales se centren en la toma de decisiones, el análisis avanzado y la respuesta estratégica. La experiencia humana seguirá siendo clave".