Deepfakes, ataques cuánticos y malware autónomo: el nuevo rostro de la ciberamenaza

El avance imparable de la inteligencia artificial está transformando el panorama de la ciberseguridad a una velocidad sin precedentes. Lejos de ser solo una aliada, la IA se convierte ahora en un arma de doble filo: alimenta nuevas generaciones de ataques mientras se perfila como herramienta clave para defenderse de ellos. Empresas, expertos y gobiernos enfrentan los retos de una era en la que la inteligencia artificial no solo predice el futuro de la ciberseguridad, sino que lo ejecuta

Mónica Ramírez

En un mundo cada vez más digitalizado, donde los datos fluyen sin fronteras y la infraestructura crítica depende de sistemas interconectados, la ciberseguridad ha dejado de ser un ámbito exclusivamente técnico para convertirse en una prioridad estratégica. Y en el epicentro de esta transformación se encuentra la inteligencia artificial (IA), protagonista indiscutible de un cambio de paradigma que afectará a gobiernos, empresas y ciudadanos por igual.

La firma Check Point Software ha sido clara en su diagnóstico: la convergencia entre IA autónoma, arquitecturas web avanzadas, amenazas cuánticas y entornos hiperconectados redefinirá la resiliencia de las organizaciones en 2026.

IA: aliada y amenaza al mismo tiempo

"La prioridad hoy para las empresas no es solo usar la inteligencia artificial, sino también defenderse de ella", sentenció Mario García, director general de Check Point para España y Portugal, durante un reciente encuentro con medios especializados. Según explicó, los atacantes ya se están aprovechando de modelos generativos para lanzar ofensivas más rápidas, más creativas y difíciles de rastrear. Sin embargo, recordó que los defensores también llevan años aplicando IA para anticiparse a los riesgos.

Uno de los casos que ilustra esta nueva era es el considerado "ataque del año", documentado por Anthropic (empresa líder en IA): un ciberataque totalmente automatizado por inteligencia artificial, capaz de ejecutar tareas de reconocimiento, explotación y movimientos laterales con autonomía completa. Este tipo de agresiones marcan un antes y un después en la forma en que entendemos la defensa digital.

El auge de los agentes autónomos

Una de las predicciones más disruptivas del informe de Check Point es la llegada de la "IA agentiva": sistemas autónomos que no solo asisten, sino que toman decisiones por sí mismos. En este nuevo entorno, agentes de IA podrán gestionar presupuestos, optimizar cadenas logísticas y tomar decisiones operativas sin intervención humana directa.

El principal riesgo en este supuesto es que se produzca una autonomía descontrolada que exija auditorías continuas y mecanismos de trazabilidad completos.

Web 4.0: nueva frontera de seguridad

A esto se suma la irrupción de la llamada Web 4.0, un ecosistema digital basado en computación espacial, realidad extendida y gemelos digitales. Las organizaciones podrán simular plantas industriales, ciudades o campus en tiempo real, pero también deberán afrontar nuevos desafíos de interoperabilidad y seguridad entre los mundos físico y virtual.

Identidad en crisis: deepfakes y fraude conversacional

Otro de los puntos críticos será la suplantación de identidad mediante voz, vídeo o texto generados por IA. Según Check Point, en 2026 será técnicamente posible falsificar una videollamada o una orden por voz con tal realismo que podría engañar incluso a expertos en seguridad. Los casos no son hipotéticos: una empresa multinacional ya sufrió un fraude por 25 millones de dólares debido a una videollamada "deepfake" que suplantó a su CFO (director financiero).

Las empresas deberán incorporar autenticación conductual, validación de contexto y análisis de patrones para verificar interacciones que, a simple vista, parecen auténticas.

Nuevos vectores de ataque: modelos como objetivo

El crecimiento de los grandes modelos de lenguaje (LLM) ha abierto un nuevo flanco: los propios modelos pueden ser atacados. A través de la inyección de "prompts" maliciosos o el envenenamiento de datos de entrenamiento, los ciberdelincuentes pueden modificar el comportamiento de una IA para obtener acceso privilegiado o sabotear decisiones críticas. En este escenario, el ciclo de vida del modelo (datos, entrenamiento, validación y despliegue) deberá estar gobernado con el mismo rigor que cualquier activo de misión crítica.

La carrera cuántica y el riesgo latente

Aunque los ordenadores cuánticos aún no han alcanzado su pleno potencial, los atacantes ya operan bajo la estrategia "harvest now, decrypt later": capturan datos cifrados hoy para descifrarlos mañana. Este ries-

Foto: Shutterstock.

go ha impulsado a organismos como el Instituto Nacional de Estándares y Tecnología de EE. UU (National Institute of Standards and Technology – NIST) a promover estándares de criptografía poscuántica, que ya comienzan a ser adoptados en organizaciones estratégicas de Europa y América.

España, por su parte, ha creado el Centro Nacional de Supercomputación Cuántica en Barcelona, reforzando su apuesta por estar en la primera línea de esta revolución tecnológica.

¿Están las empresas preparadas?

No del todo. Así lo revela un informe de Boston Consulting Group (BCG), según el cual solo el 7% de las empresas ha implementado herramientas de defensa impulsadas por IA, pese a que el 60% reconoce haber sufrido ataques con componentes de IA en el último año. El 88% planea adoptarlas próximamente, pero la brecha entre amenaza y reacción sigue siendo preocupante.

Entre los obstáculos figuran la falta de talento especializado, presupuestos limitados y marcos regulatorios aún inmaduros. En América Latina, por ejemplo, el desfase entre velocidad de ataque y capacidad de respuesta es especialmente crítico.

En el ámbito de las empresas del sector industrial, según publicaba el Barómetro Industrial del Consejo General de la Ingeniería Técnica Industrial de España (COGITI), el 75% de los ingenieros que trabajan en ellas considera que es un tema importante a la hora de hacer frente a las amenazas de los avances tecnológicos, y casi la mitad de los encuestados (46,35%) sitúa el nivel de integración actual de la misma en este sector en una opción intermedia.

En cuanto a la inteligencia artificial, se percibe un interés creciente, aunque todavía muchas empresas no han iniciado su preparación para integrarla. El 46% afirma no haber dado pasos concretos en este sentido, si bien se reconoce de forma mayoritaria su potencial para transformar el modelo industrial, mejorar la eficiencia operativa y abrir nuevas oportunidades de negocio; frente al 24% de las empresas que aseguran estar preparadas en gran medida, según el citado Barómetro Industrial.

Las amenazas que dominarán 2026

La encuesta global de BCG, por su parte, indica que los riesgos que más inquietan a los ejecutivos son el fraude financiero potenciado por IA (43%), la Ingeniería social adaptativa (39%), el Malware que aprende y evade defensas (26%) y el "Prompt injection" y manipulación de modelos (22%).

La lista deja claro que la ciberseguridad ya no se limita a firewalls o antivirus: ahora exige inteligencia adaptativa, monitorización continua y marcos de gobernanza más estrictos.

Sectores más afectados y ataques en evolución

Entre los sectores más golpeados por ciberataques con IA se encuentran la salud, con ataques que han paralizado sistemas hospitalarios y retrasado cirugías; los servicios financieros, blanco preferido del fraude conversacional y videollamadas deepfake; los gobiernos de los países, donde la ingeniería social con IA busca comprometer cuentas privilegiadas, y al ámbito de la manufactura, donde los agentes maliciosos interrumpen líneas de producción con malware adaptativo.

Por ubicación, los países más atacados han sido Estados Unidos y Corea del Sur, seguidos por economías europeas y mercados emergentes latinoamericanos.

Cuantificación del impacto económico

Con relación al impacto económico de los ciberataques con IA, se estima que el coste global de los ciberataques alcanzó los 10,5 billones de dólares en 2025, con un crecimiento anual del 15%. Por su parte, los ataques de ransomware sin cifrado (solo extorsión por filtración) aumentaron un 32% en 2023, tendencia que seguirá en 2026, y el uso de IA en ataques acelera la velocidad de penetración en redes un 300%, según datos de Veeam.

IA también para defender

Aunque los atacantes han abrazado la IA con rapidez, los defensores también cuentan con las siguientes herramientas:

- Motores de "machine learning" capaces de detectar zero-days con precisión superior al 90%. Zero-days o día cero se refiere a una vulnerabilidad de seguridad en software o hardware que es desconocida para el fabricante y los usuarios, lo que significa que no existe una solución o parche.
- Plataformas SOAR que automatizan la respuesta a incidentes y reducen el tiempo medio de contención.
- Algoritmos predictivos, como los usados por bancos españoles, que detectan movimientos sospechosos en milisegundos.
- Sistemas de análisis conductual que monitorean patrones para detectar accesos anómalos, incluso si las credenciales son válidas.

Gobernanza y responsabilidad

Sin embargo, el futuro de la ciberseguridad no puede apoyarse solo en la tecnología. Los expertos proponen la creación de Consejos de Gobernanza de IA, capaces de supervisar la adopción ética, segura y controlada de sistemas autónomos. También sugieren fortalecer la coordinación entre CEO y CISO (Chief Information Security Officer) para garantizar que la ciberseguridad se eleve al máximo nivel de decisión empresarial.

Además, un 72% de las organizaciones apoya ya la prohibición de pagos por "ransomware" (secuestro de archivos), conscientes de que perpetúan un ciclo de impunidad y financiación criminal.

El nuevo estándar: resiliencia, cumplimiento y visibilidad

Los entornos multicloud, las plataformas SaaS y la proliferación de dispositivos IoT han reducido la visibilidad sobre la ubicación de los datos. Según Veeam, solo el 29% de los responsables de TI se sienten muy seguros de poder recuperar información crítica tras un ciberataque. Esta falta de confianza evidencia que la resiliencia debe ir más allá del plano técnico: debe abarcar cumplimiento normativo, gobernanza y control de terceros.

Los estándares emergentes incluyen políticas legibles por máquina, análisis automatizados de riesgo y monitoreo continuo de la cadena de suministro.

Hacia una defensa automatizada e inteligente

Las soluciones SOAR (Security Orchestration, Automation and Response) se perfilan como aliados estratégicos para la próxima generación de defensa digital. Estas plataformas integran herramientas diversas, automatizan respuestas y reducen tiempos críticos en plena crisis.

Paralelamente, los enfoques Zero Trust y la seguridad desde la fase de diseño están ganando tracción en sectores clave como finanzas, salud y administración pública. Incluir la seguridad desde el diseño del software no solo evita vulnerabilidades, sino que reduce costes y mejora tiempos de respuesta.

La inteligencia artificial ya no es una promesa lejana. Está aquí, modelando amenazas y oportunidades en tiempo real. Las organizaciones que no adapten sus estrategias, talentos y herramientas a esta nueva realidad quedarán expuestas a un entorno más rápido, más complejo y menos predecible.

El año 2026 marcará un punto de inflexión: no solo se trata de resistir ataques, sino de construir resiliencia inteligente, basada en IA gobernada, transparencia, agilidad y responsabilidad compartida.

La IA puede proteger los sistemas de las amenazas cibernéticas a través del monitoreo automatizado, cortafuegos inteligentes y estrategias de defensa de asistentes virtuales. Foto: Shutterstock.