La nueva víctima digital y cómo la inteligencia artificial está transformando el cibercrimen

Descripción

Introducción

Durante siglos, el crimen tuvo una geografía relativamente clara. Existía un lugar físico, una víctima identificable, un agresor y una escena concreta. El delincuente necesitaba aproximarse a su objetivo, exponerse al riesgo y actuar dentro de limitaciones físicas y temporales muy definidas. El robo, la estafa o la intimidación requerían proximidad y contacto humano directo. El crimen tenía fricción.

Sin embargo, en apenas dos décadas, el escenario criminológico ha cambiado radicalmente. Hoy, una persona puede perder todos sus ahorros sin haber visto nunca al delincuente. Un adolescente puede adquirir herramientas criminales listas para usar en Internet. Un directivo puede recibir una videollamada aparentemente legítima donde la voz y el rostro de su interlocutor han sido generados mediante inteligencia artificial. El delito ya no necesita desplazarse hasta la víctima porque la víctima vive permanentemente conectada al entorno donde opera el crimen.

Este cambio no representa solo una evolución tecnológica, sino una transformación profunda de la relación entre delincuente, víctima y sociedad. El cibercrimen moderno combina automatización, inteligencia artificial, manipulación emocional y explotación masiva de datos. El objetivo ya no es únicamente vulnerar sistemas informáticos, sino influir sobre la percepción y el comportamiento humano.

La frontera entre vida física y vida digital prácticamente ha desaparecido. El teléfono móvil, las redes sociales, la banca online y los asistentes virtuales forman parte constante de nuestra identidad cotidiana. Y cuando desaparece la separación entre espacio físico y espacio digital, también desaparece la antigua idea de “lugar seguro”.

La inteligencia artificial no solo automatiza el delito, también automatiza la persuasión. Los deepfakes y las herramientas generativas están erosionando progresivamente la confianza en la evidencia audiovisual, mientras los sistemas inteligentes permiten crear manipulaciones hiperrealistas adaptadas psicológicamente a cada víctima.

Este artículo analiza cómo la inteligencia artificial está transformando el ecosistema criminal contemporáneo y de qué manera puede utilizarse también para proteger a las personas frente a un entorno crecientemente automatizado y manipulativo.

Del delito físico al ecosistema criminal digital

La historia del crimen siempre ha evolucionado junto a la tecnología. La imprenta facilitó falsificaciones documentales, el automóvil transformó la movilidad criminal y las telecomunicaciones ampliaron el alcance del fraude internacional. Sin embargo, ninguna revolución anterior había alterado tan profundamente el concepto mismo de espacio criminal como Internet combinado con inteligencia artificial.

En el delito clásico existía una escena física delimitada. Había desplazamientos, contacto directo y exposición al riesgo. En cambio, el cibercrimen opera en un entorno deslocalizado donde las fronteras geográficas prácticamente desaparecen. Un atacante puede operar desde cualquier país sobre miles de víctimas simultáneamente.

El ecosistema criminal contemporáneo se caracteriza por cinco elementos fundamentales: internacionalización, anonimato, descentralización, escalabilidad y automatización. Herramientas como VPN, redes Tor y criptomonedas dificultan enormemente la identificación de los delincuentes, aumentando la percepción de impunidad.

Además, el cibercrimen ya no depende necesariamente de organizaciones jerárquicas tradicionales. El delito digital funciona frecuentemente mediante ecosistemas colaborativos distribuidos donde diferentes actores venden servicios especializados. Este fenómeno ha dado lugar al denominado *Crime-as-a-Service* (CaaS), donde cualquier persona puede adquirir kits de *phishing*, *malware* o herramientas de fraude listas para usar.

Europol advirtió que la inteligencia artificial generativa está acelerando precisamente esta industrialización del delito, reduciendo barreras técnicas y aumentando la sofisticación de los ataques incluso para actores con poca experiencia (IOCTA 2024).

La automatización representa quizá el cambio más disruptivo. Mientras un delincuente tradicional podía atacar una víctima cada vez, un sistema automatizado puede lanzar millones de ataques simultáneamente.

Pero el cambio más importante quizá sea victimológico. Antes, el delincuente debía buscar activamente a la víctima. Hoy la víctima permanece conectada constantemente al entorno donde operan los sistemas criminales. Cada correo electrónico, red social o aplicación móvil se convierte potencialmente en una superficie de ataque.

Cómo se construye una víctima

Uno de los grandes errores al analizar el cibercrimen consiste en pensar que las víctimas "caen" únicamente por falta de conocimientos técnicos. En realidad, la mayoría de los ataques digitales explotan vulnerabilidades cognitivas y emocionales profundamente humanas.

El cerebro humano no evolucionó para detectar *phishing*, *deepfakes* o manipulación algorítmica. Evolucionó para sobrevivir rápidamente en entornos físicos hostiles. Durante miles de años, reaccionar rápido era más importante que analizar con precisión. Hoy, los ciberdelincuentes explotan precisamente esos mecanismos automáticos.

El miedo constituye una de las herramientas más efectivas. Mensajes como "actividad sospechosa detectada" o "su cuenta será suspendida" activan respuestas emocionales inmediatas. La amígdala cerebral responde reduciendo temporalmente la capacidad de razonamiento crítico.

La urgencia funciona de forma similar. El cerebro humano tolera mal la incertidumbre. Frases como "última oportunidad" o "responda antes de 10 minutos" generan impulsos automáticos de acción rápida.

La euforia también reduce el escepticismo. Mensajes relacionados con premios, bonificaciones o reembolsos activan expectativas de recompensa asociadas a la dopamina. El delincuente no necesita convencer racionalmente a la víctima, solo interrumpir temporalmente su capacidad analítica.

La inteligencia artificial amplifica enormemente estas estrategias mediante personalización masiva. Logos familiares, estilos de escritura conocidos, voces clonadas y conversaciones adaptadas psicológicamente generan falsas sensaciones de confianza.

La misma lógica explica la propagación de *fake news*. La información falsa se difunde más rápido que la verdadera porque activa emociones intensas como miedo, indignación o sorpresa.

La nueva víctima digital es, por tanto, una persona permanentemente expuesta a sistemas automatizados diseñados para captar atención, manipular emociones y alterar percepciones utilizando vulnerabilidades humanas ancestrales.

Deepfakes, IA generativa y el colapso de la evidencia visual

Durante gran parte de la historia contemporánea, la imagen audiovisual fue considerada una forma privilegiada de verdad. La fotografía o el vídeo implicaban una conexión relativamente directa con la realidad observada. Sin embargo, la inteligencia artificial generativa ha alterado radicalmente esta relación.

Hoy resulta posible fabricar rostros inexistentes, clonar voces o recrear escenas falsas con niveles de realismo extremadamente convincentes. Entramos en una era donde la

evidencia visual deja progresivamente de ser evidencia.

Los *deepfakes* utilizan redes neuronales profundas capaces de imitar expresiones faciales, voces y movimientos humanos. El problema central no es únicamente técnico, sino cognitivo. El cerebro humano utiliza atajos perceptivos y tiende a aceptar automáticamente aquello que parece coherente con sus expectativas previas.

Por eso un *deepfake* no necesita ser perfecto. Basta con que sea suficientemente creíble dentro del contexto psicológico de la víctima. Si la voz resulta familiar y el mensaje parece lógico, el cerebro rellena automáticamente pequeñas inconsistencias.

Las consecuencias criminológicas son enormes. El fraude empresarial mediante suplantación de identidad se ha sofisticado enormemente gracias a la IA generativa. Los atacantes pueden recrear videollamadas, voces y estilos comunicativos para inducir transferencias fraudulentas.

Uno de los ámbitos más preocupantes es la pornografía no consentida generada mediante IA. Más del 90 % de los *deepfakes* online tienen contenido sexual dirigido principalmente contra mujeres. El daño ya no es solo económico, sino reputacional, emocional y social.

Además, los *deepfakes* introducen una crisis epistemológica muy peligrosa. No solo las falsificaciones parecen reales, sino que las pruebas reales pueden empezar a percibirse como falsas. La sociedad entra en un escenario donde la confianza en la evidencia visual se debilita progresivamente.

La convergencia entre IA, realidad virtual y simulación inmersiva amplifica todavía más este problema. El individuo ya no observa únicamente una narrativa manipulada, puede experimentarla emocionalmente como si fuese real.

Inteligencia artificial para proteger a la víctima

La misma inteligencia artificial que amplifica el potencial ofensivo del cibercrimen también está transformando profundamente las capacidades defensivas de las organizaciones y sistemas de protección.

La gran ventaja de la IA aplicada a la seguridad reside en su capacidad para analizar enormes volúmenes de información en tiempo real detectando anomalías y patrones imposibles de identificar manualmente.

Los sistemas clásicos de ciberseguridad funcionaban mediante reglas estáticas y firmas conocidas. El problema es que los atacantes evolucionan constantemente. La IA permite construir modelos adaptativos capaces de detectar comportamientos sospechosos incluso cuando el ataque concreto nunca había sido observado previamente.

La detección de anomalías constituye uno de los pilares fundamentales. Por ejemplo, si una persona realiza transferencias bancarias inusuales desde una localización extraña, el sistema puede activar alertas preventivas antes de que el daño sea irreversible.

La IA también permite comprender patrones victimológicos. Determinados colectivos presentan vulnerabilidades específicas frente a distintos tipos de fraude. La prevención deja de centrarse únicamente en proteger dispositivos para enfocarse también en proteger procesos cognitivos humanos.

Los sistemas modernos de detección de phishing analizan lenguaje sospechoso, comportamiento histórico de remitentes y características contextuales imposibles de evaluar manualmente a gran escala.

La detección automática de *deepfakes* se ha convertido igualmente en una prioridad crítica. Los sistemas analizan inconsistencias visuales y acústicas imperceptibles para el ojo humano, como patrones de parpadeo o irregularidades pixelarias.

Además, la IA puede proporcionar asistencia continua a víctimas mediante orientación inmediata, traducción automática o apoyo psicológico inicial. También posee gran potencial en investigación criminal, reconstrucción de escenas y entrenamiento policial mediante simuladores inmersivos.

Sin embargo, estas tecnologías plantean enormes desafíos éticos. Los sistemas de IA pueden discriminar, reproducir sesgos históricos o convertirse en mecanismos excesivos de vigilancia. El problema central ya no es únicamente tecnológico, sino humano y jurídico.

Conclusiones

La inteligencia artificial está transformando profundamente la relación entre delincuentes, víctimas y sociedad. El crimen contemporáneo ya no depende únicamente de fuerza física o proximidad espacial, sino de automatización, manipulación emocional y explotación masiva de datos.

La nueva víctima digital no es solamente alguien que pierde dinero o sufre un acceso no autorizado a sus dispositivos. Es una persona cuya atención, emociones y percepción de la realidad pueden ser manipuladas sistemáticamente mediante tecnologías diseñadas para explotar vulnerabilidades cognitivas profundamente humanas.

Los *deepfakes*, la IA generativa y la automatización emocional están erosionando progresivamente la confianza social en la evidencia visual y en la autenticidad de las interacciones digitales. Entramos en una era donde la principal superficie de ataque ya no será únicamente el ordenador, sino la propia percepción humana.

Sin embargo, la inteligencia artificial también ofrece capacidades extraordinarias de protección. La detección de anomalías, los sistemas preventivos y la identificación automática de falsificaciones pueden reducir enormemente la exposición al daño.

El verdadero desafío del futuro no será únicamente tecnológico. Será ético, social y cognitivo. Requerirá alfabetización digital, pensamiento crítico y una profunda reflexión sobre cómo convivir con sistemas capaces de influir directamente sobre nuestras emociones y decisiones.

Porque quizá la gran pregunta del siglo XXI ya no sea si las máquinas pueden pensar como nosotros, sino si nosotros seremos capaces de seguir distinguiendo entre realidad, simulación y manipulación en un entorno crecientemente diseñado para influir sobre nuestra mente.

Descarga el informe completo

Si te ha gustado el artículo y quieres saber más, descarga nuestro informe completo de
ɔ de víctima y cómo la inteligencia artificial está transformación

Descarga el informe

Referencias

- Cialdini, R. B. (2021). *Influence: The Psychology of Persuasion* (Revised ed.). Harper Business.
- Cohen, L. E., & Felson, M. (1979). Social change and crime rate trends: A routine activity approach. *American Sociological Review, 44*(4), 588–608.
- Colado García, S. (2021). *Multiversos digitales: la tecnología como palanca evolutiva*. Universo de Letras.
- Colado García, S. (2019). *Influencia de la tecnología en el desarrollo del pensamiento y conducta humana*. Amazon autoedición.
- ENISA. (2024). *ENISA Threat Landscape: Finance Sector 2024*. European Union Agency for Cybersecurity.
- Europol. (2024). *Internet Organised Crime Threat Assessment (IOCTA) 2024*. European Union Agency for Law Enforcement Cooperation.
- Federal Bureau of Investigation. (2024). *Internet Crime Report 2024*. Internet Crime Complaint Center (IC3).
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- Kahneman, D. (2011). *Thinking, Fast and Slow*. Farrar, Straus and Giroux.
- Lazer, D. M. J., Baum, M. A., Benkler, Y., et al. (2018). The science of fake news. *Science, 359*(6380), 1094–1096.
- McAfee. (2023). *The State of Scams and Deepfake Fraud Report*.
- Miller, B. L. (2020). Deepfakes and synthetic media in the financial industry. *Journal of Financial Crime, 27*(4), 1219–1230.
- National Institute of Standards and Technology (NIST). (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*.
- Newman, G. R., & Clarke, R. V. (2003). *Superhighway Robbery: Preventing E-Commerce Crime*. Willan Publishing.
- OpenAI. (2024). *GPT-4 Technical Report*. OpenAI Research.
- Schneier, B. (2015). *Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World*. W. W. Norton & Company.
- Singer, P. W., & Brooking, E. T. (2018). *LikeWar: The Weaponization of Social Media*. Houghton Mifflin Harcourt.
- Taylor, S., & Fragkos, K. (2022). Cybersecurity and human vulnerability: Psychological dimensions of phishing susceptibility. *Computers & Security, 114*, 102577.
- Tufekci, Z. (2018). YouTube, the great radicalizer. *The New York Times*.
- Vishwanath, A., Herath, T., Chen, R., Wang, J., & Rao, H. R. (2011). Why do people get phished? Testing individual differences in phishing vulnerability within an integrated,

information processing model. *Decision Support Systems, 51*(3), 576–586.

- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science, 359*(6380), 1146–1151.