Inteligencia Artificial en la actividad delictiva y el control del delito

Artificial Intelligence in Criminal Activity and Crime Control

Traducido del inglés por:
Aranza Castillo Tolsá, ENALLT, UNAM

Riya Gulati

Licenciada en Derecho por Bharati Vidyapeeth Deemed University, Pune, India; maestra en Derecho por la University College Dublín, Irlanda, con especialización en Propiedad Intelectual y Tecnologías de la Información
ORCID: https://orcid.org/0000-0001-8060-774X
Correo electrónico: riyagulati0205@gmail.com

Inteligencia Artificial en la actividad delictiva y el control del delito

Artificial Intelligence in Criminal Activity and Crime Control

Riya Gulati

University College Dublín

Recepción: 17/12/2025 **Aceptación:** 14/01/2026

DOI: https://doi.org/10.57042/rmcp.v9i29.1079

Resumen

La inteligencia artificial (IA) ha revolucionado innumerables industrias, ha mejorado la productividad y la innovación y, a la vez, ha modificado la manera en que los delitos se perpetran y, por lo tanto, se previenen. La IA ha sido explotada con fines delictivos, lo que supone graves amenazas para la seguridad mundial y la estabilidad social. Al mismo tiempo, es aprovechada para prevenir delitos mediante la detección avanzada de riesgos, la supervisión en tiempo real y el análisis predictivo. Este artículo examina el doble papel de la IA como facilitadora del delito y como herramienta constructiva para su prevención, detección y respuesta.

Palabras clave

Inteligencia artificial, comisión del delito, prevención del delito, gobernanza de la IA, protección de datos, análisis predictivo.

Abstract

Artificial Intelligence (AI) has revolutionised countless industries, enhancing productivity and innovation, while also transforming the ways in which crime is both perpetrated and prevented. AI is being exploited for criminal purposes, posing serious threats to global security and societal stability. At the same time, it is being leveraged to avert crime through advanced threat detection, real-time monitoring of suspicious activity, and predictive analytics. This paper examines ai's dual role as both an enabler of crime and a constructive tool for preventing, detecting, and responding to it.

Keywords

Artificial Intelligence, Crime Commission, Crime Prevention, AI Governance, Data Protection, Predictive Analytics.

Sumario

I. Introducción. II. La IA como facilitadora del delito. III. La IA como herramienta de control del delito. IV. Retos en la prevención del delito basada en la IA. V. Regulaciones globales de la IA en el ámbito penal. VI. Conclusión. VII. Recomendaciones. VIII. Referencias.

I. Introducción

Las tecnologías han impactado primordialmente en las economías nacionales, marcando el cambio de una base industrial convencional hacia una base informacional. Es indudable que la tecnología desempeña un papel crucial en la comisión de múltiples formas de actividad delictiva, desde infracciones menores hasta delitos graves. La

información digital —los datos— constituye— la forma mediante la cual la emergente sociedad de la información lleva a cabo sus actividades. La economía global depende en gran medida del procesamiento y la transmisión de esa información a través de redes para sostener su infraestructura y llevar a cabo numerosas funciones (Walden, 2007). La inteligencia artificial se ha convertido en parte de la vida cotidiana y conlleva el riesgo de dotar a Estados, empresas y corporaciones de nuevas herramientas que pueden menoscabar la privacidad de las personas. La mayoría de las aplicaciones, los sistemas y los productos basados en IA se sustentan principalmente en la recopilación y el procesamiento de grandes volúmenes de datos, a menudo sin el conocimiento y, de manera fundamental, sin el consentimiento de quienes se encuentran en su entorno (Ryder y Naren, 2025).

Una de las técnicas más relevantes de la IA es el uso de las redes neuronales (Marwala, 2013). En el ámbito del delito, éstas se emplean tanto en dimensiones ofensivas como defensivas. Los actores delictivos explotan las capacidades de la IA y emplean redes neuronales para automatizar ataques de suplantación de identidad, evadir sistemas de ciberseguridad e identificar vulnerabilidades en infraestructuras digitales. De manera paralela, los organismos encargados de hacer cumplir la ley las emplean para detectar actividades fraudulentas, predecir comportamientos delictivos y procesar grandes volúmenes de datos de vigilancia o de redes sociales con fines de investigación. El avance de la tecnología ha permitido nuevas formas de cometer, combatir y reflexionar sobre la delincuencia, las personas que cometen delitos, la policía, los tribunales, las víctimas y la ciudadanía. La tecnología no sólo proporciona herramientas innovadoras para cometer y combatir el delito, sino también nuevos mecanismos para detectar, investigar y calificar los delitos, así como

nuevas formas de identificar, vigilar, perseguir y sancionar a quienes los cometen.

El poder de la tecnología, en constante crecimiento exponencial, ha dotado de estrategias sofisticadas y mayores oportunidades a quienes cometen delitos, al ampliar el número de objetivos potenciales, sin que hasta ahora se haya producido una mejora correspondiente en la disponibilidad o la capacidad de los sistemas de protección. Además, las redes transnacionales de información han transformado la delincuencia tecnológica de un problema local en una cuestión de seguridad global (Leman-Langlois, 2013).

La protección de datos constituye un ámbito en el que los derechos fundamentales deben ser ponderados frente al impulso hacia la innovación tecnológica y la eficacia económica. Los datos personales se han convertido en un activo esencial para la economía digital. En consecuencia, su libre flujo transfronterizo ha sido descrito como un nuevo campo de batalla para los estados que buscan salvaguardar sus principales intereses económicos y no económicos, en especial en un contexto en el que la gobernanza digital y las preocupaciones de seguridad vinculadas a la IA se han vuelto centrales (Naef, 2023).

Con las tecnologías emergentes, tener acceso a datos sin autorización se ha vuelto progresivamente más fácil. Hoy en día, los ciberataques son poco costosos y quienes los llevan a cabo asumen un riesgo muy bajo de ser descubiertos, ya que pueden ejecutarse desde cualquier parte del mundo, lo que constituye un entorno propicio para actividades delictivas con consecuencias potencialmente catastróficas (Koehler, 2018). El desafío que plantean los delitos facilitados por la IA radica en que los métodos empleados para gestionar el desorden en el mundo físico suelen resultar insuficientes en el ámbito digital. Los problemas que

surgen en el ciberespacio no permanecen confinados, sino que se trasladan al mundo real, ya que el ciberespacio no es una verdadera externalidad. Es simplemente un conducto de la actividad humana, tanto positiva como negativa. Permite que los peores elementos de cualquier lugar se infiltren en cualquier otro, y ello constituye un aspecto central del problema (Brenner, 2009).

La IA es una de las áreas emergentes de esta sociedad tecnocrática y plantea importantes desafíos. El principal reto para el derecho consiste en mantener el ritmo de los avances tecnológicos, cuyo desarrollo corresponde a los innovadores, mientras que su uso indebido para la comisión de delitos facilitados por la IA recae en quienes los cometen. La respuesta frente a estos fenómenos exige una actuación coordinada de los poderes legislativo, ejecutivo y judicial. En este contexto, al abordar los delitos habilitados por la IA, el refrán tradicional de "ojo por ojo" debe transformarse en un principio de "tecnología por tecnología", en el que herramientas tecnológicas avanzadas sean empleadas para prevenir, detectar y contrarrestar las actividades delictivas impulsadas por la IA (Kalra y Verma, 2011).

II. La IA como facilitadora del delito

La inteligencia artificial plantea amenazas diferenciadas y emergentes en el ámbito de la actividad delictiva. Éstas derivan de la autonomía de la IA, de sus capacidades de aprendizaje y de su aptitud para operar a velocidades y escalas que superan la supervisión humana. Los comportamientos emergentes, en los que los sistemas de IA actúan de formas imprevistas, pueden generar de manera inadvertida resultados delictivos, como la facilitación de actividades ilegales coordinadas por sistemas autónomos. Los marcos conven-

cionales de responsabilidad se ven cuestionados, ya que la IA puede ejecutar actos delictivos (*actus reus*) de forma independiente a la intención humana (*mens rea*), lo que plantea problemas en materia de atribución de responsabilidad y aplicación de la ley.

La supervisión de los delitos habilitados por la IA resulta igualmente compleja, ya que los agentes autónomos pueden adaptarse con mayor rapidez de la que permite la supervisión humana, operar a través de múltiples sistemas y explotar vulnerabilidades sociales y psicológicas, como la manipulación de usuarios mediante *bots* sociales. Los avances en la IA generativa, particularmente en la creación sintética de información, han posibilitado nuevas formas de acoso mediante la generación de representaciones altamente realistas, pero falsas de individuos. Este fenómeno complica la atribución jurídica y la rendición de cuentas. Además, la IA ha facilitado el robo y el fraude mediante la automatización en la recopilación de datos personales y la sofisticación en la manipulación de identidades digitales. Las campañas automatizadas de *spear phishing* han incrementado la eficacia de estos ataques, mientras que la síntesis de voz impulsada por la IA permite a quienes cometen delitos suplantar a las víctimas, eludir las medidas de seguridad y llevar a cabo actividades ilegales (King, Aggarwal, Taddeo y Floridi, 2019).

La IA ha transformado de manera sustancial la dinámica operativa de la actividad delictiva al eliminar las limitaciones humanas convencionales, lo que permite la comisión de delitos más rápidos, eficientes y sofisticados. Quienes cometen ciberdelitos han adoptado la IA para potenciar operaciones dirigidas por personas, automatizando tareas como campañas de *phishing*, el análisis de extensas bases de código en busca de vulnerabilidades ex-

plotables y la generación de comunicaciones fraudulentas convincentes. La IA también ha permitido a quienes cometen delitos ampliar sus operaciones mucho más allá de los límites humanos, impulsando el auge de las suplantaciones mediante *deepfakes*, las identidades falsas y los ciberataques a gran escala dirigidos contra individuos, empresas e instituciones financieras con una precisión sin precedentes (TRM, 2025).

La proliferación de la IA en la actividad delictiva ha generado un aumento significativo en los delitos facilitados por esta tecnología, con graves implicaciones para las fuerzas del orden y la seguridad nacional. Actualmente, la IA se emplea de manera extensiva en delitos financieros, *phishing*, ataques de Denegación de Servicio Distribuido (DDoS, por sus siglas en inglés), difusión de material de abuso sexual infantil (CSAM, por sus siglas en inglés), robo de identidad, ciberacoso, creación de *deepfakes*, desinformación; así como en sistemas de vigilancia y vulneraciones de la privacidad. El incremento de esta modalidad delictiva responde a diversos factores. En particular, las capacidades de automatización de la IA permiten escalar rápidamente las operaciones ilícitas, lo que ha aumentado la magnitud de la actividad delictiva a niveles antes inimaginables.

Quienes delinquen no sólo automatizan tipos de delitos ya existentes, como el *phishing* y el fraude, sino que también aprovechan la IA para desarrollar métodos novedosos y herramientas más avanzadas. Además, la capacidad de la IA para replicar comportamientos similares a los humanos permite a estos actores explotar vulnerabilidades psicológicas y cognitivas generalizadas, lo que incrementa la eficacia de tácticas engañosas como la ingeniería social, los *deepfakes* y la generación automatizada de contenido.

Los sistemas de IA se utilizan para crear herramientas delictivas más potentes, lo que exige que las fuerzas del orden desarrollen capacidades igualmente avanzadas para contrarrestarlas. Sin embargo, pese a las considerables ventajas que la IA proporciona a quienes cometen delitos, persisten obstáculos que limitan su implementación a mayor escala, como la necesidad de conocimientos técnicos especializados, recursos computacionales sustanciales, seguridad operativa y sistemas de IA fiables. Estos obstáculos constituyen puntos de intervención para las fuerzas del orden a fin de interrumpir y contrarrestar los delitos facilitados por esta tecnología. Al identificar y abordar estas limitaciones, las autoridades pueden ralentizar su proliferación en las actividades delictivas (Burton, Joe, Ardi Janjeva, Simon Moseley, Alice, 2025).

Los avances en este ámbito, en particular en la IA generativa, han transformado de manera significativa la actividad delictiva, al permitir que una amplia variedad de delitos se cometa con mayor rapidez, eficacia y escala. Estos desarrollos tecnológicos incluso facilitan la actuación de quienes cometen delitos con menor nivel de especialización, al reducir brechas de conocimiento, como el dominio de idiomas, la programación avanzada y otros conocimientos técnicos. Dado el rápido progreso de esta tecnología, no sólo han surgido nuevas formas de delito, sino que también las ya existentes se han transformado.

Los avances en la IA han incrementado la vulnerabilidad frente a diversas actividades delictivas, en la medida en que el contenido generado por esta tecnología se vuelve indistinguible del producido por personas. Determinados grupos en situación de vulnerabilidad, como niños, personas mayores y mujeres, presentan un riesgo particularmente elevado. Para mitigar los daños asociados con los

delitos facilitados por esta tecnología, se requiere una sólida colaboración intersectorial, mayor concientización y participación pública, el desarrollo de mecanismos de defensa innovadores y la incorporación de conocimientos y buenas prácticas en el ámbito de la ciberseguridad.

La capacidad de la IA para generar contenido altamente convincente, desde correos electrónicos y mensajes de texto hasta la clonación de la voz y los *deepfakes*, resulta particularmente evidente en delitos como el *phishing*, el robo de identidad y el fraude financiero. Estas estafas han vuelto cada vez más difíciles de detectar, ya que los actores maliciosos pueden crear mensajes personalizados y realistas que explotan vulnerabilidades humanas. La IA también posibilita nuevas formas de actividad delictiva, como los ataques contra infraestructuras críticas, en los que sistemas basados en esta tecnología podrían emplearse para controlar vehículos autónomos o infiltrarse en sistemas industriales. Asimismo, el *malware* desarrollado mediante IA ha evolucionado y se ha optimizado de forma autónoma, lo que lo ha vuelto más difícil de detectar y dificulta que las medidas convencionales de ciberseguridad mantengan el ritmo frente a estas amenazas.

Los actores delictivos pueden utilizar la IA para explotar vulnerabilidades en los sistemas de manera más rápida y eficiente, así como para desarrollar *malware* adaptativo capaz de modificar su comportamiento en respuesta a mecanismos de detección. El auge de esta tecnología en el ciberdelito ha alterado de manera fundamental las dinámicas de la delincuencia en línea, lo que vuelve inadecuados los enfoques convencionales de detección y prevención. Asimismo, su papel en la sextorsión y en la creación de material de abuso sexual infantil (CSAM) va en aumento, ya que los *deepfakes* y el contenido explícito generado mediante

estas herramientas dificultan que las víctimas demuestren su carácter fabricado, mientras causan un daño emocional y psicológico significativo.

La asequibilidad y facilidad del uso de las herramientas de IA han reducido la barrera de entrada para quienes cometen delitos, lo que ha convertido estas tecnologías en una amenaza extendida en diversos sectores. Estas herramientas también han facilitado a los grupos terroristas la difusión de propaganda, el reclutamiento de seguidores y la generación de desinformación persuasiva a través de múltiples plataformas. Asimismo, han propiciado la aparición de nuevos ataques ciberfísicos, en los que actores maliciosos explotan sistemas integrados en infraestructuras críticas. Entre ellos se incluyen ataques contra vehículos autónomos, drones y sistemas de ciudades inteligentes, donde su manipulación puede causar daños físicos, interrumpir servicios públicos o incluso provocar consecuencias catastróficas. Estos ataques representan una alarmante convergencia de amenazas digitales y físicas, en la que los sistemas podrían utilizarse como armas para causar daños a gran escala. Además, la expansión de esta tecnología en el procesamiento de información sensible ha convertido a estos sistemas en un objetivo prioritario para la vulneración de datos privados. Su capacidad para gestionar grandes volúmenes de información incrementa la vulnerabilidad frente al acceso no autorizado y la explotación de datos personales (U.S. Department of Homeland Security, 2024).

III. La IA como herramienta de control del delito

El delito se ha vuelto más complejo debido a los avances tecnológicos. Para hacer frente a este desafío, las fuerzas del orden y el sector privado han adoptado la inteligencia artificial con el fin de fortalecer las investigaciones, mejorar la eficiencia y aumentar la eficacia de las estrategias preventivas. Esta tecnología permite a las autoridades procesar grandes volúmenes de datos, identificar patrones y detectar tendencias vinculadas con la delincuencia organizada y otras actividades ilícitas.

La IA desempeña un papel cada vez más relevante en el análisis de la *dark web*, al facilitar la detección de una amplia gama de actividades ilícitas mediante técnicas de procesamiento automático de grandes volúmenes de datos y la identificación de patrones sospechosos. Esta herramienta puede reconocer comportamientos delictivos, supervisar plataformas en línea, detectar posibles víctimas de delitos como la trata de personas y descubrir redes ocultas. Asimismo, permite aplicar análisis predictivo, minería de datos, reconocimiento facial y análisis de patrones para facilitar la localización de los responsables y la prevención de delitos futuros.

A pesar de sus ventajas, el uso de la IA en las fuerzas del orden plantea diversos desafíos, entre ellos el presupuesto, las carencias de formación y las preocupaciones éticas relacionadas con la privacidad y los derechos humanos. Por ello, resultan fundamentales el desarrollo responsable de esta tecnología, la regulación adecuada y la capacitación continua del personal encargado de hacer cumplir la ley. Al equilibrar la innovación con la rendición de cuentas, la IA puede consolidarse como una herramienta poderosa para fortalecer los sistemas de justicia penal y mejorar la seguridad pública (Sadulski, 2024).

La seguridad de una comunidad constituye una prioridad fundamental, lo que impulsa a los gobiernos a adoptar medidas orientadas a reducir la actividad delictiva y contribuye al crecimiento económico y a la mejora de la calidad de vida. El análisis del fenómeno delictivo desempeña un papel relevante en la comprensión de los patrones de comportamiento y en la identificación de indicadores de actividad ilícita. No obstante, su prevención resulta compleja debido a la diversidad de conductas, motivaciones, repercusiones, métodos de actuación y estrategias preventivas. Para hacer frente a estos desafíos, la predicción del delito ha surgido como una herramienta poderosa que permite a las fuerzas del orden identificar tendencias y asignar recursos con mayor eficacia.

Debido a los avances en inteligencia artificial y aprendizaje automático, los modelos predictivos pueden analizar grandes volúmenes de datos, detectar patrones y anticipar posibles conductas delictivas. Estos modelos integran múltiples variables, como el tipo de infracción, la ubicación, el momento y los factores demográficos o socioeconómicos, lo que facilita la identificación de focos de riesgo y la anticipación de la actividad delictiva. A diferencia de los métodos convencionales, como el mapeo de zonas de concentración, que se basan en datos históricos para desplegar recursos de manera reactiva, la IA genera información prospectiva que optimiza la asignación de recursos, mejora las estrategias de patrullaje y fortalece la seguridad pública. Además, permite desarrollar estrategias preventivas que además de anticipar comportamientos ilícitos, también revelan dinámicas sociales subyacentes y respaldan intervenciones focalizadas y decisiones de política pública basadas en evidencia.

La incorporación de la IA en la predicción del delito representa un avance significativo en la labor policial contemporánea, con potencial para reducir los índices de delincuencia y mejorar la eficacia institucional y el bienestar comunitario (Dakalbab, Talib, Waraga, Nassif, Abbas y Nasir, 2022). Asimismo, la inteligencia artificial ha transformado las fuerzas del orden al posibilitar enfoques más proactivos y sustentados en datos para la prevención e investigación. Mediante la automatización del análisis de conjuntos complejos de información, como redes sociales, registros financieros, sistemas de videovigilancia y expedientes policiales, las autoridades pueden detectar patrones ocultos, identificar amenazas emergentes y responder con mayor rapidez ante situaciones en curso.

Los análisis predictivos y los modelos de aprendizaje automático han demostrado especial eficacia en la identificación de zonas de concentración delictiva, individuos de alto riesgo y comportamientos financieros anómalos, lo que, sin duda, contribuye a una gestión más estratégica de los recursos. La implementación de herramientas basadas en IA en la seguridad urbana, la prevención del ciberdelito y la detección de fraudes financieros han contribuido a reducciones medibles en las tasas de delincuencia en los contextos en que se han implementado (Frenkel, 2025).

La integración de tecnologías intensivas en los datos de las investigaciones delictivas ha ampliado de forma significativa el papel del tratamiento de la información en la labor policial proactiva. Herramientas como los sistemas de vigilancia y las técnicas de minería de datos permiten a las fuerzas del orden recopilar, analizar y aprovechar grandes volúmenes de información, lo que fortalece las estrategias de prevención e investigación. Al mismo tiempo, el uso de esta información acentúa la necesidad de respetar

los derechos fundamentales, en particular, la protección de datos, en las fases de acceso, tratamiento e intercambio de contenidos procedentes de bases policiales y judiciales. Este principio se aplica tanto a los archivos generados en el marco de investigaciones como a los derivados de la persecución penal, lo que asegura que los avances tecnológicos se equilibren con las debidas salvaguardias jurídicas y éticas (Hijmans y Kranenborg, 2014).

Los investigadores digitales requieren competencias técnicas adecuadas, recursos suficientes y facultades de investigación claras. No obstante, el volumen y la diversidad del material susceptible de análisis forense, junto con las limitaciones temporales y presupuestarias, dificultan la implementación de soluciones integrales. Parte de los costos puede trasladarse o compartirse entre las entidades que generan y almacenan información, como los proveedores de servicios de comunicación, aunque las limitaciones estructurales persistirán. Además, quienes cometen delitos emplean herramientas destinadas a mantener las pruebas digitales fuera del alcance de las autoridades de manera permanente.

Si bien los avances tecnológicos y del mercado pueden complicar las labores de las fuerzas del orden, también ofrecen mecanismos de control del delito que contribuyen a reducir la incidencia y el impacto de la ciberdelincuencia (Walden, 2007). Las fuerzas del orden han perfeccionado técnicas de policía predictiva que permiten identificar tendencias delictivas y prevenir, en lugar de limitarse a reaccionar ante futuros delitos mediante la agregación y el análisis de diversos datos sobre el comportamiento de las personas (Friedman, Ekdahl y Thomas, 2015). La integración de la IA con la robótica y los drones para tareas de vigilancia e intervención constituye una alternativa más

segura que el despliegue de agentes en situaciones de alto riesgo, al reducir el potencial de daño y mantener operaciones eficaces de seguridad pública (Rigano, 2019).

La predicción del delito se ha consolidado como un componente clave de la labor policial moderna, al permitir que las fuerzas del orden transiten de respuestas reactivas a un enfoque preventivo proactivo. Los sistemas basados en IA desempeñan un papel central en la denominada policía inteligente, al apoyar la toma de decisiones investigativas, optimizar la asignación de recursos, reducir los costos operativos y, potencialmente, disminuir las tasas de delincuencia. Sin embargo, pese a su promesa técnica, estos sistemas plantean importantes preocupaciones éticas, jurídicas y sociales. Los modelos de IA pueden reproducir de manera inadvertida sesgos históricos presentes en los datos sobre delincuencia, afectando de forma desproporcionada a comunidades desfavorecidas y socavando la equidad en los procesos de justicia penal.

Si bien la inteligencia artificial ha impulsado de manera significativa la predicción del delito, su implementación plantea desafíos críticos relacionados con la equidad, la rendición de cuentas, la explicabilidad, la interpretabilidad, la privacidad y la seguridad. Abordar estas cuestiones resulta fundamental para garantizar que los sistemas basados en IA no sólo sean técnicamente eficaces, sino también éticamente sólidos y socialmente responsables. A medida que esta tecnología se integra cada vez más en las prácticas de las fuerzas del orden, es esencial mantener un equilibrio entre el rendimiento predictivo y la confianza pública, a fin de prevenir usos indebidos, sesgos y daños no intencionados.

En las investigaciones penales, la capacidad de comprender, explicar y justificar las predicciones generadas por la IA es fundamental. La explicabilidad resulta, por tan-

to, central para generar confianza en estos sistemas, ya que los resultados transparentes e interpretables favorecen la rendición de cuentas en las fuerzas del orden y refuerzan la legitimidad de su aplicación ética. Al mismo tiempo, salvaguardar la integridad de los datos, la privacidad y la seguridad es un requisito esencial, dada la naturaleza sensible de la información vinculada a la delincuencia. Asimismo, debe mantenerse un énfasis constante en las consideraciones éticas, incluidas la equidad, la rendición de cuentas y el impacto social (Ersöz, Filiz, Taner Ersöz, Marcelloni y Ruffini, 2025).

La IA también tiene el potencial de mejorar los resultados en la justicia penal al aumentar la eficacia, reducir el error humano, reforzar la coherencia en la toma de decisiones y apoyar una asignación de recursos basada en datos; cuando se utiliza de manera responsable, puede fortalecer la confianza pública al incrementar la fiabilidad y la equidad en el funcionamiento del sistema de justicia. No obstante, la supervisión humana debe mantenerse como eje central, con mecanismos de reparación disponibles en caso de errores o daños. Asimismo, las instituciones necesitan formación en fundamentos de IA para comprender las limitaciones de los modelos, los sesgos de los datos, la degradación del rendimiento y los riesgos asociados a los sistemas de "caja negra".

Las estrategias proactivas, como los programas piloto, la colaboración interdisciplinaria, las evaluaciones independientes y la supervisión continua son fundamentales para evaluar la confiabilidad y el impacto real. Asimismo, se requieren marcos sólidos de gobernanza y regulación. La participación de las partes interesadas es esencial para garantizar la legitimidad. Generar confianza pública exige una comunicación transparente, consultas inclusivas, ex-

plicaciones en un lenguaje claro y un compromiso sostenido con las comunidades afectadas, los profesionales, los expertos tecnológicos y la sociedad civil. La capacitación del personal de primera línea y la incorporación de su experiencia en la implementación de la IA fortalecen la eficacia y la aceptación de estas tecnologías (Council on Criminal Justice, 2024).

IV. Retos en la prevención del delito basada en la IA

A medida que las tecnologías digitales avanzan, también aumenta el número de herramientas disponibles para lanzar ciberataques. Estos pueden proporcionar a quienes cometen delitos una gran cantidad de información sobre una persona o una ubicación determinada. Los riesgos cibernéticos y los desafíos en materia de ciberseguridad son generalizados. Resulta necesario reforzar la protección de los activos de información frente a la exposición, los ataques, la pérdida, el robo o el daño. Las cuestiones éticas, las deficiencias en la validación de la identidad, las preocupaciones relativas a la privacidad, los desafíos técnicos en la recopilación y conservación de datos, así como las dificultades para corroborar pruebas plantean serias dudas sobre la eficacia práctica de las medidas actuales de ciberseguridad (Johnson, 2013).

La era de la información plantea una cuestión clave: el uso y abuso de la información por parte de actores económicos y políticos como medio de control. El Estado procura utilizar los datos, a menudo generados a través de comunicaciones privadas, como herramienta para reforzar el orden público. La impotencia inicial del Estado frente a la actividad en internet ha dado paso al reconocimiento de su potencial singular para llevar un seguimiento de las co-

nexiones, el paradero y los movimientos de las personas, bajo la premisa de que toda persona representa un riesgo potencial de delinquir.

Persiste un conflicto entre los límites legítimos de la vigilancia y la expectativa de privacidad, libre de injerencias gubernamentales. El paradigma de la privacidad se encuentra en tensión con el estado de vigilancia y también con el fenómeno de la accesibilidad de los datos. Generalmente, la privacidad se reconoce como un derecho humano fundamental y goza de protección específica en los tratados internacionales de derechos humanos y en las constituciones nacionales.

Pueden identificarse tres factores de riesgo que pueden considerarse dimensiones de la privacidad. El primero de ellos es el riesgo de injusticia derivado de inexactitudes significativas en los datos personales, inferencias indebidas, la reutilización progresiva de los datos para fines distintos de aquellos para los que fueron recopilados, o la inversión de la presunción de inocencia, como ocurre en el cruce de datos, cuando la combinación de información procedente de fuentes dispares puede generar una impresión superior a la suma de sus partes. El segundo riesgo se refiere al control del individuo sobre la recopilación de su información personal como consecuencia de la vigilancia excesiva e injustificada, la obtención de datos sin consentimiento y la obstaculización de los medios destinados a remediar estos riesgos, como el uso del cifrado y de programas de anonimización. Por último, el riesgo que existe para la dignidad personal, derivado de la exposición o la vergüenza ocasionadas por la falta de transparencia en los procedimientos de tratamiento de la información, la intrusión física en espacios privados, la identificación innecesaria o la ausencia de anonimato, así como la divulgación injustificada de datos personales sin consentimiento.

En contraste con las normas relativas a la protección de la privacidad y de los datos personales, se encuentran aquellas que formalizan, justifican y regulan las actuaciones del Estado y de particulares que pueden afectar las expectativas razonables de privacidad de las personas, en aras de otros objetivos sociales, económicos y políticos legítimos. Éstas incluyen normas que regulan las tecnologías de la información, como las telecomunicaciones e internet, al facilitar el rastreo de vínculos entre individuos (lo que permite la recopilación de "datos de tráfico" que identifican cuándo y con quién se comunican los usuarios de dichas tecnologías); la recopilación de información detallada sobre las interacciones de las personas (mediante la interceptación del contenido de las comunicaciones); o al impedir el uso efectivo de mecanismos de protección frente a la vigilancia (lo que prohíbe o limita el uso de tecnologías de cifrado).

En el entorno digital, el objetivo principal de la vigilancia estatal ha sido garantizar que las fuerzas del orden y los organismos de seguridad nacional dispongan de acceso suficiente y facultades necesarias para mantener prácticas de investigación eficaces en la amplia diversidad de canales públicos de comunicación. Alcanzar y mantener un equilibrio proporcional entre eficacia y legitimidad en un ámbito en el que la tecnología se encuentra en constante transformación dista mucho de ser una tarea sencilla. Un factor que añade complejidad es que las facultades otorgadas a los organismos estatales para tener acceso y recopilar datos digitales generados por el público suelen producir o permitir la generación de conjuntos de datos relevantes para organizaciones comerciales (Rowland, Kohl y Charlesworth, 2012).

V. Regulaciones globales de la IA en el ámbito penal

Las regulaciones que rigen el uso de la IA en materia penal han surgido a nivel regional, nacional e internacional, y reflejan un equilibrio entre la innovación y la protección de los derechos humanos. En la Unión Europea, la Ley de Inteligencia Artificial prohíbe los sistemas de IA destinados a predecir la probabilidad de que una persona cometa delitos basándose exclusivamente en la elaboración de perfiles (artículo 5, letra d) y restringe la identificación biométrica en tiempo real por parte de las fuerzas del orden a investigaciones específicas (artículo 5, letra h), de forma complementaria al Reglamento General de Protección de Datos (RGPD), el cual limita la toma de decisiones automatizadas y exige evaluaciones de impacto en materia de protección de datos para aplicaciones sensibles de IA.

La información inmaterial se ha convertido en un activo clave, el combustible que impulsa la "economía de la información", entre cuyos activos los datos personales constituyen una categoría fundamental. De conformidad con el artículo 10 del RGPD, los datos personales relativos a condenas penales o infracciones sólo pueden ser objeto de tratamiento al amparo de una autoridad o en virtud de una autorización legal, con las garantías adecuadas para la protección de los derechos de las personas, y los registros de condenas penales deben permanecer sujetos a control oficial.

A nivel global, organismos de las Naciones Unidas como la UNODC y la UNESCO proporcionan orientación en materia de ética de la IA y ciberseguridad, haciendo hincapié en los derechos humanos, la rendición de cuentas y el doble papel de la IA como herramienta tanto para la comisión de delitos como para el cumplimiento de la ley (Jejelola, 2024). En Estados Unidos, la ley *Take It Down* (TIDA) establece la

obligación de eliminar de las plataformas en línea, en un plazo de 48 horas desde la notificación, las imágenes íntimas difundidas sin consentimiento, incluidos los *deepfakes*. La ley impone sanciones penales y obligaciones de restitución a los infractores, con el objetivo de salvaguardar la privacidad de las personas y prevenir su explotación.

Asimismo, las medidas para el etiquetado de contenido sintético generado por IA de China exigen que todo contenido generado por ésta sea identificado mediante marcadores explícitos e implícitos. Estas etiquetas permiten identificar el contenido generado por IA y prevenir su uso indebido, como la desinformación. Asimismo, las medidas prohíben la alteración maliciosa o la eliminación de dichas etiquetas, con el fin de reforzar la transparencia y proteger a los usuarios frente a contenido sintético engañoso (Zhang, 2025). En el Reino Unido, la *Online Safety Act* 2023 tipificó como delito la difusión o la amenaza de difusión de imágenes íntimas generadas mediante tecnología *deepfake* sin consentimiento.

No obstante, las modificaciones propuestas en 2025 amplían el alcance de esta normativa al tipificar como delito la creación de *deepfakes* con contenido sexual explícito sin consentimiento, con penas de hasta dos años de prisión para quienes actúen con la intención de causar daño o angustia, o con fines de gratificación sexual (Patishman, 2025). Asimismo, la *Criminal Code Amendment* (*Deepfake Sexual Material*) *Act 2024* de Australia refuerza la seguridad en línea al tipificar como delito la creación y difusión no consentidas de material de contenido sexual explícito en internet, incluido aquel creado o modificado mediante IA generativa (Swan, 2024).

La cooperación internacional es fundamental para combatir el ciberdelito transfronterizo (Chang, 2012). Las

regulaciones internacionales en materia de IA y delincuencia deben priorizar la cooperación global para enfrentar eficazmente el ciberdelito transfronterizo. Por sí solas, las regulaciones nacionales resultan insuficientes, ya que los ciberdelitos trascienden las jurisdicciones y explotan las brechas entre distintos marcos normativos. Una regulación eficaz requiere un enfoque colaborativo en el que los estados armonicen sus legislaciones en materia de ciberseguridad, compartan recursos e inteligencia, e inviertan en el fortalecimiento de capacidades, especialmente en los países en desarrollo, con el fin de reforzar la lucha global contra la delincuencia facilitada por la IA.

VI. Conclusión

A nivel mundial, el uso generalizado de la inteligencia artificial ha dado lugar a un rápido aumento de las actividades de ciberdelito. Las características propias del ciberespacio, entre ellas el anonimato, su naturaleza sin fronteras, el fácil acceso, la rapidez y la creciente dependencia tecnológica, han acelerado la expansión de los delitos facilitados por la IA. En este tipo de delitos, a menudo, las víctimas son inconscientes de su victimización, ya que los autores suelen eliminar los mecanismos de trazabilidad, manipular los activos de información objeto del ataque y, además, muchas víctimas muestran reticencia a denunciar los hechos.

Los activos intangibles, en forma de información, han adquirido hoy un alto valor, y la inteligencia artificial ha intensificado tanto su importancia estratégica como su vulnerabilidad frente a la explotación delictiva. La amplia disponibilidad de información en el ciberespacio, accesible para quienes cuentan con los medios técnicos necesarios, ha puesto de manifiesto el problema de la vulneración del dere-

cho a la privacidad, reconocido como un derecho fundamental. Si bien el uso de la IA ha incrementado sustancialmente la productividad y ampliado el acceso a la información, también ha expuesto a personas y organizaciones a una amplia gama de conductas delictivas.

Para abordar eficazmente los delitos facilitados por la IA, los gobiernos y las fuerzas del orden deben pasar de un modelo reactivo de actuación a estrategias proactivas basadas en inteligencia criminal. Ello requiere actualizar los marcos jurídicos para tipificar las conductas facilitadas por la IA, reforzar la capacidad institucional mediante la formación especializada y el fortalecimiento de la informática forense, así como desplegar sistemas de detección basados en IA para identificar *deepfakes*, contenido sintético y comportamientos anómalos.

Las respuestas eficaces también requerirán un intercambio seguro de información entre los sectores público y privado, el fortalecimiento de la cooperación transfronteriza y la adopción de salvaguardias éticas en el desarrollo de la IA para prevenir su uso indebido. Las acciones coordinadas en estos ámbitos reforzarán la resiliencia, garantizarán una aplicación uniforme de la ley y mitigarán los riesgos que plantea la delincuencia facilitada por la IA. Asimismo, resulta esencial promover programas sostenidos de alfabetización digital y en IA, a fin de dotar a personas y organizaciones de los conocimientos necesarios para reconocer, prevenir y responder a los delitos facilitados por esta última, lo que reforzaría su capacidad de protección frente a estas amenazas emergentes.

VII. Recomendaciones

La aplicación de la inteligencia artificial en el abordaje del delito debe ir más allá de la mera identificación de riesgos y orientarse al desarrollo de mecanismos de mitigación concretos y exigibles. Entre las medidas clave se encuentran la realización obligatoria de auditorías algorítmicas a cargo de órganos de supervisión independientes, la implementación de evaluaciones de impacto algorítmico antes de su puesta en funcionamiento y la adopción de estándares técnicos verificables, como parámetros de equidad y requisitos de transparencia. Estos mecanismos permitirían traducir los principios éticos en herramientas de gobernanza operativas, lo que contribuiría a prevenir resultados sesgados o injustos y a fortalecer la rendición de cuentas institucional en el despliegue de sistemas policiales asistidos por IA.

El uso creciente de drones autónomos o semiautónomos para labores de vigilancia o ataques dirigidos, así como el sabotaje de infraestructuras críticas asistido por IA, demuestra que las amenazas impulsadas por IA trascienden el ámbito de la ciberdelincuencia y exigen medidas integrales. Abordar tanto los riesgos digitales como los físicos es esencial para desarrollar estrategias eficaces de control del delito que permitan anticipar amenazas emergentes y reforzar la seguridad pública.

Los actores del sector privado desempeñan un papel central en el desarrollo, la implementación y la regulación de los sistemas de inteligencia artificial, por lo que asumen una responsabilidad ética y jurídica significativa. Los marcos regulatorios deben enfatizar la responsabilidad corporativa mediante obligaciones exigibles en materia de protección de datos, requisitos de seguridad desde el diseño y estándares de responsabilidad claramente definidos en casos de brechas de datos o implementación negligente

de sistemas. Asimismo, resulta necesaria una colaboración más estrecha entre los gobiernos y las empresas tecnológicas para mitigar de manera eficaz la actividad delictiva facilitada por la IA.

Por último, la adopción de sistemas de inteligencia artificial explicable en el ámbito de la seguridad pública puede mejorar la transparencia, la rendición de cuentas y la confianza pública, al hacer más comprensibles las decisiones algorítmicas para los agentes policiales, las autoridades judiciales y la ciudadanía. Al mismo tiempo, la implementación de incentivos económicos, como el financiamiento para mejorar la infraestructura de seguridad, las recompensas por el cumplimiento de estándares de protección de datos o las sanciones por el despliegue negligente de sistemas de IA, puede fortalecer las salvaguardas operativas y promover un uso responsable de estas tecnologías en la actividad policial.

VIII. Referencias

Brenner, Susan (2009, 26 de marzo). *Cyber Threats: The Emerging Fault Lines of the Nation State*. Reino Unido: Oxford University Press. https://academic.oup.com/book/10778

Burton, Joe, Ardi Janjeva, Simon Moseley, Alice (2025, 31 de marzo). "AI and Serious Online Crime". *Centre for Emerging Technology and Security*. London: The Alan Turing Institute. https://cetas.turing.ac.uk/publications/ai-and-serious-online-crime

Chang, Lennon Yao-Chung (2012, 30 de noviembre). *Cybercrime in the Greater China Region: Regulatory Responses and Crime Prevention Across the Taiwan Strait*. Reino Unido: Edward Elgar.

Council on Criminal Justice (2024, octubre). "The Implications of AI for Criminal Justice. Key Takeaways from a convening of leading stakeholders". *Council on Criminal Justice*. Washington, DC. https://counciloncj.org/the-implications-of-ai-for-criminal-justice/

Dakalbab, Fatima, Manar Aba Talib, Omnia Abu Waraga, Ali Bou Nassif, Sohail Abbas y Qassim Nasir (2022). "Artificial intelligence and crime prediction: A systematic literature review". *Social Sciences and Humanities Open*, 6(1). https://www.sciencedirect.com/science/article/pii/S2590291122000961

Ersöz, Filiz, Taner Ersöz, Francesco Marcelloni, Fabrizio Ruffini (2025, 24 de marzo). Artificial Intelligence in Crime Prediction: A Survey with a Focus on Explainability. *IEEE Access*, 13. https://ieeexplore.ieee.org/document/10937481

Frenkel, Omer (2025, 30 de enero). "AI and Crime: How Artificial Intelligence is Advancing Crime Prevention". *Cognyte*. https://www.cognyte.com/blog/ai-and-crime/

Friedman, Francine, Kristofer Ekdahl y Matthew Thomas (2015). "Big Data and Privacy Concerns: The Federal Policy Debate Around Consumer Data". Kumar Jayasuriya y Kathryn A. Ritcheske (eds.). *Big Data, Big Challenges in Evidence-Based Policymaking*. California: West Academic Publishing.

Hijmans, Hielke y Herke Kranenborg (2015, 21 de julio). "Data Protection Anno 2014: How to Restore Trust?" Cambridge: Intersentia, 5.

Jejelola, Folajimi (2024, 4 de diciembre). "The Role of Artificial Intelligence in the Eradication of Transnational Crime". *International Journal of Research and Innovation in Social Science*. University of Ibadan. https://rsisinternational.org/journals/ijriss/articles/the-role-of-artificial-intelligence-in-the-eradication-of-transnational-crime/

Johnson, Mark (2013). *Cyber Crime, Security and Digital Intelligence*. Farnham, Surrey: Gower.

Kalra Kush and Ayush Verma (2011). "Cyber Crime: Patterns and Control". Girish Shankar Bajpai (ed.). *On Cyber Crime and Cyber Law*. Serial Publications, pp. 170-171.

King, Thomas, Nikkita Aggarwal, Mariarosaria Taddeo y Luciano Floridi (2019, 14 de febrero). "Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions". *Science and Engineering Ethics*, 26, pp. 89-120. https://ec.europa.eu/futurium/en/system/files/ged/king2019_article_artificialintelligencecrimeani.pdf

Koehler, Thomas (2018). *Understanding Cyber Risk: Protecting Your Corporate Assets*. Routledge.

Leman-Langlois, Stéphane (2013). *Technocrime: Policing and Surveillance*. Routledge.

Marwala, Tshilidzi (2013). *Economic Modeling Using Artificial Intelligence Methods*. London: Springer Nature.

Naef, Tobias (2023). *Data Protection without Data Protectionism: The Right to Protection of Personal Data and Data Transfers in EU Law and International Trade Law*. Springer Nature.

Patishman, Henry (2025, 12 de agosto). *Global Legal Actions Against AI Deepfakes: Five Laws of 2025*. Regula. https://regulaforensics.com/blog/deepfake-regulations/

Rigano, Christopher (2019, enero). "Using Artificial Intelligence to Address Criminal Justice Needs". *National Institute of Justice Journal*, 280. https://www.ojp.gov/pdffiles1/nij/252038.pdf

Rowland, Diane, Uta Kohl y Andrew Charlesworth (2011). *Information Technology Law* (4a ed.). Routledge.

Ryder, Rodney y Nikhil Naren (2025). *Artificial Intelligence and Law*. Indian: Law and Justice.

Sadulski, Jarrod (2024, 6 de diciembre). *Artificial intelligence in crime detection: How it's useful*. Charles Town: American

Military University. https://www.amu.apus.edu/area-of-study/information-technology/resources/artificial-intelligence-in-crime-detection/

Swan, Michael (2024, 18 de octubre). The Criminal Code Amendment (Deepfake Sexual Material) Act 2024. *Policy reform to strengthen online safety in Australia*. Sydney: Carroll and O'Dea Lawyers. https://www.codea.com.au/publication/the-criminal-code-amendment-deepfake-sexual-material-act-2024-policy-reform-to-strengthen-online-safety-in-australia/

TRM Team (2025, 29 de enero). *The rise of Alenabled crime: Exploring the evolution, risks, and responses to Alpowered criminal enterprises*. TRM Labs. https://www.trmlabs.com/resources/blog/the-rise-of-ai-enabled-crime-exploring-the-evolution-risks-and-responses-to-ai-powered-criminal-enterprises?utm_source=chatgpt.com

United Nations Office on Drugs and Crime (2025, septiembre). *Emerging threats: The intersection of criminal and technological innovation in the use of automation and artificial intelligence in the cybercrime landscape of Southeast Asia*. Bangkok: Cybercrime Technical Brief Series: https://www.unodc.org/roseap/uploads/documents/Publications/2025/UNODC_Report_Emerging_threats_-_The_intersection_of_criminal_and_technological_innovation_in_the_use_of_automation_and_AI.pdf

U.S. Department of Homeland Security (2024). *Impact of Artificial Intelligence on Criminal and Illicit Activities. Homeland Security:* https://www.dhs.gov/sites/default/files/2024-10/24_0927_ia_aep-impact-ai-on-criminal-and-illicit-activities.pdf

Walden, Ian (2007). *Computer Crimes and Digital Investigations*. Oxford University Press.

Zhang, Justina (2025, 19 de marzo). "China Releases Laws to Mandate Labelling of AIGC". *Herbert Smith Freehills Kramer*. https://www.hsfkramer.com/notes/tmt/2025-posts/China-Releases-Laws-to-Mandate-Labelling-of-AIGC