2024

Inteligencia artificial y crimen organizado

AI ACTION SUMMIT

EL PACCTO 2.0
EU-LAC Partnership on justice and security

Financiado por la Unión Europea

Edición: Programa EL PACCTO 2.0
Calle Almansa 105
28040 Madrid (España)

Con la Coordinación de:

Marc Reina Tortosa
Senior Executive Manager de EL PACCTO 2.0
Emilie Breyne
Técnica de Proyecto de EL PACCTO 2.0

Autores:

Cristos Velasco, Jean Garcia Periche, Juan De Dios Gómez Gómez y Miguel Bueno Benedí

Con la revisión de:

Alfonso Peralta Gutiérrez

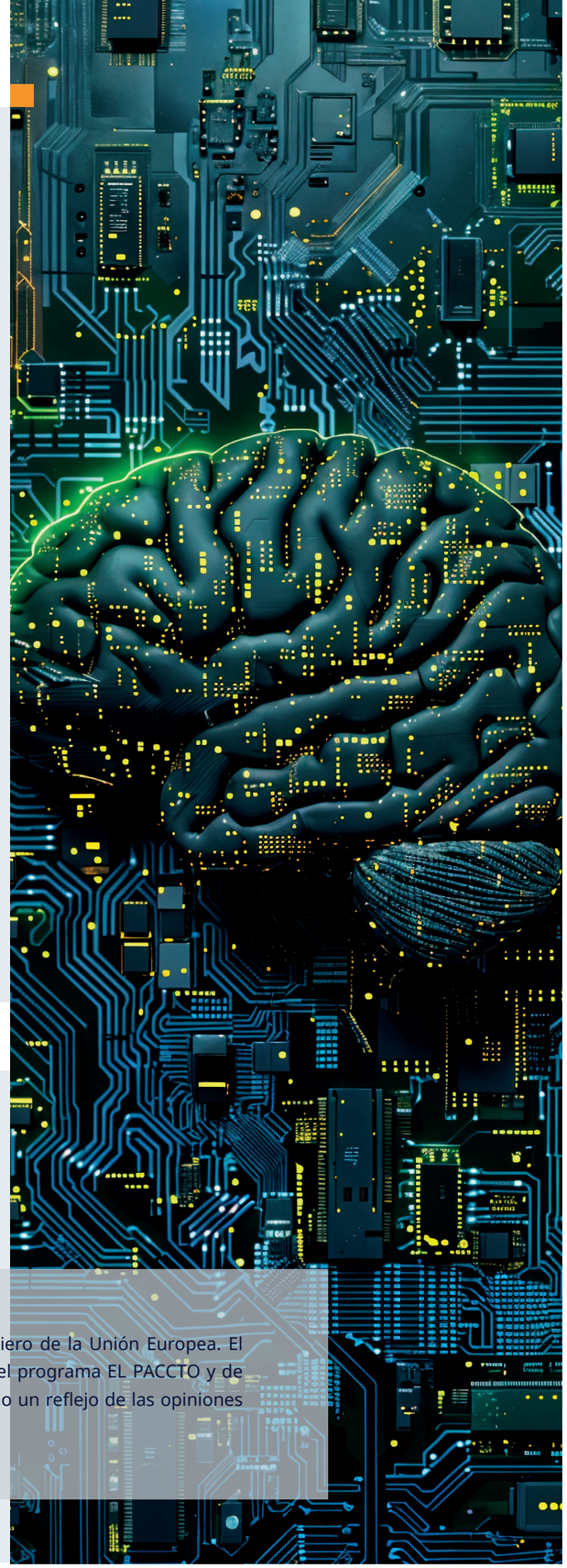
Este documento fue realizado con la contribución de las siguientes instituciones:

Expertise France

Fundación Internacional y para Iberoamérica de Administración y Políticas Públicas

Edición no venal. Madrid, noviembre de 2024

Este documento fue elaborado con el apoyo financiero de la Unión Europea. El contenido de esta publicación es responsabilidad del programa EL PACCTO y de sus autores, en ningún caso debe considerarse como un reflejo de las opiniones de la Unión Europea.



ÍNDICE

04 **ACRÓNIMOS**

06 **INTRODUCCIÓN**

08 **ANÁLISIS DE CONTEXTO Y DESAFÍOS**

- Contexto
- Tipos de IA
- Desafíos
- Retos de género
- Retos de Derechos Humanos

20 **NORMATIVA INTERNACIONAL, ESTRATEGIAS Y OTRAS INICIATIVAS NO VINCULANTES**

- Legislación y estrategias
  - Convención Marco sobre Inteligencia Artificial del Consejo de Europa
  - Reglamento Europeo sobre Inteligencia Artificial
  - Estrategias regionales y nacionales
- Principios e iniciativas internacionales no vinculantes relevantes
  - Directrices Éticas para una IA Fiable del Grupo Independiente de Expertos de Alto Nivel sobre Inteligencia Artificial (HLEG)
  - Recomendación de la OCDE sobre Inteligencia Artificial
  - Recomendación sobre Ética e Inteligencia Artificial de la UNESCO
  - Otras Iniciativas Relevantes

30 **PRINCIPALES DELITOS COMETIDOS UTILIZANDO HERRAMIENTAS DE IA**

- Fraude financiero y bancario
- Suplantación de Identidad
- Ransomware como Servicio (RasS)Phishing e Ingeniería Social
- Tráfico de personas: reclutamiento y explotación online
- Delitos de abuso y explotación sexual
- Conductas de ciberviolencia

42 **UTILIZACIÓN DE HERRAMIENTAS DE IA POR PARTE DE LAS INSTITUCIONES DE JUSTICIA Y SEGURIDAD**

- La IA en las instituciones de justicia
  - Proyectos e iniciativas de cooperación jurídica internacional
  - Gestión de casos judiciales
- La IA en las instituciones de seguridad
  - Aplicaciones de la IA a la investigación criminal
  - Proyectos e iniciativas para fortalecer la cooperación en materia de seguridad y las investigaciones de delitos
- Herramientas de IA para temáticas u áreas específicas
  - Análisis y evaluación de pruebas
  - Asistencia en la toma de decisiones y resoluciones judiciales asistidas por IA
  - Seguimiento en la ejecución de las penas impuestas en sentencia
- Herramientas de IA utilizadas en países de América Latina y el Caribe

68 **RECOMENDACIONES DE ACTUACIÓN Y CONCLUSIONES**

- Recomendaciones de actuación
- Conclusiones

74 **BIBLIOGRAFÍA**

ACRÓNIMOS

AIAB	Consejo Asesor de IA de la Comisión Europea
ALC	América Latina y el Caribe
ALPR	Reconocimiento Automático de Matrículas
ANDJE	Agencia Nacional de Defensa Jurídica del Estado de Colombia
AVENUE	Proyecto “Analysis of Video Evidence with Novel Enhanced Understanding Engine”
AVIDICUS	Proyecto “Assessment of Video-Mediated Interpreting in the Criminal Justice System“
BID	Banco Interamericano de Desarrollo
BKA	Bundeskriminalamt
Blockchain	Tecnología de cadena de bloques
CaaS	Crime-as-a-Service
CAF	Banco de Desarrollo de América Latina y el Caribe
CAI	Comité de Inteligencia Artificial del Consejo de Europa
CEF	Connecting Europe Facility
CENIA	Centro Nacional de Inteligencia Artificial de Chile
CEPAL	Comisión Económica para América Latina y el Caribe
CEPEJ	Comisión Europea para la Eficacia de la Justicia
CJI	Cumbre Judicial Iberoamericana
CJNG	Cártel Jalisco Nueva Generación
C4	Centro de Comando, Control, Comunicaciones y Cómputo de Colombia
DDOS	Ataques de denegación de servicio
EE.UU.	Estados Unidos de América
EL PACCTO 2.0	Europa, Latinoamérica y el Caribe, Programa de Asistencia contra el Crimen Transnacional Organizado
Eurojust	Agencia de Cooperación de la Unión Europea en materia de Justicia Penal
Europol	Agencia de la Unión Europea de Cooperación Policial
Frauke	Proyecto “Fraud Analysis Using Knowledge Extraction”
FRICoRe	Proyecto “Fundamental Rights in Courts and Regulation”
GDO	Grupos de delincuencia organizada
HRCN	Redes Criminales de Alto Riesgo
IA	Inteligencia Artificial
IAG	Inteligencia Artificial Generativa
iBorderCtrl	Proyecto para mejorar el control fronterizo mediante el uso de tecnologías avanzadas
IC	Investigación criminal
iCOP	Proyecto “Identifying and Catching Online Predators”
ILIA	Índice Latinoamericano de Inteligencia Artificial
INSPECTr	Proyecto “Intelligence Network and Secure Platform for Evidence Correlation and Transfer”
Interpol	Organización Internacional de Policía Criminal
JuLIA	Proyecto “Justice, fundamentaL rIghts and Artificial Intelligence Applications”
MARCELL	Proyecto “Multilingual Resources for CEF.AT in the Legal Domain”
ML	Aprendizaje Automático
NCMEC	Centro Nacional para Niños Desaparecidos y Explotados
OCDE	Organización para la Cooperación y el Desarrollo Económico
OEA	Organización de los Estados Americanos
OLGA	Proyecto “Online-Strafverfahrensregister für Organisierte Kriminalität und Geldwäsche”
ONU	Organización de las Naciones Unidas
PCC	Primeiro Comando da Capital
PLN	Procesamiento de Lenguaje Natural
PRISMA	Perfil de Riesgo de Reincidencia para Solicitud de Medidas de Aseguramiento
RasS	Ransomware como Servicio
REIA	Reglamento Europeo sobre Inteligencia Artificial
ROXANNE	Proyecto “Real-time network, text, and speech analytics for combating organized crime and terrorism”
SCJN	Suprema Corte de Justicia de la Nación de México
STOA	Panel para el Futuro de la Ciencia y la Tecnología del Parlamento Europeo
TAJ	Traitement d’Antécédents Judiciaires
TENSOR	Proyecto “Retrieval and analysis of heterogeneous data for predicting and mitigating violent actions”
UE	Unión Europea
UNESCO	Organización de la Naciones Unidas para la Educación, la Ciencia y la Cultura
UNICRI	Centro de Inteligencia Artificial y Robótica del Instituto Interregional de las Naciones Unidas para Investigaciones sobre la Delincuencia y la Justicia
VioGén	Sistema de Seguimiento Integral en los casos de Violencia de Género
VMI	Interpretación mediada por vídeo

INTRODUCCIÓN

La adopción de sistemas y tecnologías basadas en IA ha registrado un enorme crecimiento desde la introducción de ChatGPT en noviembre de 2022. El uso de sistemas de IA resulta sumamente útil en diversas áreas de la economía tales como educación, logística y transporte y en general en la prestación de servicios digitales para el ciudadano. La integración de sistemas de IA en la labor y actividades de las fuerzas de seguridad y las autoridades del sistema de justicia incluido el poder judicial está teniendo una gran trascendencia, puesto que ofrece herramientas versátiles y poderosas para llevar acabo el análisis de grandes cantidades de datos, la identificación de patrones y conductas delictivas e inclusive cuestiones más avanzadas relacionadas con la vigilancia predictiva y la predicción de sentencias en contra de sujetos imputados de algún delito, lo que permite a las autoridades de justicia penal, por un lado, optimizar la labor y eficacia de sus investigaciones ofreciendo tiempos de respuesta más rápidos, y por otro lado, la posibilidad de asignar y administrar recursos humanos y financieros de manera más eficiente.

No obstante, los sistemas de IA también están siendo utilizados y explotados en beneficio del crimen organizado sin la necesidad de contar con habilidades o destrezas técnicas avanzadas. Los delincuentes pueden utilizar la IA para facilitar y potenciar sus ataques maximizando las oportunidades de obtener beneficios en un menor tiempo, explotando nuevas víctimas y creando modelos de negocio delictivos más innovadores, al tiempo que reducen las posibilidades de ser atrapados. Por ejemplo, la automatización de ataques mediante el uso de bots controlados por IA que tienen la capacidad de potenciar sus actividades a gran escala tales como ataques de denegación de servicio (DDOS), la propagación y distribución de programas maliciosos (malware), la identificación de imágenes en redes sociales para explotar a víctimas de tráfico de personas o de delitos de abuso y explotación sexual, la creación de correos phishing con mayor precisión y con la posibilidad de adaptarlos a las características específicas de sus víctimas, automatización de falsedades documentales, violaciones masivas de la propiedad intelectual e industrial, utilizar técnicas manipulativas para persuadir a personas a realizar conductas no deseadas, incluso intentar influir en un proceso electoral o engañar al juez con pruebas falsas en un juicio. o el uso de deepfakes para suplantar la identidad de sus víctimas para posteriormente cometer delitos de extorsión y fraude. Todas estas conductas y vectores presentan retos y oportunidades que las autoridades del sistema de justicia deben abordar y explorar para poder ofrecer respuestas concretas a las víctimas de este tipo de delitos.

Bajo este contexto, el presente estudio tiene como objetivo proporcionar una visión completa de las interacciones entre la IA y el crimen organizado en América Latina, el Caribe y la Unión Europea, dando ejemplos de acciones realizadas, amenazas o casos públicos, así como herramientas de trabajo existentes y destacando tanto los peligros como las oportunidades que presenta esta tecnología. Además, el estudio explora estas dinámicas, proporcionando una visión integral de cómo la IA está siendo utilizada tanto por los grupos criminales como por las fuerzas del orden en América Latina, y las implicaciones sociales y éticas de esta evolución tecnológica en el ámbito del crimen organizado.

En particular, el estudio busca: (i) identificar los principales desarrollos legislativos y de políticas en materia de IA actualmente desarrollados por organismos internacionales y regionales, el alcance de la legislación existente vigente en la UE, así como una revisión del estado actual de las estrategias nacionales sobre IA desarrolladas principalmente en países de ALC; (ii) proporcionar un análisis de las principales conductas y delitos cometidos a través del uso de herramientas de IA; (iii) proporcionar un mapeo y análisis de las herramientas de IA actualmente utilizadas por las instituciones del sector justicia y de seguridad en la UE y en países de ALC para fortalecer la administración de justicia y el combate al crimen organizado; y (iv) proponer acciones y actividades específicas para abordar el vínculo entre la IA y el crimen para que puedan ser implementadas y desarrolladas por los países del Programa de la Unión Europea EL PACCTO 2.0 (Europa, Latinoamérica y el Caribe, Programa de Asistencia contra el Crimen Transnacional Organizado) tomando en cuenta los retos de género y la protección de los derechos fundamentales de las partes en el proceso penal.

La información que se aporta es una fotografía del momento actual que debe ser interpretada como tal ya que la tecnología de IA evoluciona muy rápidamente y algunas herramientas, sistemas o aplicaciones desarrolladas y mencionadas en el presente estudio pueden quedar obsoletas en cuestión de pocos años o incluso meses. Además, las tendencias criminales varían mucho de país en país e incluso entre regiones y sub-regiones. No obstante, se puede esperar que el uso de herramientas de IA por parte de los grupos de delincuencia organizada (GDO), particularmente de las redes criminales de alto riesgo (HRCN, por sus siglas en inglés), vaya en aumento, incluso en algunos casos su uso aumente exponencialmente en los próximos años.

BLOQUE 1: ANÁLISIS DE CONTEXTO Y DESAFÍOS

1.1. CONTEXTO

La inteligencia artificial (IA) ha transformado múltiples sectores de la sociedad global, proporcionando nuevas oportunidades para el desarrollo económico y social, así como desafíos importantes, especialmente en la lucha contra el crimen organizado. En América Latina, donde el crimen organizado ha adoptado tradicionalmente actividades como el narcotráfico y el tráfico de armas, la IA ha ampliado el alcance y la sofisticación de las operaciones criminales. Esta adopción tecnológica plantea nuevas dificultades para las fuerzas del orden, que ya enfrentan recursos limitados y estructuras institucionales frágiles en varios países de la región.

La IA ha introducido nuevas modalidades de delitos que van más allá de las prácticas tradicionales. Por ejemplo, Europol ha señalado el modelo de Crime-as-a-Service (CaaS), el cual permite que criminales sin experiencia técnica accedan a herramientas sofisticadas en el submundo digital, aumentando su capacidad para realizar ataques complejos[1]. Esto ha facilitado que tecnologías emergentes, como la IA, se conviertan en un motor clave para actividades delictivas. Asimismo, en Sudáfrica, los delitos de suplantación de identidad aumentaron en un 284% entre 2021 y 2022, impulsados por el uso de IA para la creación de identidades falsas y fraudes financieros[2].

Un ejemplo particular de fraude sofisticado vinculado al uso de IA ocurrió en China, donde un individuo fue engañado para transferir cerca de 500.000 USD a un estafador que utilizó tecnología de intercambio de rostros y de imitación de voz para hacerse pasar por un amigo cercano. Aunque las autoridades lograron detener parte del dinero, el incidente demostró cómo la IA está siendo empleada para llevar a cabo fraudes financieros de alto nivel[3].

En América Latina, los grupos del crimen organizado han aprovechado la IA en diversas formas, integrando herramientas avanzadas para aumentar la efectividad de sus operaciones. Los carteles de la droga en México, por ejemplo, han empezado a utilizar drones controlados por IA no solo para transportar drogas, sino también para realizar ataques físicos contra miembros de otros carteles o para interrumpir

1 Europol. (2017). Europol. European Union Serious and Organised Crime Threat Assessment (SOCTA). Retrieved from https://www.europol. europa.eu/activities-services/main- reports/european-union-serious-and-organised-crime-threat-assessment-2017.

2 ENACT, (2023) AI and organised crime in Africa. Sigsworth, R., ENACT Observer (2023). Retrieved from https://enactafrica.org/enact-observer/ ai-and-organised-crime-in-africa.

3 UNODC, (2024). Casino Underground Banking Report 2024. UNODC Publications (2024). Retrieved from https://www.unodc.org/roseap/uploads/documents/Publications/ 2024/ Casino_Underground_Banking_Report_2024.pdf

cadenas de suministro rivales[4]. Estos drones se han convertido en una herramienta fundamental para el contrabando, permitiendo el cruce de fronteras con mayor facilidad y evitando los controles de seguridad convencionales. En muchos casos, estos drones operan de manera autónoma, por lo que se les ha denominado “las nuevas mulas de drogas”[5].

La IA también está transformando la forma en que los carteles optimizan sus rutas de contrabando. Mediante modelos de aprendizaje automático, los carteles pueden analizar datos de rutas, horarios de patrullas fronterizas y métodos de envío para prever los mejores momentos y ubicaciones para transportar drogas, reduciendo el riesgo de detección. Además, el uso de sistemas de reconocimiento facial ha permitido a los carteles identificar a agentes encubiertos, lo que complica aún más los esfuerzos de las autoridades para infiltrarse en estas organizaciones[6]. En 2018, un dron armado fue utilizado para atacar la casa del secretario de Seguridad Pública del estado de Baja California, Gerardo Sosa Olachea, en la ciudad de Tecate, a lo largo de la frontera entre Estados Unidos y México. Parece que se emplearon al menos dos drones en el ataque. El primero llevaba equipos de audio y video, además de dos artefactos explosivos improvisados (IED) que no detonaron tras caer en el patio del funcionario, mientras el segundo realizaba la vigilancia[7].

Este posible uso de robot de acción remota o sistemas de armas letales autónomos (LAWS) puede suponer en supuestos de terrorismo y narcoterrorismo no sólo efectos y daños indiscriminados al no tener un control humano significativo con consecuencias desastrosas para la seguridad y la vida humana, sino también permitir potencialmente que miembros de organizaciones criminales y terroristas operen a distancia del entorno donde se desenvuelve una acción, y proporcionarles una mejor protección personal en un escenario de guerra, de explosiones así como por ejemplo en entornos contaminados por agentes químicos o escenarios catastróficos[8].

4 ENACT, (2023) AI and organised crime in Africa. Sigsworth, R., ENACT Observer (2023). Retrieved from https://enactafrica.org/enact-observer/ai-and-organised-crime-in-africa.

5 idem

6 National School of Political and Administrative Studies. Artificial Intelligence – a Double-Edged Sword. Organized Crime`s AI vs Law Enforcement`s AI

7 ASMANN, P. (2018, Agosto 15). Are armed drones the weapon of the future for Mexico’s cartels? InSight Crime. https://insightcrime.org/news/brief/armed-drones-weapon-future-mexico-cartels/

8 LAMAS LOPEZ, F., & PERALTA GUTIERREZ, A. (2023). Marco de Derecho Internacional Público y Usos Militares de la Inteligencia Artificial en la UE. Revista Electrónica De Estudios Internacionales, (46), 505–525. https://doi.org/10.36151/reei.46.17

Otras formas en que los grupos criminales utilizan la IA incluyen la explotación de imágenes satelitales de plataformas como *Google Earth* para planificar rutas de contrabando con precisión, monitoreo a tiempo real de los movimientos de las fuerzas y cuerpos de seguridad y la manipulación de redes sociales a través de minería de datos y generación automática de contenido, herramientas que les permiten gestionar sus operaciones de manera más eficiente[9]. De manera similar a las empresas legítimas, las organizaciones criminales emplean la IA para la gestión de la cadena de suministro y para la mitigación de riesgos[10].

La capacidad de la IA para predecir vulnerabilidades también ha sido utilizada en esquemas de chantaje, extorsión y difamación. Los crímenes tradicionales como la extorsión y el terrorismo ahora se llevan a cabo con nuevos métodos, incluyendo la generación automática de contenido, lo que facilita la creación de pornografía infantil y la manipulación de redes sociales para influir en la opinión pública y defender los intereses del crimen organizado[11]. Los criminales también han comenzado a desarrollar su propio software de IA, empleando modelos sin restricciones para generar malware como el ransomware, que cifra datos y exige rescates[12].

Sin embargo, las fuerzas del orden en América Latina también están aprovechando la IA como una herramienta en su lucha contra el crimen organizado. Tecnologías como el análisis predictivo, el reconocimiento de patrones y el Reconocimiento Automático de Matrículas (ALPR) ayudan a las agencias de seguridad a procesar grandes cantidades de datos, como registros financieros, imágenes de vigilancia y datos de redes sociales, lo que mejora su capacidad para identificar y rastrear redes criminales[13]. A pesar de estos avances, la capacidad para implementar estas tecnologías de manera efectiva sigue siendo limitada en muchos países de la región, lo que subraya la necesidad de fortalecer las instituciones y los recursos para enfrentar el creciente impacto de la IA en el crimen organizado.

A medida que las organizaciones criminales en América Latina continúan adoptando la IA, las formas tradicionales de delitos como el tráfico de drogas, el contrabando de armas y el lavado de dinero se están transformando. El uso de IA no solo permite a estos grupos operar con mayor sigilo, sino que también les otorga una capacidad para escalar sus operaciones de manera exponencial. Uno de los ejemplos más notables, como hemos comentado, es el uso de drones controlados por IA para la vigilancia y el transporte de contrabando a través de las fronteras, algo que permite a los criminales operar con una autonomía sin precedentes y evitar los métodos tradicionales de control y vigilancia fronteriza. Estos avances no solo incrementan la eficiencia de las organizaciones criminales, sino que también dificultan la tarea de las fuerzas del orden.

Por otro lado, la optimización de fraudes financieros también ha sido facilitada por la IA. Casos como los deepfakes y la automatización de estafas mediante chatbots han permitido a los criminales ampliar el alcance de sus actividades. Estos chatbots, como el LoveGPT, son capaces de generar conversaciones automáticas en aplicaciones de citas para estafar emocionalmente a las víctimas, pidiendo dinero bajo falsas emergencias o presentando oportunidades de inversión fraudulentas. Este tipo de fraude demuestra la capacidad de la IA no solo para automatizar procesos, sino para hacerlo a una escala que hubiera sido imposible sin la intervención de la tecnología. Además, algunos de los grupos criminales más poderosos de América Latina, como el Cártel Jalisco Nueva Generación (CJNG) en México y el Primeiro Comando da Capital (PCC) en Brasil, están involucrados en este tipo de fraudes, lo que sugiere que la tecnología no está limitada a grupos pequeños, sino que también es empleada por grandes organizaciones delictivas.

Otro aspecto clave es el phishing, una técnica que ha evolucionado considerablemente gracias a los avances en IA. Anteriormente, el phishing se limitaba a correos electrónicos sencillos y fácilmente identificables. Sin embargo, hoy en día los delincuentes pueden personalizar sus mensajes utilizando modelos de lenguaje que imitan a personas o instituciones de confianza, incrementando la tasa de éxito de estos fraudes. En Brasil, un grupo de ciberdelincuentes llamado PINEAPPLE ha utilizado estas técnicas para enviar correos que imitan al servicio de impuestos federales, logrando engañar a víctimas para que descarguen malware al intentar acceder a documentos falsos (ENACT, 2023). Esta técnica no solo representa un reto para las fuerzas del orden, sino que también revela cómo la IA está refinando y profesionalizando los métodos delictivos.

Además, la creación de malware ha sido facilitada por la IA, lo que ha llevado a un incremento en la sofisticación de los ataques cibernéticos. Aunque los modelos de IA más populares, como ChatGPT, están diseñados para rechazar solicitudes maliciosas, hay soluciones como WormGPT que se anuncian sin ningún tipo de restricción, y por supuesto sin que la empresa asuma ningún tipo de responsabilidad por usos ilícitos, utilizando IA para desarrollar software dañino o adaptar herramientas para su uso delictivo. Esto supone una "democratización" de los ciberataques donde los cibercriminales ya no necesitan grandes costes, ni conocimiento técnico ni tiempo, es barato, fácil, sencillo, rápido y automatizado[14]. El malware bancario en Brasil es un ejemplo de cómo los criminales están empleando estas herramientas para robar información confidencial de manera casi idéntica a cómo lo haría un software legítimo, lo que complica aún más la labor de las fuerzas del orden.

Frente a estos desafíos, las autoridades en América Latina están comenzando a integrar la IA en sus operaciones, aunque el progreso ha sido desigual. En países con mayores recursos, como Brasil y México, se han implementado tecnologías como el Reconocimiento Automático de Matrículas (ALPR) y herramientas de análisis predictivo que permiten rastrear vehículos sospechosos y anticipar actividades criminales. Sin embargo, en muchas naciones de la región, la falta de infraestructura adecuada y la escasez de recursos limitan la capacidad para adoptar estas tecnologías de manera efectiva.

A pesar de los avances tecnológicos disponibles, la lucha contra el crimen organizado que utiliza IA en América Latina enfrenta desafíos estructurales y éticos. Los marcos legales que regulan el uso de estas tecnologías no están completamente desarrollados en muchos países, y la colaboración internacional es esencial para combatir las redes criminales transnacionales que operan a través de fronteras. Además, la vigilancia predictiva y otras herramientas de IA plantean preocupaciones sobre la privacidad y los derechos humanos, lo que requiere un equilibrio cuidadoso entre la seguridad pública y las libertades individuales[15].

Desde el uso de deepfakes y drones hasta la optimización de fraudes y la creación de malware, la IA ha permitido a las organizaciones criminales evolucionar y aumentar su alcance de maneras sin precedentes. Sin embargo, también ofrece herramientas poderosas para las fuerzas del orden, que, si se implementan de manera adecuada, podrían nivelar el campo de batalla y mejorar significativamente la capacidad de las autoridades para combatir el crimen organizado.

INICIATIVAS DE ÁMBITO REGIONAL

La Alianza Digital Unión Europea – América Latina y el Caribe[16] fue creada y lanzada en Colombia en marzo de 2023 y consiste en un marco de cooperación informal basado en valores compartidos, abierto a todos los países de ALC y a los Estados miembros de la UE quienes podrán participar a través de sus respectivos gobiernos y agencias relacionadas con la agenda digital. El objetivo de la Alianza es fomentar el desarrollo de infraestructuras digitales seguras, resilientes y centradas en el ser humano sobre la base de un marco basado en valores, garantizando un entorno democrático y transparente y haciendo hincapié en la privacidad y los derechos digitales, y en particular busca promover la cooperación en una amplia gama de temas, entre los que se incluyen el diálogo sobre políticas digitales y la IA entre la UE y los países de ALC. Como parte de esta Alianza, se acordó la creación de dos plataformas de coordinación, el Centro D4D (Digital for Development)[17] para los socios y participantes europeos, y la Comisión Económica de las Naciones Unidas para América Latina y el Caribe (CEPAL) que será el órgano coordinador para los socios de los países de ALC.

9 Europol, (2020). Malicious Uses and Abuses of Artificial Intelligence, Trend Micro Research, European Union Agency for Law Enforcement Cooperation 2020. Retrieved from https://www.europol.europa.eu/sites/default/files/documents/malicious_uses_and_abuses_ of_artificial_intelligence_europol.pdf.

10 ENACT, (2023) AI and organised crime in Africa. Sigsworth, R., ENACT Observer (2023). Retrieved from https://enactafrica.org/enact-observer/ai-and-organised-crime-in-africa.

11 Caldwell, M., Andrews, J.T.A., Tanay, T. et al. (2020) AI-enabled future crime. Crime Sci 9, 14 (2020). Retrieved from https://doi.org/10.1186/s40163-020-00123-8

12 idem

13 AIplusInfo, (2023). How Will Artificial Intelligence Affect Policing and Law Enforcement? Artificial Intelligence + (2023). Retrieved from https://www.aiplusinfo.com/blog /artificial-intelligence-ai-and-policing

14 MARTIN, Nacho. El Independiente. (2024, November 9). WormGPT: El ChatGPT sin restricciones que usan los ciberdelincuentes. El Independiente. https://www.elindependiente.com/futuro/inteligencia-artificial/2024/11/09/wormgpt-el-chatgpt-sin-restricciones-que-usan-los-ciberdelincuentes/

15 Deloitte, (2021). Surveillance and Predictive Policing Through AI. Study Overview by Deloitte Insights (2021). Retrieved from https://www.deloitte.com/global/en/Industries/government-public/perspectives/urban-future-with-a-purpose/surveillance-and-predictive- policing-through-ai.html

16 Comisión Europea, "Global Gateway: los socios de la UE, América Latina y el Caribe ponen en marcha en Colombia la Alianza Digital UE-ALC", comunicado de prensa 14 de marzo de 2023, en: https://ec.europa.eu/commission/presscorner/detail/es/ip_23_1598

17 Ver: https://d4dhub.eu/

El Banco Interamericano de Desarrollo (BID) cuenta con una iniciativa conocida como "fAIr LAC+"[18] que es una iniciativa conformada por empresas del sector privado, grupos del sector académico, agencias gubernamentales y organismos internacionales especializados cuyo propósito es promover el uso ético y responsable de la IA en países de ALC. A través de esta iniciativa, se han creado un conjunto de guías, entre ellas destaca una guía de autoevaluación ética para el sector público que fue diseñada para permitir la mitigación de riesgos éticos asociados al uso y aplicación de nuevas tecnologías dirigida a instituciones y organizaciones del sector público[19]. fAIr LAC+ ha generado asociaciones académicas y proyectos como GuIA[20] en donde se abordan discusiones relacionadas con la ética, los principios, normas y políticas de los sistemas de IA, y problemas específicos en ALC.

La Comisión Económica para América Latina y el Caribe (CEPAL), el Centro Nacional de Inteligencia Artificial de Chile (CENIA) junto con el apoyo del Banco Interamericano de Desarrollo (BID), el Banco de Desarrollo de América Latina y el Caribe (CAF) y la Organización de los Estados Americanos (OEA) crearon en 2023 un Índice Latinoamericano de Inteligencia Artificial (ILIA)[21] cuyo propósito es evaluar un conjunto de elementos relacionados con la infraestructura, capital humano, disponibilidad de datos, regulaciones, áreas estratégicas y participación ciudadana para ofrecer indicadores y métricas sobre el nivel de desarrollo en la adopción de IA en ALC e identificar los principales desafíos en la región. El índice comprende tres dimensiones: (i) factores habilitantes que incluye infraestructura, datos y desarrollo de talento; (ii) investigación, adopción y desarrollo que incluye tres subdimensiones: Investigación, Innovación y Desarrollo y Adopción; y, (iii) gobernanza que incluye tres subdimensiones: visión e institucionalidad, internacional, y regulación[22].

El lanzamiento de la segunda versión del ILIA fue presentado por la CEPAL y CENIA el 24 de septiembre de 2024. El índice incluye a 19 países y dentro de los tres países que obtuvieron la mejor puntuación destacan Chile (73,07), seguido de Brasil (69,30) y Uruguay (64,98). De acuerdo con la CEPAL, Chile, Brasil y Uruguay no solo han avanzado en la implementación de tecnologías basadas en la IA, sino que también están orientando sus estrategias nacionales hacia la consolidación y expansión de estas tecnologías en todos los sectores de su economía y sociedad y, además cuentan, con un entorno favorable que potencia la investigación, el desarrollo y la adopción de tecnologías, promoviendo la innovación y aplicación de IA[23].

La Segunda Cumbre Ministerial Latinoamericana y del Caribe por la Inteligencia Artificial se llevó a cabo en Cartagena, Colombia el 8 y 9 de agosto de 2024. En esta cumbre participaron representantes de alto nivel de 16 países, así como representantes de la Unión Europea. Durante esta cumbre, se acordó una Declaración para la Gobernanza de Ecosistemas de Inteligencia Artificial y el Fomento a la Educación en IA de Manera Ética y Responsable en América Latina y el Caribe[24].

18 Ver: https://fairlac.iadb.org/

19 La guía de autoevaluación ética para el sector público del BID se encuentra en: https://view.genially.com/62aa57549a8ebc001038afe0

20 Ver: https://proyectoguia.lat/

21 El portal del Índice Latinoamericano de Inteligencia Artificial (ILIA) se encuentra en: https://indicelatam.cl/

22 CEPAL/CENIA, «Índice Latinoamericano de Inteligencia Artificial (ILIA)», Biblioteca del Congreso Nacional de Chile. Departamento de Estudios, Extensión y Publicaciones, 7 de agosto de 2023 en: https://obtienearchivo.bcn.cl/obtienearchivo?id=repositorio/10221/34598/1/Indice_Latinoamericano_de_Inteligencia_Artificial.pdf

23 CEPAL, «Índice Latinoamericano de Inteligencia Artificial (ILIA) mantiene a Chile, Brasil y Uruguay como líderes en la región». Comunicado de prensa de 24 de septiembre de 2024 en: https://www.cepal.org/es/comunicados/indice-latinoamericano-inteligencia-artificial-ilia-mantiene-chile-brasil-uruguay-como

24 El portal de la Segunda Cumbre Ministerial Latinoamericana y del Caribe por la Inteligencia Artificial, incluida la declaración y las memorias se encuentran en: https://mintic.gov.co/cumbre-ia/secciones/Cumbre-Ministerial-de-IA/Memorias/

SENTENCIAS JUDICIALES RELEVANTES

En el ámbito judicial destaca una reciente sentencia de la Corte Constitucional de Colombia en un caso presentado por una madre de familia que solicitó la exoneración del cobro de pagos y cuotas por los servicios y medicamentos del tratamiento de su hijo con trastorno autista. Un juez de segunda instancia es ese país empleó *ChatGPT* para formular interrogantes jurídicos sobre el derecho fundamental a la salud de menores de edad diagnosticados con trastorno de espectro autista incorporando las preguntas y respuestas en la motivación de su sentencia.

La Sala de Revisión de la Corte Constitucional concluyó que no hubo un remplazo de la función judicial por parte de *ChatGPT,* debido a que el juez de segunda instancia utilizó la IA luego de haber fundamentado y tomado la decisión. Sin embargo, la Corte Constitucional en su sentencia advirtió sobre las implicaciones de utilizar IAG y la importancia de evaluar el uso adecuado de herramientas como ChatGPT y recomendó la aplicación de criterios éticos y de respeto a los mandatos superiores para garantizar los derechos fundamentales de los ciudadanos. Exhorto a los funcionarios y empleados de la Rama Judicial a aplicar los principios de (i) transparencia, (ii) responsabilidad, (iii) privacidad, (iv) no sustitución de la racionalidad humana, (v) seriedad y verificación, (vi) prevención de riesgos, (vii) igualdad y equidad, (viii) control humano, (ix) regulación ética, (x) adecuación a buenas prácticas y estándares colectivos, (xi) seguimiento continuo y adaptación e (xii) idoneidad y ordenó al Consejo Superior de la Judicatura adoptar una guía, en relación con la implementación de la IAG en la Rama Judicial, especialmente en cuanto al uso de la herramienta *ChatGPT*[25].

25 Sentencia, T-323 de 2024. Corte Constitucional República de Colombia. Sala Segunda de Revisión, en: https://www.corteconstitucional.gov.co/relatoria/2024/T-323-24.htm

1.2. TIPOS DE IA

Existen diferentes tipos y clasificaciones relacionadas con el concepto de IA que han sido desarrolladas por la comunidad científica. Entre las clasificaciones relevantes de IA para las autoridades de justicia penal destacan las siguientes:

- **Procesamiento de Lenguaje Natural (PLN).** Es un subcampo de la informática y la IA que utiliza el aprendizaje automático para permitir que los sistemas comprendan y se comuniquen a través del lenguaje humano. El PLN permite que los sistemas y los dispositivos digitales reconozcan, comprendan y generen texto y voz mediante la combinación de la lingüística computacional con base en modelos estadísticos, el aprendizaje automático y el aprendizaje profundo.

- **Aprendizaje Automático (ML).** El aprendizaje automático se basa en modelos matemáticos entrenados con datos que aprenden con base en las experiencias. Mediante el aprendizaje automático, los algoritmos pueden hacer predicciones o tomar decisiones sin la necesidad de tener que ser programados. Existen tres subcategorías de algoritmos de ML que son esenciales para las investigaciones: (i) aprendizaje supervisado, (ii) aprendizaje no supervisado y (iii) aprendizaje reforzado.

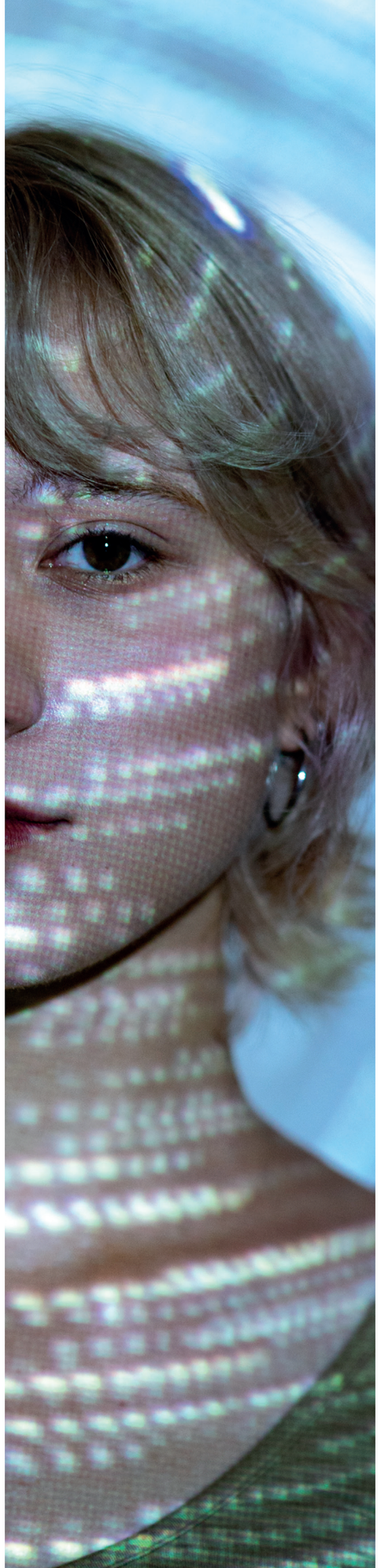
- **Aprendizaje Profundo.** El aprendizaje profundo es un subconjunto de métodos de aprendizaje automático basados en redes neuronales con aprendizaje de representación. El aprendizaje profundo genera que muchas aplicaciones y servicios mejoren la automatización y permite realizar tareas analíticas sin la intervención humana.

- **Reconocimiento Facial.** El reconocimiento facial es una categoría de tecnología biométrica que analiza los rasgos faciales de una persona para analizar y confirmar su identidad. El reconocimiento facial se utiliza comúnmente en sistemas de seguridad para identificar a sospechosos y para identificar a víctimas del delito. Esta tecnología compara imágenes digitales o fotogramas de vídeo con huellas faciales previamente almacenadas, lo que permite la identificación rápida de personas en espacios públicos e incluso en lugares con grandes aglomeraciones.

- **IA generativa.** La IA generativa o inteligencia artificial generativa hace referencia al uso de la IA para crear contenido, como texto, imágenes, música, audio y vídeos. La IA generativa se basa en modelos fundacionales, es decir, grandes modelos de IA, que pueden realizar varias tareas a la vez y llevar a cabo tareas pre configuradas, como resúmenes, preguntas y respuestas, clasificación, etc. Además, al requerir una preparación mínima, los modelos básicos se pueden adaptar a casos prácticos concretos con muy pocos datos de ejemplo.

1.3. DESAFÍOS

La IA ofrece grandes oportunidades para mejorar y optimizar la administración del sistema de justicia pero al mismo tiempo también presenta grandes desafíos técnicos, jurídicos y de cooperación internacional que los países tendrán que resolver a través de la creación de estrategias nacionales sobre IA, creación de políticas públicas diseñadas para el sector público y privado, actualización de marcos jurídicos nacionales para penalizar conductas cometidas a través de sistemas de IA, formación y capacitación de los operadores del sistema de justicia en el manejo de las tecnologías y especialmente la creación de alianzas público-privadas para que el uso de estas tecnologías pueda ser mejor entendido, tenga una mayor impacto y puedan resolverse problemas que se presenten en el uso e implementación de dichas tecnologías, -tales como sesgos, discriminación de determinados grupos de la sociedad, desinformación-, de forma conjunta entre las autoridades reguladoras nacionales, las autoridades del sistema de justicia y los desarrolladores y proveedores de tecnologías de IA con el apoyo y expertise de organismos internacionales y regionales, grupos académicos y de la sociedad civil especializados en la materia.



1.4. RETOS DE GÉNERO

El análisis de los problemas de la IA desde una perspectiva de género también es crucial, ya que las tecnologías pueden reproducir, amplificar o incluso crear nuevas formas de discriminación y violencia de género.

La IA se basa en grandes volúmenes de datos para entrenar sus modelos predictivos y si estos datos contienen sesgos de género o reflejan desigualdades preexistentes, los sistemas de IA pueden perpetuar estos patrones, afectando desproporcionadamente a las mujeres y a personas no conformes con el género tradicional. El uso de sistemas de inteligencia artificial utilizados en el empleo, la gestión de trabajadores y el acceso al trabajo por cuenta propia, en particular para el reclutamiento y la selección de personas, para la toma de decisiones que afecten a los términos de la relación laboral, la promoción y la terminación de relaciones contractuales relacionadas con el trabajo que causen una discriminación grave en el empleo, público o privado, así como servicios privados esenciales y servicios y prestaciones públicos esenciales puede provocar discriminaciones y sesgos contra personas por razón de su ideología, religión o creencias, su situación familiar, su etnia, raza o nación, origen nacional, sexo, edad, orientación o identidad sexual o de género, razones de género, aporofobia o exclusión social, enfermedad o discapacidad.

En el contexto del crimen organizado, esto puede llevar a la subrepresentación de los impactos diferenciales que sufren las mujeres, como la trata de personas, la violencia de género y la explotación sexual.

En relación con esto, un estudio de la ONU sugiere que los algoritmos de IA aplicados en áreas como la seguridad y la justicia tienden a estar entrenados con datos que no representan adecuadamente a mujeres y grupos vulnerables. Por ejemplo, en los sistemas de reconocimiento facial, se ha demostrado que los algoritmos tienen tasas de error más altas para las mujeres y especialmente para mujeres negras, lo que puede afectar la identificación correcta de víctimas y perpetradores en el contexto del crimen organizado. En concreto, un informe de AI Now Institute introdujo la idea que los sistemas de reconocimiento facial desarrollados en EE. UU tienen una tasa de error significativamente mayor para mujeres afroamericanas en comparación con hombres caucásicos, lo que podría traducirse en problemas de identificación errónea o en la falta de identificación de víctimas y criminales en redes de explotación sexual o tráfico de personas[26]. Es el ejemplo del servicio de análisis facial de Amazon, Rekognition, que ya demostró sesgos de género y raza peores que los de herramientas comparables, sesgos que tomaban la forma de literalmente 'no ver' a mujeres de piel oscura, mientras que era más competente para detectar a hombres de piel clara.

Y no sólo en reconocimiento facial, según una auditoría realizada por investigadores independientes de la Universidad del Sur de California (USC, en EE. UU.), en 2021 revela que el sistema de publicación de anuncios de Facebook muestra diferentes novedades de

26 West, S.M., Whittaker, M. and Crawford, K. (2019). Discriminating Systems: Gender, Race and Power in AI. AI Now Institute. Disponible en https://ainowinstitute.org/ discriminatingsystems.html.

empleo a las mujeres y a los hombres, aunque los puestos de trabajo en cuestión requieren las mismas cualificaciones de tal manera que oculta ciertos trabajos a mujeres a pesar de tener la misma formación. Los investigadores no lograron descubrir por qué es así ya que el sistema de alguna manera detectan la actual distribución demográfica de estos trabajos, porque Facebook no quiere explicar cómo funciona su sistema de publicación de anuncios.[27]

Organismos internacionales como el Consejo de Europa y la ONU[28] han advertido acerca de los riesgos, y sesgos sistémicos que presentan los sistemas de IA y que tienen el potencial de dañar y poner en desventaja a grupos vulnerables de la sociedad, incluidas las mujeres y las niñas. El Informe de la UNESCO titulado: *Prejuicios Sistemáticos. Una investigación de prejuicios contra las mujeres y las niñas en los grandes modelos de lenguaje*[29] contiene algunos ejemplos de los principales sistemas basados en IA que a menudo perpetúan e incluso amplían y amplifican los sesgos humanos estructurales y sociales dirigidos a mujeres y niñas y propone una lista de recomendaciones y acciones tanto a los creadores de política pública y desarrolladores de IA para mitigar los riesgos de los sesgos sistemáticos que presentan algunos de los grandes modelos de lenguaje actualmente utilizados.

El crimen organizado afecta a las mujeres de manera distinta que, a los hombres, especialmente en áreas como la trata de personas, el tráfico de drogas y la explotación laboral y sexual. La falta de una perspectiva de género en el diseño y aplicación de sistemas de IA puede generar que estas realidades no sean adecuadamente priorizadas o visibilizadas en los procesos de investigación y prevención de delitos. Es decir, Si las herramientas de detección y prevención se centran en crímenes dominados por hombres, como el tráfico de armas o de drogas, se podrían pasar por alto delitos que afectan mayoritariamente a mujeres y niñas.

Por otro lado, las mujeres, particularmente las jóvenes y las pertenecientes a comunidades marginadas, enfrentan formas crecientes de violencia digital. El uso de IA por parte del crimen organizado para difundir contenido sexual no consentido o extorsionar a víctimas se ha documentado en varios países. Si los sistemas de monitoreo no están diseñados para identificar esta violencia digital con una perspectiva de género, pueden no ser eficaces en su detección. Según estudios de Amnistía Internacional, las mujeres son el objetivo principal del ciberacoso y la violencia digital. Las redes criminales han intensificado el uso de deepfakes y otros métodos de manipulación de contenido digital para extorsionar y explotar a mujeres[30].

En 2019, en Bogotá se publicitaba el proyecto emprendido por la Universidad Nacional de Colombia, la Secretaría Distrital de Seguridad, Convivencia y Justicia y la empresa de matemática aplicada Quantil, que contaría con la financiación de uno 3 000 millones de pesos del Fondo de Ciencia, Tecnología e Innovación del Sistema General de Regalías y que tenía como finalidad construir, en treinta meses, modelos para describir cuatro problemáticas relacionadas con el crimen, la seguridad y la convivencia en Bogotá (López, 2019): los homicidios, los crímenes al patrimonio con uso de violencia, las lesiones personales y las dinámicas detrás de la percepción de seguridad que tiene la ciudadanía, de tal manera que se pudiera predecir y anticipar las cuatro preguntas tradicionales con “w”: when, where, who y why (cuándo, dónde, quién y por qué). Por supuesto, el profesor Francisco Gómez, del Departamento de Matemáticas de la Universidad Nacional de Colombia, reconocía que uno de los principales desafíos del proyecto era identificar y corregir los sesgos31.

Otro aspecto preocupante puede ser la escasa y limitada participación de mujeres en el desarrollo de IA y políticas tecnológicas. Según informes de la UNESCO32, una minoría de los profesionales de IA a nivel mundial son mujeres, lo que limita la diversidad de perspectivas en la creación de tecnologías destinadas a combatir el crimen organizado. La falta de una participación equitativa no solo refuerza los sesgos existentes, sino que también afecta la capacidad para crear soluciones inclusivas y que protejan a todos los sectores de la sociedad.

Por último, el uso de IA en la vigilancia y prevención del crimen puede llevar a la criminalización injusta de mujeres que se encuentran en situaciones de vulnerabilidad, como las trabajadoras sexuales o migrantes. Sin una consideración cuidadosa del contexto, los algoritmos pueden clasificar a estas mujeres como potenciales delincuentes, lo que refuerza estereotipos negativos y aumenta su marginalización.

27 HAO, Karen trad. MILUTINOVIC Ana (2021, 14 de abril). La IA de Facebook discrimina a las mujeres en los anuncios de trabajo. Technology Review. https://www.technologyreview.es/s/13219/la-ia-de-facebook-discrimina-las-mujeres-en-los-anuncios-de-trabajo

28 Naciones Unidas, “La inteligencia artificial ya reproduce estereotipos de género”, 7 de marzo de 2024 en: https://news.un.org/es/story/2024/03/1528182

29 UNESCO, « Challenging systematic prejudices: an investigation into bias against women and girls in large language models », 2024 en : https://unesdoc.unesco.org/ark:/48223/pf0000388971

30 En *Toxic Twitter: Violence and Abuse Against Women Online* (2018), disponible en https://www.amnestyusa.org/wp-content/uploads/2018/03/Toxic-Twitter.pdf

31 LÓPEZ B., Joaquín M. El sistema de inteligencia artificial para adelantarse a crímenes en Bogotá. En: La República. 22 abril 2019. Disponible en: https://www.larepublica.co/internet-economy/asi-seria-elsistema-de-inteligencia-artificial-para-adelantarse-a-crimenes-en-bogota-2854179

32 Por ejemplo, ”I’d Blush if I Could” disponible en https://unesdoc.unesco.org/ark:/48223/pf0000367416 o ” Artificial intelligence and gender equality: key findings of UNESCO’s Global Dialogue” disponible en https://unesdoc.unesco.org/ark:/48223/pf0000374174

1.5. RETOS DE DERECHOS HUMANOS

Los algoritmos y los datos utilizados en el entrenamiento y funcionamiento de los sistemas de IA pueden generar información falsa que puede poner en riesgo los derechos fundamentales de las personas, como por ejemplo en los sistemas de predicción policial, si los datos utilizados no son actualizados, verificados y auditados regularmente pueden presentar sesgos de información que ponen en riesgo el derecho a la vida privada de los individuos, el riesgo de ser discriminado e inclusive la transgresión del derecho a gozar de un juicio justo e imparcial[33].

El sistema de protección de derechos fundamentales de la Unión Europea es uno de los más completos y complejos ya que participan distintas instituciones e instancias judiciales en asegurarse de que estos derechos sean debidamente protegidos conforme al Tratado sobre el Funcionamiento de la Unión Europea, la Carta de los Derechos Fundamentales de la Unión Europea y el Convenio Europeo de Derechos Humanos[34]

Cabe destacar que organismos internacionales
han elaborado documentos y recomendaciones sobre la protección de los derechos humanos en relación al uso e implementación de algoritmos en el contexto del procesamiento de datos en el sistema judicial, entre ellos destaca el Estudio sobre las dimensiones de derechos humanos de las técnicas de procesamiento automatizado de datos y su posible implicación regulatoria[35] del Consejo de Europa que examina los efectos del uso de algoritmos y la forma en que se ejercen y garantizan los derechos humanos de conformidad con la normativa internacional de derechos humanos, incluidos los principios del estado de derecho dentro los procesos en el ámbito judicial.

La Recomendación del Comité de Ministros del Consejo de Europa sobre las repercusiones de los sistemas algorítmicos en los derechos humanos de abril de 2020 afirma que existen importantes desafíos en materia de derechos humanos asociados a la creciente dependencia de los sistemas algorítmicos en la vida cotidiana, como por ejemplo en lo que respecta al derecho a un juicio justo; el derecho a la privacidad y la protección de datos; el derecho a la libertad de pensamiento, conciencia y religión; el derecho a la libertad de expresión; el derecho a la libertad de reunión; el derecho a la igualdad de trato y los derechos económicos y sociales[36].

33 Council of Europe, «Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems. A*dopted by the Committee of Ministers on 8 April 2020*», ver Apéndice de la Recomendación Párrafo A numeral 4 en: https://search.coe.int/cm#{%22CoEIdentifier%22:[%2209000016809e1154%22],%22sort%22:[%22CoEValidationDate%20Descending%22]}

34 Gonzalez Fuster, Gloria. Study requested by the LIBE Committe of the European Parliament, «Artificial Intelligence and Law Enforcement. Impact on Fundamental Rights», julio de 2020 en:

https://www.europarl.europa.eu/RegData/etudes/STUD/2020/656295/IPOL_STU(2020)656295_EN.pdf

35 Council of Europe, «Algorithms and human rights. Study on the human rights dimensions of automated data processing techniques and possible regulatory implications, 2018» en: https://edoc.coe.int/en/internet/7589-algorithms-and-human-rights-study-on-the-human-rights-dimensions-of-automated-data-processing-techniques-and-possible-regulatory-implications.html

36 European Union Agency for Fundamental Rights (FRA), «Bias in Algorithms. Artificial Intelligence and Discrimination » 8 de diciembre de 2022, disponible en: https://fra.europa.eu/en/publication/2022/bias-algorithm

El Comité de Ministros encomendó al Comité Europeo de Problemas Criminales (CDPC) la responsabilidad de supervisar y coordinar las actividades del Consejo de Europa en el ámbito de la prevención y el control del delito, teniendo en cuenta su "Estudio de viabilidad sobre un futuro instrumento del Consejo de Europa sobre inteligencia artificial y derecho penal (2020)"[37]. A raíz de estos proyectos, en la 77ª sesión plenaria del CDPC celebrada en Estrasburgo del 3 al 6 de diciembre de 2019, se encargó al Grupo de trabajo « realizar un estudio de viabilidad que identificara el alcance y los principales elementos de un futuro instrumento del Consejo de Europa sobre IA y derecho penal, preferiblemente una convención» [38]. En concreto, se le ha encomendado la tarea de redactar un instrumento jurídico sobre la responsabilidad penal relacionada con el uso de la IA, cuya redacción está prevista para finales de 2025. En la 86.ª reunión plenaria en el Consejo de Europa, Palacio de Europa, Estrasburgo, del 20 al 22 de noviembre de 2024 se presentará el primer Documento de debate del CDPC sobre la responsabilidad penal relacionada con los sistemas de IA[39].

Teniendo presente además el Convenio Marco del Consejo de Europa sobre Inteligencia Artificial y Derechos Humanos, Democracia y Estado de Derecho (CETS No. 225) y otros documentos importantes elaborados por el Comité de Inteligencia Artificial como "Posibles elementos de un marco legal sobre inteligencia artificial, basado en las normas del Consejo de Europa sobre derechos humanos, democracia y estado de derecho (2021); Hacia la regulación de los sistemas de IA (2020); Estudio de viabilidad sobre un marco legal sobre diseño, desarrollo y aplicación de IA basado en las normas del Consejo de Europa (2020). Teniendo en cuenta el Convenio sobre la ciberdelincuencia (Convenio de Budapest, ETS nº 185) [40].

Asimismo, no hay que olvidar los trabajos del Comité de Ministros del Consejo de Europa a los Estados miembros y, en particular, en las recomendaciones: Aspectos jurídicos de los vehículos «autónomos»: Resolución 2346 (2020) y Recomendación 2187 (2020[41]); Justicia mediante algoritmos: el papel de la inteligencia artificial en los sistemas policiales y de justicia penal: Resolución 2342 (2020) y Recomendación 2182 (2020)[42]

37 Council of Europe European Committee on Crime Problems (CPDC), 'Feasibility study on a future Council of Europe instrument on artificial intelligence and criminal law', 2020. Available at: https://rm.coe.int/cdpc-2020-3-feasibility-study-of-a-future-instrument-on-ai-and-crimina/16809f9b60

38 CDPC – Lista de decisiones – 77.ª sesión plenaria, CDPC (2019) 23, punto 7, pág. 4.

39 CDPC meeting: 86th Plenary Session 20-22 November 2024 en : https://rm.coe.int/cdpc-2024-oj2-en-draft-agenda-november-2024-2788-1036-6986-v-1/1680b22c02

40 Council of Europe, 'Convention on Cybercrime', ETS No.185, 2001. ; Council of Europe, 'Chart of signatures and ratifications of the Convention on Cybercrime'.

41 Text adopted by the Standing Committee, acting on behalf of the Assembly, on 22 October 2020 (see Doc. 15143, report of the Committee on Legal Affairs and Human Rights, rapporteur: Mr Ziya Altunyaldiz) en : http://www.europeanrights.eu/public/atti/Resolution_2346_(2020)_ENG.pdf See also Recommendation 2187 (2020) : http://assembly.coe.int/nw/xml/XRef/Xref-DocDetails-en.asp?FileId=28817

42 Text adopted by the Standing Committee, acting on behalf of the Assembly, on 22 October 2020 (see Doc. 15156, report of the Committee on Legal Affairs and Human Rights, rapporteur: Mr Boriss Cilevičs) en : https://pace.coe.int/en/files/28805/html .See also Recommendation 2182 (2020) : https://pace.coe.int/en/files/28806

En el ámbito europeo, destaca el estudio solicitado por el Comité LIBE del Parlamento Europeo sobre Inteligencia Artificial y la Ejecución de la Ley. Impacto sobre Derechos Fundamentales cuyo objetivo es evaluar los avances recientes en relación con la IA en el ámbito de la aplicación de la ejecución de la ley y la justicia penal, con el fin de revisar su impacto en los derechos fundamentales de la UE y presentar recomendaciones de política al Parlamento Europeo. El estudio contiene un análisis detallado de las principales leyes europeas relacionadas que impactan la protección de derechos humanos, análisis de casos y la labor de las instituciones responsables de su ejecución incluida una sección sobre IA y justicia penal[43].

El efecto de la 'caja negra' por medio del cual los algoritmos entrenados pueden generar respuestas, resultados y outputs plantea ciertas dudas y cuestionamientos sobre su capacidad para tomar decisiones o hacer predicciones que carecen de una explicación clara y lógica respecto a su fundamento. De acuerdo con Europol, en el ámbito de la aplicación de la ley, la opacidad de la 'caja negra' plantea un desafío importante. «Cuando un sistema impulsado por IA plantea inquietudes sobre la amenaza potencial de una persona o recomienda agentes de la policía desplegar patrullas en zonas específicas, resulta imperativo que los agentes de la ley y las personas afectadas por tales decisiones comprendan la lógica subyacente ». Europol indica que la ausencia de esta información de fundamental importancia abre la puerta a la posibilidad de sesgos, errores o malas interpretaciones, lo que plantea cuestiones fundamentales de responsabilidad y justicia[44].

Para hacer frente a los retos que plantea el efecto de la 'caja negra' en sistemas de IA, principios éticos tales como transparencia, confiabilidad, explicación, imparcialidad y otros conceptos como human-in-the-loop empiezan a tomar relevancia en el ámbito internacional precisamente para evitar que los sistemas de IA transgredan y vulneren derechos fundamentales de los individuos.

Los sistemas de IA no deben remplazar el juicio y las decisiones humanas, sino que deben ser utilizados como una herramienta que puede guiar y ayudar a la toma de decisiones de las autoridades y las fuerzas del orden en investigaciones penales, incluidas las relacionadas con crimen organizado transfronterizo.

43 European Parliament, «Protecting Fundamental Rights within the Union» en: https://www.europarl.europa.eu/about-parliament/en/democracy-and-human-rights/fundamental-rights-in-the-eu

44 EUROPOL, «AI and Policing. The Benefits and Challenges of Artificial Intelligence for Law Enforcement. An Observatory Report from the Europol Innovation Lab», 2024, p.35 en: https://www.europol.europa.eu/publication-events/main-reports/ai-and-policing

BLOQUE 2: NORMATIVA INTERNACIONAL, ESTRATEGIAS Y OTRAS INICIATIVAS NO VINCULANTES

2.1. LEGISLACIÓN Y ESTRATEGIAS

CONVENCIÓN MARCO SOBRE INTELIGENCIA ARTIFICIAL DEL CONSEJO DE EUROPA

El Comité de Inteligencia Artificial (CAI) del Consejo de Europa y su grupo predecesor CAHAI grupo de expertos conformado por representantes de la industria, gobierno, academia y sociedad civil tuvo el mandato de crear un instrumento jurídico sobre IA desde 2019.

En 2022, el CAI comenzó a elaborar el texto de la convención junto con los 46 Estados Miembros del Consejo de Europa. El proceso de negociación y consulta duro más de 2 años. La *Convención Marco sobre Inteligencia Artificial*[45] se abrió para firma de los Estados Parte del Consejo de Europa el 5 de septiembre de 2024 en Vilna, Lituania y es el primer tratado internacional vinculante destinado a garantizar que el uso de sistemas de IA sea plenamente compatible con los derechos humanos, la democracia y el estado de derecho[46].

El alcance de la convención marco es bastante amplio e incluye todas las actividades dentro del ciclo de vida de los sistemas de IA llevadas a cabo tanto por autoridades públicas como por actores privados que actúen en su nombre y que tengan el potencial de interferir con los derechos humanos, la democracia y el estado de derecho. Quedan excluidas actividades relacionadas con la protección de la seguridad nacional. Entre algunas de las disposiciones que contiene la convención marco:

- Insta a los países a adoptar medidas y aplicar una lista de principios generales relacionados con el ciclo de vida de los sistemas de IA, tales como dignidad humana, autonomía, transparencia y vigilancia, responsabilidad, equidad y no discriminación, privacidad y protección de datos, confiabilidad e innovación segura.

- Una sección que recomienda a los países adoptar y mantener medidas para garantizar la disponibilidad de recursos judiciales accesibles y efectivos a los ciudadanos por posibles violaciones de los derechos humanos resultantes de las actividades dentro del ciclo de vida de los sistemas de IA.

- Establecimiento de medidas de salvaguardia para proteger los derechos de las personas afectadas.

- Un capítulo que insta a los países a adoptar o mantener medidas para la identificación, evaluación, prevención y mitigación de los riesgos que plantean los sistemas de IA, considerando los impactos reales y potenciales a los derechos humanos, la democracia y el estado de derecho, así como otras disposiciones relacionadas con los derechos de las personas con discapacidad, niños, la promoción de alfabetización digital y habilidades digitales y la salvaguardia de los derechos humanos, así como

- Disposiciones relacionadas con la Conferencia de los Estados Parte (que será el órgano de seguimiento), cooperación internacional e implementación de mecanismos de control y vigilancia.

A la fecha del presente estudio, dicho instrumento ha sido firmado por 10 países (Andorra, Georgia, Islandia, Noruega, la República de Moldavia, San Marino, Reino Unido, Israel, EE.UU. y la UE[47]. Requiere la

45 Council of Europe Framework Convention on Artificial Intelligence, Human Rights and the Rule of Law, CETS No. 225, Vilnius 5.IX.2024 en: https://rm.coe.int/1680afae3c

46 Council of Europe, «The Framework Convention on Artificial Intelligence» en: https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence

47 Ver Tabla de firmas y ratificaciones en: https://www.coe.int/en/web/conventions/full-list?module=signatures-by-treaty&treatynum=225

ratificación de 5 países -incluidos 3 miembros del Consejo de Europa- para entrar oficialmente en vigor. Países como Argentina, Costa Rica, México, Perú y Uruguay han formado parte de las negociaciones de la convención.

La implementación de esta convención requerirá que los países que la firmen y ratifiquen tengan que crear, adoptar e implementar marcos jurídicos nacionales consistentes con las disposiciones de ese instrumento, incluido el desarrollo de estrategias nacionales y políticas relacionadas con sistemas de IA que incluya la dimensión de la protección de los derechos fundamentales de las personas.

REGLAMENTO EUROPEO SOBRE INTELIGENCIA ARTIFICIAL

La Unión Europea a través de instituciones como la Comisión Europea, el Parlamento Europeo (a través de sus distintos comités AIDA, IMCO y LIBE) y el Consejo Europeo han jugado un rol decisivo y clave en generar e impulsar un marco jurídico sobre IA con el propósito de proteger los derechos de las personas ante los riesgos que genera la utilización de sistemas de IA durante su ciclo de vida con base en la categorización del nivel riesgo que presentan a la sociedad y estableciendo prohibiciones de determinadas prácticas que atenten contra los principios democráticos de la UE y las libertades de los ciudadanos[48].

El texto final del Reglamento Europeo sobre IA (REIA) fue publicado en el Diario Oficial de la Unión Europea el 12 de julio de 2024[49]. El REIA entró oficialmente en vigor el 1º de agosto de 2024 y se aplicará de forma gradual, en un plazo de dos años, con algunas excepciones: las disposiciones generales y las prohibiciones se aplicarán después de seis meses; las normas de gobernanza y las obligaciones para los modelos de IA de uso general se aplicarán después de doce meses; y las normas para los sistemas de IA integrados en productos regulados entrarán en vigor después de tres años[50].

El REIA reglamentara disposiciones en otras leyes secundarias europeas conforme a su entrada en vigor. El texto del REIA contiene 180 recitales y un total de 113 artículos distribuidos a lo largo de trece capítulos que establece un conjunto de reglas generales para abordar los riesgos creados específicamente por los sistemas y aplicaciones de IA que:

- Prohíbe prácticas de IA que supongan riesgos inaceptables para el individuo;
- Establece requisitos para los sistemas de IA destinados a aplicaciones consideradas como de alto riesgo;
- Contiene obligaciones específicas para los proveedores, distribuidores e importadores de aplicaciones de IA;
- Exige llevar a cabo una evaluación de la conformidad antes de que un determinado sistema de IA se ponga en servicio o se comercialice en la UE;
- Establece medidas de cumplimiento después de que un determinado sistema de IA se comercialice;
- Establece una estructura de gobernanza a nivel europeo y nacional en donde participaran diversas instituciones y autoridades reguladoras para asegurar su cumplimiento, así como multas y sanciones en caso de incumplimiento por parte los proveedores de aplicaciones de IA[51].

48 Para una síntesis de las diversas instituciones de la Unión Europea que han participado en la creación, negociación e implementación del Reglamento Europeo sobre IA, ver: Centre for AI and Digital Policy (CAIDP), *Artificial Intelligence and Democratic Values Index 2023*, pp. 56-77, disponible en: https://www.caidp.org/reports/aidv-2023/

49 Véase: Reglamento (UE) 2024/1689 de 13 junio de 2023 por el que se establecen normas armonizadas en materia de inteligencia artificial y por el que se modifican los Reglamentos (CE) no. 300/2008, (UE) no. 167/2013, (UE) no. 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828 (Reglamento de Inteligencia Artificial) en: https://eur-lex.europa.eu/eli/reg/2024/1689/oj

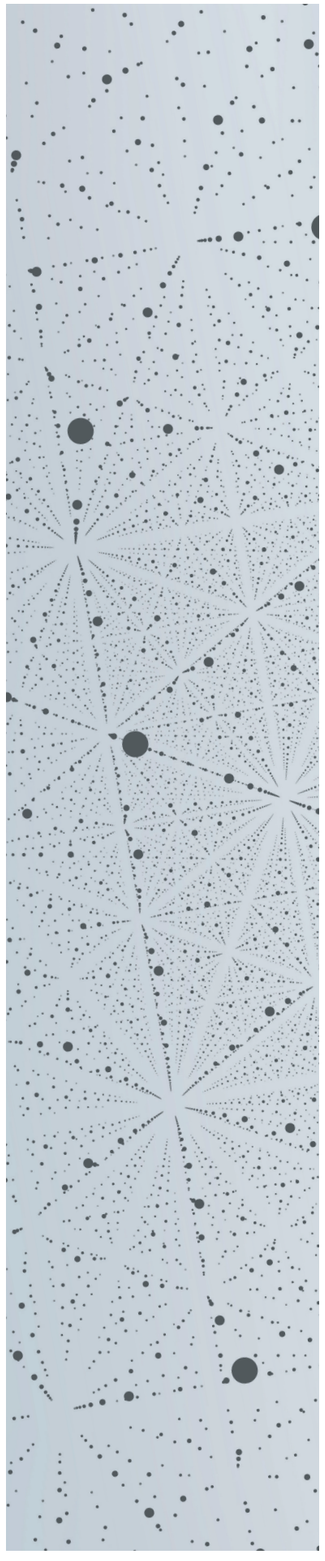
50 Para una línea de tiempo y fechas clave relevantes para la implementación del REIA, ver: https://artificialintelligenceact.eu/ai-act-implementation-next-steps/

51 European Commission AI Act https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

El REIA establece algunas excepciones. El apartado cuarto del Art. 2 establece que el Reglamento no será aplicable a las autoridades públicas de terceros países ni a las organizaciones internacionales cuando utilicen sistemas de IA en el marco de acuerdos internacionales con fines de aplicación de la ley y cooperación judicial con la UE o con uno o varios Estados miembros de la UE siempre que tal tercer país u organización internacional ofrezca garantías suficientes con respecto a la protección de los derechos y libertades fundamentales de las personas.

Entre algunas de las disposiciones de relevancia para las fuerzas del orden y las autoridades investigadoras del sistema de justicia que establece el REIA se encuentran:

- El Anexo III que comprende ocho ámbitos de categorización de sistemas de Alto Riesgo a que se refiere el Art. 6 Apartado 2. El numeral 6 del Anexo III incluye: «Garantía del cumplimiento del Derecho, en la medida en que su uso esté permitido por el Derecho de la Unión o nacional aplicable» y prevé cinco supuestos adicionales:
  - Sistemas de IA destinados a ser utilizados por las autoridades garantes del cumplimiento del Derecho, o en su nombre, o por las instituciones, órganos y organismos de la Unión en apoyo de las autoridades garantes del cumplimiento del Derecho, o en su nombre, para evaluar el riesgo de que una persona física sea víctima de delitos.
  - Sistemas de IA destinados a ser utilizados por las autoridades garantes del cumplimiento del Derecho, o en su nombre, o por las instituciones, órganos y organismos de la Unión en apoyo de las autoridades garantes del cumplimiento del Derecho como polígrafos o herramientas similares;
  - Sistemas de IA destinados a ser utilizados por las autoridades garantes del cumplimiento del Derecho, o en su nombre, o por las instituciones, órganos y organismos de la Unión en apoyo de las autoridades garantes del cumplimiento del Derecho para evaluar la fiabilidad de las pruebas durante la investigación o el enjuiciamiento de delitos;
  - Sistemas de IA destinados a ser utilizados por las autoridades garantes del cumplimiento del Derecho, o en su nombre, o por las instituciones, órganos y organismos de la Unión en apoyo de las autoridades garantes del cumplimiento del Derecho para evaluar el riesgo de que una persona física cometa un delito o reincida en la comisión de un delito atendiendo no solo a la elaboración de perfiles de personas físicas mencionada en el artículo 3, punto 4, de la Directiva (UE) 2016/680 o para evaluar rasgos y características de la personalidad o comportamientos delictivos pasados de personas físicas o colectivos



- Sistemas de IA destinados a ser utilizados por las autoridades garantes del cumplimiento del Derecho, o en su nombre, o por las instituciones, órganos y organismos de la Unión en apoyo de las autoridades garantes del cumplimiento del Derecho para elaborar perfiles de personas físicas, como se menciona en el artículo 3, punto 4, de la Directiva (UE) 2016/680, durante la detección, la investigación o el enjuiciamiento de delitos

→ El numeral 7 del Anexo III «Migración, asilo y gestión del control fronterizo, en la medida en que su uso esté permitido por el Derecho de la Unión o nacional aplicable».

→ El numeral 8 del Anexo III «Administración de justicia y procesos democráticos» incluye el apartado a): «Sistemas de IA destinados a ser utilizados por una autoridad judicial, o en su nombre, para ayudar a una autoridad judicial en la investigación e interpretación de hechos y de la ley, así como en la garantía del cumplimiento del Derecho a un conjunto concreto de hechos, o a ser utilizados de forma similar en una resolución alternativa de litigios»

El Art. 5 apartado 1 h) del REIA establece que el uso de sistemas de identificación biométrica remota «en tiempo real» en espacios de acceso público con fines de aplicación de la ley se encuentra prohibido, salvo que su uso sea estrictamente necesario para alcanzar alguno de los siguientes objetivos:

→ "la búsqueda selectiva de víctimas concretas de secuestro, trata de seres humanos o explotación sexual de seres humanos, así como la búsqueda de personas desaparecidas;

→ la prevención de una amenaza específica, importante e inminente para la vida o la seguridad física de las personas físicas o de una amenaza real y actual o real y previsible de un atentado terrorista;

→ la localización o identificación de una persona sospechosa de haber cometido un delito a fin de llevar a cabo una investigación o un enjuiciamiento penales o de ejecutar una sanción penal por alguno de los delitos mencionados en el anexo II que en el Estado miembro de que se trate se castigue con una pena o una medida de seguridad privativas de libertad cuya duración máxima sea de al menos cuatro años".

Los usos previstos en el REIA están sujetos a la autorización de un órgano judicial u otro órgano administrativo independiente y al cumplimiento de salvaguardias y condiciones necesarias y proporcionadas en relación con el uso, en particular en lo que respecta a las limitaciones temporales, geográficas y personales conforme al numerales 2, 3, y 4 del Art. 5º del REIA.

ESTRATEGIAS REGIONALES Y NACIONALES

La Comisión Europea presentó inicialmente un comunicado el 25 de abril de 2018 que contiene la **Estrategia sobre Inteligencia Artificial de la UE** que incluye diversos objetivos entre los cuales se encuentran: (i) posicionar a la UE en un competitivo panorama internacional; (ii) impulsar la capacidad tecnológica e industrial de la UE y la adopción de la IA en toda la economía ; (iii) acercar la IA a las pequeñas empresas y usuarios potenciales ; (iv) garantizar el establecimiento de un marco ético y jurídico adecuado, basado en los valores de la UE y en consonancia con la Carta de los Derechos Fundamentales de la UE[52].

De acuerdo con información de Comisión Europea[53] y la Asociación Europea de Agencias de Comunicación[54], hasta 2022, veinte Estados Miembros de la UE y Noruega cuentan con una estrategia nacional sobre IA. No obstante, muy pocos países de la UE cuentan con una estrategia de IA específica para las fuerzas del orden y las autoridades del sistema de justicia.

52 Comisión Europea, «Comunicacion de la Comisión. Inteligencia Artificial para Europa » COM(2018) 237 final, 25.4.2018. disponible en: https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:52018DC0237

53 European Commision JRC Technical Report. AI Watch National Strategies on Artificial Intelligence: A European Perspective. 2022 Edition, disponible en: https://op.europa.eu/en/publication-detail/-/publication/54e385d8-eac0-11ec-a534-01aa75ed71a1

54 Véase https://eaca.eu/news/national-ai-strategies-in-europe/

Por su parte, algunos países en ALC han comenzado a elaborar estrategias nacionales sobre IA con el fin de establecer un marco de políticas y lineamientos que pueden ser útiles para el sector público y privado en el desarrollo y la utilización de sistemas de IA. La Recomendación sobre Inteligencia Artificial de la OCDE y la Recomendación sobre la Ética de la Inteligencia Artificial de la UNESCO están generando un impulso en algunos países de la región para el desarrollo de estrategias nacionales sobre IA más amplias que incluyan los principios éticos y recomendaciones de dichos organismos.

La OCDE reporta que siete países de ALC ya han desarrollado o están en proceso de desarrollar una estrategia nacional de IA, entre ellos destacan Argentina, Brasil, Chile, Colombia, México, Perú y Uruguay. Ese organismo indica que la mayoría de los países de ALC aún y cuando carecen de una estrategia de IA ya han publicado una estrategia nacional de gobierno digital más amplia, o una agenda o programa digitales vinculados que incluyen componentes que actúan como bases fundacionales de la IA (por ejemplo, interoperabilidad, infraestructura, herramientas y procesos analíticos, integración de servicios, etc.), a pesar de que esta no suele incorporarse como objeto principal[55]. Ese organismo reporta que países como Ecuador, Costa Rica, República Dominicana y Panamá ya se encuentran explorando abordar estrategias nacionales sobre IA aunque a la fecha no han sido oficialmente adoptadas y publicadas por sus respectivos gobiernos. El gobierno de Costa Rica, a través del Ministerio de Ciencia, Innovación, Tecnología y Telecomunicaciones (MICITT) publicó su Estrategia Nacional de Inteligencia Artificial 2024-2027 la tercera semana de octubre de 202456.

[1]

Si bien las estrategias nacionales de IA en la región abordan cuestiones muy amplias en diferentes áreas relacionadas con la adopción de tecnologías de IA en el sector público tales como financiamiento, educación, salud, innovación y desarrollo y aspectos de gobernanza, ninguna de ellas hace particular referencia a la forma en que deben ser abordadas por el sector de justicia penal incluido el poder judicial.

55 OECD/CAF Development Bank of Latin America (2022), "Estrategias de inteligencia artificial en América Latina y el Caribe", in The Strategic and Responsible Use of Artificial Intelligence in the Public Sector of Latin America and the Caribbean, OECD Publishing, Paris, pp.24-25. DOI: https://doi.org/10.1787/03c4e7eb-es. Para un análisis de las Estrategias Nacionales de IA en Argentina, Chile, Colombia, México y Uruguay, ver: CeTyS «Inteligencia Artificial en América Latina y el Caribe. Ética, Gobernanza y Políticas», ISN 2684-0278, pp. 134-147, en: https://proyectoguia.lat/wp-content/uploads/2020/10/compilado-espanol-compressed.pdf

56 La Estrategia Nacional de Inteligencia Artificial 2024-2027 de Costa Rica en: https://observatorioecuadordigital.mintel.gob.ec/wp-content/uploads/2024/10/Estrategia-Nacional-de-IA-Costa-Rica-.pdf

2.2. PRINCIPIOS E INICIATIVAS INTERNACIONALES NO VINCULANTES RELEVANTES

DIRECTRICES ÉTICAS PARA UNA IA FIABLE DEL GRUPO INDEPENDIENTE DE EXPERTOS DE ALTO NIVEL SOBRE INTELIGENCIA ARTIFICIAL (HLEG)

En 2018, la Comisión Europea conformo un grupo independiente de expertos de alto nivel cuyo mandato fue facilitar y asesorar a la Comisión Europea en el desarrollo de la estrategia europea sobre IA. Dentro de los entregables de dicho grupo destaca un documento sobre *Lineamientos Éticos para una IA Fiable* publicado en abril de 2019. El documento establece una definición sobre IA fiable y desarrolla siete principios clave junto con una lista de evaluación concreta cuyo propósito es ayudar a verificar la aplicación de cada uno de esos principios. Entre los principios que contiene la guía se encuentran: (i) agencia y supervisión humanas; (ii) robustez técnica y seguridad; (iii) privacidad y gobernanza de datos; (iv) transparencia; (v) diversidad, no discriminación y equidad; (vi) bienestar social y ambiental; y (vii) responsabilidad[57].

Este documento fue una de las primeras guías a nivel internacional que desarrollo una definición sobre sistema de IA y principios generales éticos para el desarrollo y utilización de tecnologías de IA.

RECOMENDACIÓN DE LA OCDE SOBRE INTELIGENCIA ARTIFICIAL

El Consejo de la OCDE adoptó una Recomendación sobre Inteligencia Artificial el 22 de mayo de 2019[58]. Se trata de la primera normativa intergubernamental que tiene como propósito fomentar la innovación y la confianza de los sistemas de IA y promover la gestión responsable de una IA fiable, garantizando el respeto de los derechos humanos y los valores democráticos entre los países miembros de ese organismo.

La Recomendación contiene una sección para la gestión responsable de una IA fiable que establece cinco principios complementarios relevantes para todas las partes interesadas: (i) crecimiento inclusivo, desarrollo sostenible y bienestar; (ii) respeto del estado de derecho, los derechos humanos y los valores democráticos, incluida la equidad y la privacidad; (iii) transparencia y explicabilidad; (iv) solidez, seguridad y protección; y, (v) responsabilidad. Esta sección insta además a los diferentes actores de la IA a promover e implementar estos principios en el ámbito de sus respectivas actividades y funciones.

La segunda sección ofrece cinco recomendaciones dirigidas a los países Miembros y no Miembros que se han adherido a la Recomendación para que implementen en el marco de sus políticas nacionales lo siguiente: (i) inversión en investigación y desarrollo de la IA; (ii) fomentar un ecosistema inclusivo que facilite la IA; (iii) implementar un entorno de gobernanza y políticas interoperables que facilite la IA; (iv) desarrollar la capacidad humana y prepararse para la transformación del mercado laboral; y, (v) fomentar la cooperación internacional para una IA fiable.

La OCDE cuenta con un observatorio de políticas de IA cuyo propósito es difundir el trabajo de ese organismo, monitorear la labor de los países adherentes en la implementación de los principios establecidos en la recomendación y promover la cooperación con los distintos organismos internacionales y regionales que se encuentren trabajando en políticas sobre IA[59].

La Recomendación sobre Inteligencia Artificial fue modificada en 2024 para actualizar la definición sobre "sistema de IA" y clarificar el alcance de algunos de los principios entre los que se encuentra abordar cuestiones de seguridad, de tal forma que, si los sistemas de IA corren el riesgo de causar algún daño o exhibir un comportamiento no deseado, puedan ser anulados, reparados y/o revocados de forma segura mediante la interacción humana, entro otros aspectos.[60]

Destaca la labor de la OCDE en el desarrollo de una clasificación sobre sistemas de IA[61]; un catálogo de herramientas y métricas sobre IA confiable[62] y una herramienta para el monitoreo de incidentes de IA[63] cuyo propósito es ayudar a los a los profesionales de la IA, responsables de políticas y a las diversas partes interesadas a obtener información relevante sobre los riesgos y los daños que pueden generar los sistemas de IA.

RECOMENDACIÓN SOBRE ÉTICA E INTELIGENCIA ARTIFICIAL DE LA UNESCO

UNESCO adoptó una *Recomendación sobre la Ética de la Inteligencia Artificial* el 23 de noviembre de 2021 cuyo objetivo principal es establecer un marco universal de valores, principios y acciones para orientar a los países en la formulación de leyes, políticas u otros instrumentos relativos a la IA, de conformidad con el derecho internacional[64].

Esa recomendación contiene cuatro principales objetivos y además una sección sobre 'Valores' que incluye cuatro rubros entre los que

57 European Commision, *Ethics guidelines for trustworthy AI*, 8 de abril de 2019 en: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

58 OECD Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449. Adoptada el 22.05.2019 y reformada el 03.05.2024 en: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

59 OECD Policy AI Observatory en: https://oecd.ai/en/ai-principles

60 Para algunos aspectos relevantes de la revisión de la Recomendación sobre IA de la OCDE, ver: OECD, «Report of the Implementation of the OECD Recommendation on Artificial Intelligence» C/MIN(2024)17, 24 de abril de 2024 en: https://one.oecd.org/document/C/MIN(2024)17/en/pdf

61 OECD (2022), «OECD Framework for the Classification of AI systems », *OECD Digital Economy Papers*, No. 323, OECD Publishing, Paris, en: https://doi.org/10.1787/cb6d9eca-en

62 OECD.AI Policy Observatory, Catalogue of Tools & Metrics for Trustworthy AI en: https://oecd.ai/en/catalogue/overview

63 OECD.AI Policy Observatory, OECD AI Incidents Monitor (AIM) en: https://oecd.ai/en/incidents

64 UNESCO, «Recomendación sobre la ética de la inteligencia artificial» Adoptada el 23 de noviembre de 2021, SHS/BIO/PI/2021/1, 2022 en: https://unesdoc.unesco.org/ark:/48223/pf0000381137

destaca: “el respeto, protección y promoción de los derechos humanos y las libertades fundamentales y de la dignidad humana”; así como una sección con diez ‘Principios’: (i) Proporcionalidad e inocuidad, (II) Seguridad y protección; (iii) Equidad y no discriminación; (iv) Sostenibilidad, (v) Derecho a la intimidad y protección de datos, (vi) Supervisión y decisión humanas; (vii) Transparencia y explicabilidad; (viii) Responsabilidad y rendición de cuentas; (ix) Sensibilización y educación; y (x) Gobernanza y colaboración adaptativas y de múltiples partes interesadas.

La Recomendación incluye además once ámbitos de acción política en diferentes áreas y un mecanismo de seguimiento y evaluación apoyado principalmente en tres metodologías relacionadas con la evaluación del impacto ético de las tecnologías de la IA, del estado de preparación para ayudar a los Estados Miembros y la evaluación ex ante y ex post respecto a la eficacia y la eficiencia de las políticas y los incentivos relacionados con la ética de la IA.

UNESCO en colaboración con un organismo internacional, un centro de investigación, una fundación y el gobierno de Japón creo un Observatorio sobre Gobernanza y Ética Global de IA[65] cuyo propósito es mostrar información sobre el estado actual de preparación de sus países miembros en la adopción de la IA de manera ética y responsable conforme a su recomendación. El observatorio cuenta además con un Laboratorio de Ética y Gobernanza de la IA, que reúne contribuciones, investigaciones, kits de herramientas y buenas prácticas en la implementación de los tres documentos que ha generado dicho organismo: (i) la Recomendación sobre Ética de la IA[66], (ii) la metodología de evaluación del estado de preparación de los países (RAM)[67]; y (iii) la Evaluación del impacto ético[68].

Destacan los siguientes reportes e iniciativas de la UNESCO:

Un kit de herramientas global sobre IA y el estado de derecho para el poder judicial cuyo propósito es ofrecer a los operadores judiciales el conocimiento y las herramientas necesarias para comprender los beneficios y riesgos del uso de la IA en el ámbito de sus actividades y ofrecer orientación sobre las instancias, principios y regulaciones del derecho internacional de los derechos humanos y la jurisprudencia emergente que sustenta el uso responsable de la IA[69].

Dos reportes sobre IA conforme a su metodología de evaluación del estado de preparación de los países (RAM) de dos países de ALC: México[70] y Chile[71].

65 UNESCO, Global AI Ethics and Governance Observatory en: https://www.unesco.org/ethics-ai/en?hub=32618

66 Ver: https://unesdoc.unesco.org/ark:/48223/pf0000381137

67 Ver: https://unesdoc.unesco.org/ark:/48223/pf0000385198_spa

68 Ver: https://unesdoc.unesco.org/ark:/48223/pf0000386276

69 UNESCO, «Kit de herramientas global sobre IA y el Estado de derecho para el poder judicial» 2023, en: https://unesdoc.unesco.org/ark:/48223/pf0000387331_spa

70 Ver: https://unesdoc.unesco.org/ark:/48223/pf0000390568?posInSet=2&queryId=d8165a01-53a2-4058-b878-84794bf39668

71 Ver: https://unesdoc.unesco.org/ark:/48223/pf0000387216_spa

Recientemente anunció una alianza estratégica con el Consejo Superior de la Judicatura de Colombia para promover el uso ético de la IA en el sistema judicial de ese país. La alianza tiene como objetivo desarrollar directrices y capacidades sobre el uso responsable de la IA en los despachos judiciales principalmente en cuatro áreas clave, incluido el desarrollo de capacidades a través de talleres de formación y capacitación[72].

OTRAS INICIATIVAS RELEVANTES

La Asamblea General de las Naciones Unidas aprobó la resolución general 78/L.49 el 11 de marzo de 2024[73] que contienen algunos lineamientos y directrices relacionados con la gobernanza de IA. Entre las recomendaciones que contiene ese instrumento resalta el numeral 5:

> *“Destaca que se deben respetar, proteger y promover los derechos humanos y las libertades fundamentales durante todo el ciclo vital de los sistemas de inteligencia artificial, exhorta a todos los Estados Miembros y, en su caso, a otros interesados, a que se abstengan o dejen de usar sistemas de inteligencia artificial que sean imposibles de operar en consonancia con el derecho internacional o que supongan riesgos indebidos para el disfrute de los derechos humanos, en especial de quienes se encuentran en situaciones vulnerables, y reafirma que los derechos de las personas también deben estar protegidos en Internet, también durante el ciclo vital de los sistemas de inteligencia artificial.”*

El Centro de Inteligencia Artificial y Robótica del Instituto Interregional de las Naciones Unidas para Investigaciones sobre la Delincuencia y la Justicia (UNICRI) e Interpol desarrollaron junto con un grupo de expertos multisectorial y con el apoyo financiero de la Unión Europea un documento conocido como ‘*Kit de herramientas sobre Inteligencia Artificial. Innovación Responsable en materia de IA para las fuerzas de la aplicación de la ley*[74].’

Este kit de herramientas sobre IA tiene como finalidad ayudar a las fuerzas del orden a mejorar su comprensión y desarrollo en el uso de sistemas de IA de una manera alineada y consistente con la legislación y normativa sobre protección de derechos humanos, los principios éticos y policiales y la innovación en el uso de este tipo de herramientas. Contiene información y ejemplos prácticos para asistir a las fuerzas del orden a cumplir con principios y fundamentos éticos en el proceso de innovación y utilización de herramientas de IA en el ámbito de sus actividades e inclusive contiene una guía para el desarrollo de una estrategia de IA dirigida a las unidades de policía con una explicación sobre los pasos y las actividades que se deben cumplir e implementar.

Dentro de las futuras actividades de UNICRI se contempla incluir una versión interactiva del kit de herramientas basada en un portal, brindar mayor capacitación a las agencias sobre innovación responsable en IA, brindar tutoría dirigida a puestos ejecutivos en unidades policiales clave y ampliar la comprensión de las limitaciones y los factores que influyen en el uso de la IA por parte de autoridades policiales y fuerzas de seguridad[75].

72 UNESCO, «UNESCO y Colombia: Lideres en el uso ético y responsable de la IA en el Poder Judicial» 10 de octubre de 2024 en: https://www.unesco.org/es/articles/unesco-y-colombia-lideres-en-el-uso-etico-y-responsable-de-la-ia-en-el-poder-judicial

73 Resolución General de la ONU 78/L.49, Aprovechar las oportunidades de sistemas seguros, protegidos y fiables de inteligencia artificial para el desarrollo sostenible, 11 de marzo de 2024, en: https://digitallibrary.un.org/record/4040897?v=pdf

74 UNICRI, INTERPOL «Responsible AI Innovation in Law Enforcement. AI Toolkit » Revised February 2024 en: https://unicri.it/sites/default/files/2024-02/03_Organizational_Roadmap_Feb24.pdf

75 UNICRI, «The Toolkit for Responsible Artificial Intelligence Innovation in Law Enforcement» en: https://unicri.it/topics/Toolkit-Responsible-AI-for-Law-Enforcement-INTERPOL-UNICRI

BLOQUE 3: PRINCIPALES DELITOS COMETIDOS UTILIZANDO HERRAMIENTAS DE IA

Al igual que ha ocurrido en otras ocasiones con nuevas tecnologías que han ido apareciendo, la IA será aprovechada por los criminales para sus actividades ilícitas con toda la intensidad que les sea posible. En el caso específico del crimen organizado, sus grades facilidades de financiación eliminarán cualquier barrera económica que pueda limitar la integración de la IA en los procedimientos criminales. Hemos sido testigos de este tipo de episodios en los que las nuevas tecnologías son empleadas de formas muy eficientes por el crimen organizado, es el caso, por ejemplo, del aprovechamiento que las organizaciones criminales han hecho de las tecnologías criptográficas para garantizar el secreto de sus comunicaciones con plataformas de comunicaciones cifradas como ENCROCHAT o SKYECC[76]. Otro factor que, sin duda, dinamizará la integración de la IA en los procedimientos criminales es la falta de sujeción por su parte a cualquier tipo de normativa que legal que se establezca a nivel nacional o internacional para evitar los efectos nocivos de esta tecnología. Es de esperar, por tanto, que el grado y el ritmo al que la IA se incorpore a los procedimientos empleados por el crimen organizado suponga un reto extraordinario para la sociedad y, en especial, para las autoridades y organismos encargados de la lucha activa contra esta lacra.

76 La operación Trojan Shield/Green Light demostró el uso criminal de esta tecnología de forma masiva por las organizaciones criminales de todo el mundo. https://www.europol.europa.eu/media-press/newsroom/news/800-criminals-arrested-in-biggest-ever-law-enforcement-operation-against-encrypted-communication

Blauth et al. (2022)[77] distinguen las formas en las que la IA puede ser empleada con intenciones maliciosas en dos categorías, por un lado, aquellas en las que se abusa de las vulnerabilidades de la propia IA (abuse of IA) y, por otro, aquellas en las que se usa la IA (IA-enabled and IA-enhanced criemes) para la comisión de un crimen. Dentro de la primera categoría, encontraríamos actividades maliciosas como los ataques a la integridad de los modelos de IA, el aprovechamiento criminal de resultados no intencionales ni esperados de la IA, la manipulación de los sistemas IA de comercio algorítmico de alta velocidad empleados en bolsa o los ataques de inferencia que permiten desanonimizar datos empleados para el entrenamiento de los algoritmos.

En lo que respecta a la segunda categoría, es decir, a aquellos usos maliciosos que aprovechan la IA para posibilitar o facilitar el crimen, Bhelauth et al. (2022) distinguen los ataques basados en ingeniería social que emplean la decepción y el phishing para manipular al ser humano, los ataques de desinformación y fake news, los ataques de hacking, que abarcan el *malware*, las deepfakes y los ataques repetitivos, y por último los sistemas armamentísticos autónomos.

Hayward y Maas (2021)[78], por su parte, identifican la IA como un fenómeno potencialmente criminógeno, tanto en el sentido de su capacidad para escalar crímenes ya existentes como de facilitar nuevas transgresiones digitales. Estos autores establecen tres categorías criminógenas de la IA. La primera categoría es la de los crímenes contra la IA (crimes on AI), en la línea de lo que Blauth et al. (2022) denominan abuso de la IA, y que consideran la IA como la superficie o el objeto que sufre el ataque. La segunda categoría identificada por Hayward & Maas (2021) es la de los crímenes con IA (AI as a tool), que se corresponde con la categoría denominada por Blauth et al. (2022) como ataques posibilitados o facilitados por la IA (AI-enabled and AI-enhanced). Por último, Hayward & Maas (2021) establecen una categoría que denominan crímenes cometidos por la IA, que podrían tener un origen en resultados no esperados por los desarrolladores de una IA, y que, inevitablemente, da pie a suscitar el debate sobre el estatus legal de los algoritmos de IA y su posible responsabilidad.

77 Blauth, T. F., Gstrein, O. J., & Zwitter, A. (2022). Artificial Intelligence Crime: An Overview of Malicious Use and Abuse of AI. IEEE Access, 10, 77110–77122. https://doi.org/10.1109/ACCESS.2022.3191790

78 Hayward, K. J., & Maas, M. M. (2021). Artificial intelligence and crime: A primer for criminologists. Crime, Media, Culture, 17(2), 209–233. https://doi.org/10.1177/1741659020917434

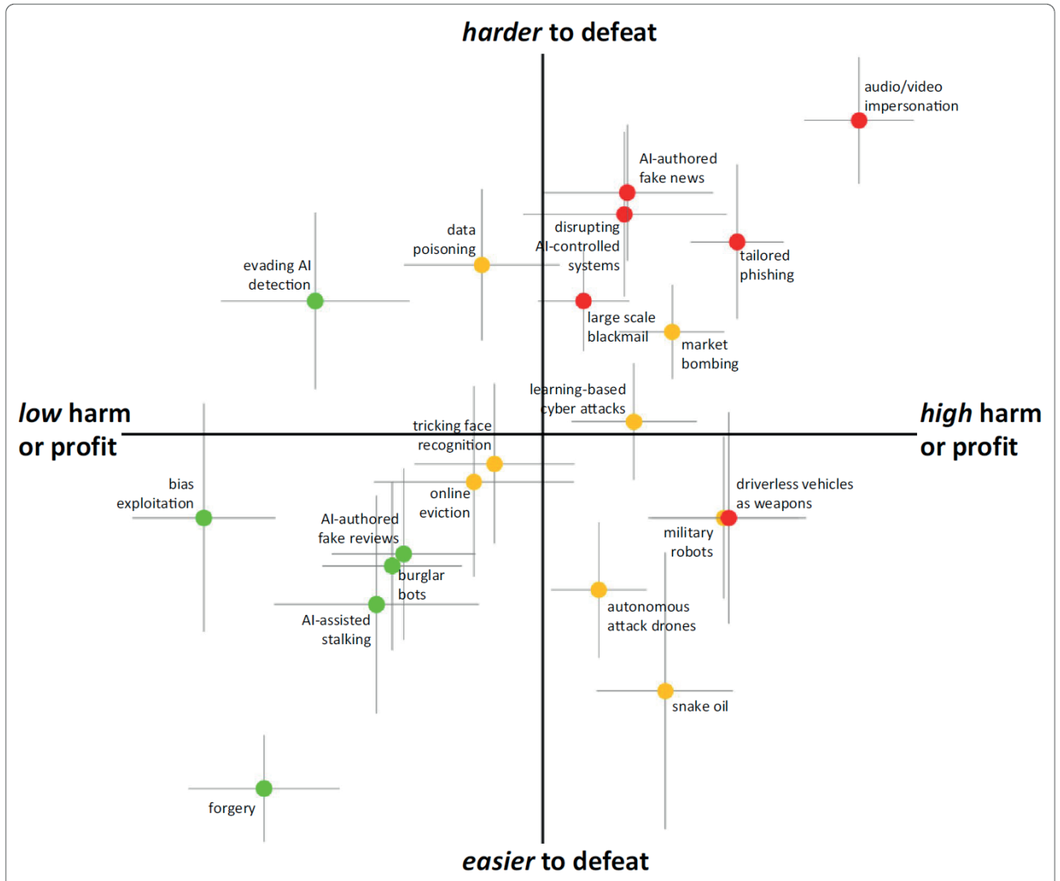


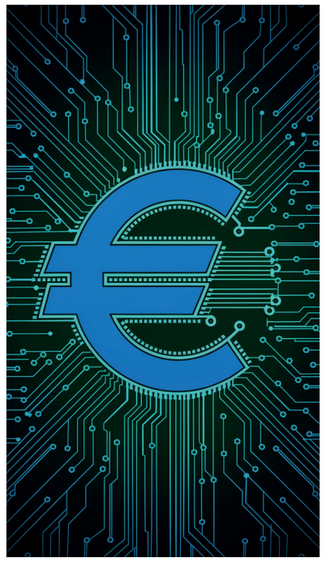
Fig. 3 Difficulty of defeat relative to the harmfulness or profitability of the crime. The most concerning crimes are those in the upper right quadrant, being both very harmful and hard to defeat. Crimes in the lower right quadrant potentially offer the strongest potential for intervention, being both harmful and defeatable. (Error bars indicate the between-group interquartile range for the ratings.)

Desde un punto de revisión documental, King et al. (2020)[79] identifican algunas tipologías criminales susceptibles de ser facilitadas por la IA, que denominan AI-Crime (AIC), y que han sido tratadas por la literatura científica. En primer lugar, mencionan el ámbito del comercio, los mercados financieros y la insolvencia. En este caso, la literatura analizada plantea incertidumbre sobre el posible uso de la IA en la manipulación de los mercados, la fijación de precios o en posibles delitos de colusión contra la competencia. Otro ámbito delictivo identificado es el de las drogas, en el que la literatura revisada sugiere que la IA puede ser empleada como instrumento para el tráfico y la venta de estas sustancias ilícitas, desde el uso de vehículos autónomos para su trasporte, hasta el uso de algoritmos de machine learning para la publicidad de la venta de las sustancias mediante bots en redes sociales. En cuanto a los delitos contra las personas, la literatura incide en los casos de acoso y torturas, por ejemplo, en lo que se refiere al acoso, a través de bots en redes sociales o el uso de contenido fake de manera sistemática o no contra una persona. En cuanto a delitos sexuales, la literatura se refiere al impacto de la IA tanto en agresiones y abusos sexuales como en la pederastia, si bien la cuestión no es pacífica y esta implicación la IA en delitos sexuales sigue siendo una cuestión a debate como una hipotética área del AIC. Por último, King et al. (2020) constatan que la literatura conecta la falsificación y la suplantación de identidad a través de AIC con el robo y el fraude no corporativo, e implican el uso el machine learning en el fraude corporativo. En el caso del robo y el fraude no corporativo, la literatura contempla el uso de la IA para la recopilación de datos personales de un individuo en una primera fase y para la suplantación de su identidad en una segunda.

79 King, T. C., Aggarwal, N., Taddeo, M., & Floridi, L. (2020). Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions. Science and Engineering Ethics, 26(1), 89–120. https://doi.org/10.1007/s11948-018-00081-0

Otro estudio interesante es el realizado por Caldwell et al. (2020)[80], publicado en la revista Crime Science, en el que identificaban y evaluaban posibles usos de la IA con intereses criminales. En el estudio participaron representantes académicos, de policía, de fuerzas armadas, gubernamentales y del sector privado. Los criterios conforme a los cuales se catalogaron esos posibles usos maliciosos de la IA fueron cuatro: le daño a las víctimas, el beneficio criminal, posibilidad de realización criminal y su dificultad para ser contrarrestado. Las conclusiones del estudio distinguían 18 posibles usos de la IA para la actividad criminal o terrorista, destacando 6 de ellos como los más preocupantes: usurpación de identidad mediante vídeo o audio, el phishing a medida, el uso de vehículos autónomos como arma, las fake news, el chantaje a gran escala y el sabotaje de sistemas controlados por IA. A continuación, se reproduce el gráfico generado por Caldwell et al. (2020) en su trabajo y que resume el resultado del estudio por considerarse de alto interés.

El uso de herramientas y tecnologías de IA ya empiezan a ser utilizadas y explotadas por grupos del crimen organizado en países de ALC. Ejemplos de algunas conductas delictivas en las que se utilizaron herramientas de IA para perpetrar delitos se encuentran las siguientes:



3.1. FRAUDE FINANCIERO Y BANCARIO

El cibercrimen es uno de los ámbitos más afectados por la adopción de IA. Se usan algoritmos avanzados para automatizar ataques como el phishing y el fraude financiero o bancario.

En 2021, Europol, en colaboración con autoridades policiales de diversos países de la Unión Europea, desmanteló una red internacional de ciberdelincuentes conocida como BazarCall, que utilizaba IA para llevar a cabo ataques de phishing a gran escala contra empresas europeas y de otros lugares. Esta red empleaba IA y técnicas avanzadas de ingeniería social para analizar comportamientos en línea, segmentar víctimas potenciales y enviar correos electrónicos fraudulentos altamente personalizados a empleados de alto nivel.

BazarCall centraba sus ataques en organizaciones del sector financiero y de tecnología, donde los correos electrónicos suplantaban a proveedores o directivos de confianza, instando a los empleados a realizar transferencias bancarias o a abrir documentos adjuntos que contenían malware. El objetivo principal era obtener acceso a información confidencial y ejecutar transferencias fraudulentas de dinero.

Uno de los casos más destacados fue el ataque a la empresa tecnológica Eurofins Scientific, líder global en análisis de laboratorio, que sufrió pérdidas millonarias debido a transferencias bancarias ilegales facilitadas por la IA de los atacantes. La red también se dirigió a empresas de otros sectores clave en Alemania, Francia, Países Bajos y el Reino Unido, logrando infiltrarse en sistemas corporativos a través de correos electrónicos diseñados para evadir sistemas tradicionales de detección de fraudes.

El reporte de Interpol sobre Evaluación del fraude financiero[81], indica que ha habido casos recientes en países miembros de ese organismo en los que se han generado fotografías deepfake para abrir cuentas bancarias en línea con el fin de expandir las redes de muleros que prestan sus servicios al crimen organizado. Dicho reporte señala que existe evidencia reciente de que grupos criminales latinoamericanos como Commando Vermelho, Primeiro Comando da Capital (PCC) y Cartel Jalisco Nueva Generación (CJNG) de México están involucrados en la comisión de fraude financiero.

El lavado de dinero mediante el empleo de IA también está a la orden, permitiendo a las organizaciones criminales ocultar el origen ilícito de sus ingresos mediante transacciones automatizadas y complejas redes financieras que evaden los controles tradicionales.

En 2023, las autoridades italianas y alemanas, en colaboración con Europol, desmantelaron una sofisti-

80 Caldwell, M., Andrews, J. T. A., Tanay, T., & Griffin, L. D. (2020). AI-enabled future crime. Crime Science, 9(1). https://doi.org/10.1186/s40163-020-00123-8

81 Evaluación mundial de INTERPOL sobre la amenaza que plantean las estafas (2023), Interpol.

cada red criminal transnacional dedicada al lavado de dinero que empleaba IA y tecnologías avanzadas para mover fondos ilícitos a través de criptomonedas. Esta organización utilizaba algoritmos de IA para identificar las rutas más seguras y rápidas para transferir dinero entre distintas cuentas de criptomonedas, ocultando el origen ilícito de los fondos.

El esquema consistía en una combinación de lavado de dinero tradicional y el uso de plataformas de criptomonedas, aprovechando los vacíos regulatorios en ciertas jurisdicciones. La red utilizaba herramientas de IA para analizar grandes volúmenes de transacciones financieras y detectar vulnerabilidades en los sistemas de monitoreo de lavado de dinero, evitando los controles tradicionales de las autoridades bancarias. Este enfoque les permitía mover fondos entre distintas jurisdicciones en cuestión de minutos, dificultando la detección y el seguimiento del dinero por parte de las agencias reguladoras.

Uno de los casos más destacados de esta operación ocurrió en Italia, donde la red lavaba dinero a través de plataformas de criptomonedas como Binance y Kraken, facilitando transferencias rápidas hacia países con regulaciones laxas o inexistentes sobre el uso de criptodivisas. Finalmente se incautaron activos por valor de más de 40 millones de euros en criptomonedas y propiedades en varias partes de Europa. Esta operación, denominada "Operazione Colossus", fue coordinada por Guardia di Finanza en Italia y la Bundeskriminalamt (BKA) en Alemania, con el apoyo técnico de Europol y expertos en delitos financieros internacionales.

Por último, la negociación algorítmica es entendida como un proceso de ejecución de órdenes por medio instrucciones de negociación automatizadas y preprogramadas para tener en cuenta variables como el precio, el momento y el volumen de ejecución. A su vez, la llamada negociación de alta frecuencia (high frequency trading o "HFT" por sus siglas en inglés) es un subtipo de negociación algorítmica que sigue estrategias de inversión soportadas por modelos matemáticos complejos y que se centra en aprovechar las ineficiencias del mercado y los movimientos de precios a corto plazo.

Los sistemas HFT llevan a cabo un gran número de transacciones de compra y/o venta, de volumen reducido y en periodos muy cortos de tiempo, con estrategias previamente diseñadas para ejecutar operaciones en cuestión de microsegundos y aprovechar así las pequeñas fluctuaciones de precios que se dan en esos periodos. Esto ha ayudado a la aparición de flash crashes: caídas repentinas e inesperadas en un periodo muy corto, a las que sigue una recuperación rápida del precio de los valores o activos en un mercado. Esto produce un aumento de la volatilidad y gran incertidumbre en los mercados financieros[82]. El HFT se ha convertido en una estrategia comercial prominente en los mercados financieros actuales que está ganando peso a medida que se perfeccionan sus sistemas.

La integración de la IA en el HFT ha revolucionado los mercados financieros ofreciendo beneficios como decisiones más rápidas y precisas, y mayor eficiencia y gestión del riesgo. Sin embargo, como

plantean García Pedroviejo y Marina[83] se plantean desafíos éticos y regulatorios debido a la complejidad y diversidad de los algoritmos de IA. La preocupación radica en la posibilidad de reacciones en cadena, la acentuación de la volatilidad del mercado y la posible manipulación de los precios y del mercado a través de ejecución de estrategias ilícitas dirigidas a la manipulación del mercado, como el spoofing y el layering como sostiene Januário[84]. En Brasil, la BM&F-Bovespa Supervisão de Mercados define el layering como «una práctica abusiva que crea liquidez artificial en el libro de activos a través de capas de ofertas en niveles de precios sucesivos con el objetivo de influir en los inversores para superar la barrera creada por la capa y generar negocios en el lado opuesto del libro» y el spoofing como «una práctica abusiva que crea liquidez artificial con ofertas de tamaño fuera del estándar del libro de órdenes con el objetivo de influir en los inversores para superar la oferta artificial y generar negocios en el lado opuesto del libro»[85]

Como cuenta este autor, en los Estados Unidos, ganó notoriedad el caso *USA v. Coscia*, juzgado en 2015. El Acusado era socio y gerente de una sociedad de inversión que adoptó HFT y implementó un algoritmo que permitía enviar y cancelar ofertas en muy poco tiempo. Con esto, ganó aproximadamente 1,4 millones de dólares en tan solo 10 semanas, con la compra y venta de contratos en 10 mercados del Grupo CME y 3 mercados del *ICE Future Europe Exchange*. El 26 de octubre de 2015, fue sentenciado a 36 meses de prisión y dos años de libertad supervisada[86].

Una segunda estrategia que se puede mencionar es la llamada quotte stuffing. Esta es la introducción/cancelación rápida y sucesiva de grandes cantidades de órdenes, provocando volatilidad en el mercado y congestionando el sistema, dificultando así la actuación y reacción de otros traders ante la gran cantidad de información producida

En materia de mercado y consumidores, deben encuadrarse las infracciones penales relativas a la propiedad intelectual e industrial. Recientemente se han publicado distintas noticias[87] [88] de que OpenAI podría enfrentarse a una posible demanda de la actriz Scarlett Johansson después de que ésta afirmara que la voz de Sky, el chat-

82 HERNANDO CUÑADO, J. (2023, 29 de junio). ¿Llegará la inteligencia artificial a controlar los mercados financieros? The Conversation. https://theconversation.com/llegara-la-inteligencia-artificial-a-controlar-los-mercados-financieros-205942

83 JOSEFINA GARCÍA, JON MARINA Uso de inteligencia artificial en el mercado de valores (high-frequency trading). Pérez Llorca Techlaw 2024. https://www.perezllorca.com/wp-content/uploads/2024/03/04-TechLaw-IA_IA-y-Sectores-Regulados.pdf

84 Januário, Túlio Felippe Xavier (2024). Manipulacion de mercado y nuevas tecnologias: el caso de las negociaciones de alta frecuencia (High Frequency Trading) VVAA, CALAZA, FONTESTAD Y SUAREZ, Paideia : Perspectivas jurídico-procesales en un mundo digital cambiante. COLEX

85 Januário (2024) con cita de Torres, Marcos José Rodrigues et. al., 2016. Painel: Monitoração de ofertas — Spoofing e layering Workshop sobre Monitoração de Práticas Abusivas de Ofertas, de Prevenção à Lavagem de Dinheiro e de Controles Internos de Suitability. En: BM&FBovespa Supervisão de Mercado — BSM. Disponible en: https://www.bsmsupervisao.com.br/assets/file/noticias/Monitoraca_Ofertas.pdf [consulta: 04.07.2023]; Costa, Isac Silveira da, 2018. High frequency trading..., op. cit., p. 216-217.

86 Costa, Isac Silveira da, 2018. High frequency trading..., op. cit., p. 230-232. Ver también: Sousa, Susana Aires de, 2020. «Não fui eu, foi a máquina»: teoria do crime, responsabilidade e inteligência artificial. En: Rodrigues, Anabela Miranda (coord.). A inteligência artificial no direito penal. Coimbra: Almedina, 2020, p. 59-94, p. 64.

87 Hart, R. (2024, mayo 4). El conflicto entre Scarlett Johansson y OpenAI podría generar una guerra de las celebridades contra las empresas de IA. Forbes Argentina. Consultado el 1 de junio de 2024. Disponible en: https://www.forbesargentina.com/innovacion/javier-milei-entusiasmo-ceos-apple-google-meta-openai-empresarios-software-locales-esperan-una-nueva-era-gracias-ia-n53767

88 Pérez Colomé, J. (2024, mayo 21). Scarlett Johansson no permitió que ChatGPT usara su voz, pero OpenAI lo hizo igualmente: "Me enfadé, no podía creerlo". El PAIS. Consultado el 1 de junio de 2024. Disponible en: https://elpais.com/tecnologia/2024-05-21/scarlett-johansson-no-permitio-que-chatgpt-usara-su-voz-pero-openai-lo-hizo-igualmente-me-enfade-no-podia-creerlo.html

bot del fabricante de ChatGPT, ahora retirado del mercado, sonaba inquietantemente parecida a ella misma.

Esta podría ser la segunda vez que Johansson ejercite acciones legales por lo mismo. Hace unos meses, ya saltó la noticia. Johansson apareció en un anuncio de 22 segundos publicado en X/Twitter por una aplicación de generación de imágenes de inteligencia artificial llamada Lisa AI: 90s Yearbook & Avatar.

El anuncio, reseñado por Variety[89], comienza con un antiguo clip de Johansson detrás de escena de "Black Widow" de Marvel.

3.2. SUPLANTACIÓN DE IDENTIDAD

Organizaciones criminales en la región utilizan herramientas de IA para crear audios y videos falsos convincentes, imitando voces o imágenes de familiares de las víctimas. Esta tecnología ha sido usada para cometer fraudes y extorsiones. Por ejemplo, en países como Perú y Argentina, los delincuentes están utilizado deepfakes para simular la voz de familiares secuestrados y exigir pagos de rescate y extorsiones[90].

En este rubro, uno de los fraudes más emblemáticos fue el acontecido a la empresa inglesa de diseño e ingeniería Arup en mayo de 20024 a través del cual los criminales utilizaron video conferencias pasadas de los ejecutivos de la empresa para entrenar a la herramienta de IA y recrear un escenario en el que director financiero junto con otros empleados solicitan hacer distintas depósitos y transferencias bancarias a uno de sus empleados de sus oficinas en Hong-Kong. El empleado accedió a hacer las transferencias y se estima que la empresa reporto una pérdida de alrededor de $25.6 millones de dólares[91]. No existe algún caso similar reportado en algún país de ALC.

Herramientas de IA también han sido empleadas para cometer estafas relacionadas con migrantes desaparecidos, los criminales generan imágenes falsas de sus víctimas para convencer a sus familiares o parientes de que han sido secuestrados a cambio de obtener el pago de rescates[92]. En la frontera entre EE.UU. y México se reporta que grupos criminales utilizan y manipulan imágenes para estafar a las familias de migrantes desaparecidos.

Se ha hecho viral un video de un joven que muestra como a través de WhatsApp su mamá le envía una serie de audios, pidiéndole dinero. Toda esta situación es grabada con la mamá al lado y en los audios se evidencian errores en la dicción y desde los audios le insisten que no lo haga a la cuenta de siempre porque no tiene la tarjeta para retirar el dinero de esa cuenta habitual y le envían los datos de transferencia para otra personaEsta modalidad de ataque se conoce como visihing, en la que se usa la IA para clonar la voz de una persona para suplantar su identidad y enviar un mensaje falso para solicitar dinero o datos personales, todo justificado mediante una petición cercana a la víctima[93].

89 Shanfeld, E. (2023, noviembre 1). Scarlett Johansson Takes Legal Action Against AI App That Ripped Off Her Likeness in Advertisement. Variety. Consultado el 1 de junio de 2024. Disponible en: https://variety.com/2023/digital/news/scarlett-johansson-legal-action-ai-app-ad-likeness-1235773489/

90 El Comercio, «Clonan voces de personas con IA para estafar o fingir secuestros: al menos 55 casos en Perú», 16 de julio 2023, en: https://elcomercio.pe/lima/clonacion-de-voz-para-estafar-con-inteligencia-artificial-como-funciona-esta-modalidad-y-que-recomendaciones-seguir-inseguridad-deepfake-ciberdelincuencia-hackers-secuestros-noticia/

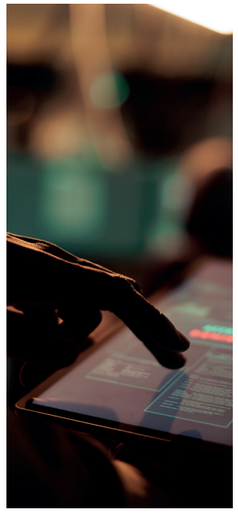
91 FORTUNE, «A deepfake 'CFO' tricked the British design firm behind the Sydney Opera House in $25 million scam», 17 de mayo de 2024 en: https://fortune.com/europe/2024/05/17/arup-deepfake-fraud-scam-victim-hong-kong-25-million-cfo/

92 ASMANN, Parker. InSight Crime, «4 Ways AI is Shaping Organized Crime in Latin America» 15 de agosto de 2018, en: https://insightcrime.org/news/four-ways-ai-is-shaping-organized-crime-in-latin-america/

93 RIOS, Juan. Infobae. (2024, 29 de octubre). Cuidado en WhatsApp: copian la voz de tu mamá, usan IA para crear la estafa y roban dinero del banco. Infobae. https://www.infobae.com/tecno/2024/10/29/cuidado-en-whatsapp-copian-la-voz-de-tu-mama-usan-ia-para-crear-la-estafa-y-roban-dinero-del-banco/

En México, delincuentes lograron suplantar la imagen del empresario Carlos Slim que fue manipulada con IA y utilizada para promocionar esquemas de inversión entre la población mexicana a través de un enlace phishing que prometía obtener grandes ganancias diarias[94].

En el área de transacción e intercambio de criptomonedas, una empresa de ciberseguridad reporta acerca de una innovadora herramienta de IA utilizada por organizaciones criminales que tiene la capacidad de evadir los sistemas de autenticación de dos factores (2FA) mediante el uso de deepfakes, y a través de la identidad falsa creada por los delincuentes, intentan utilizarla en el proceso de autenticación con las empresas intercambiadoras de criptomonedas para posteriormente llevar a cabo transacciones relacionadas con criptomonedas[95].



3.3 RANSOMWARE COMO SERVICIO (RASS)

La IA puede ser utilizada para optimizar la eficacia de los ataques de ransomware. Los algoritmos pueden alterar dinámicamente el código del ransomware para evadir la detección de los sistemas de ciberseguridad, identificar archivos valiosos para cifrarlos e incluso decidir las cantidades necesarias para el pago del rescate. Existe evidencia de que el grupo de ransomware 'BlackCat', utilizo en 2023 técnicas y herramientas de IA para eludir las defensas tradicionales de ciberseguridad y propagarse rápidamente a través de las redes, cifrando los datos de las víctimas antes de solicitar el pago de los rescates[96].

De acuerdo con TRM Labs, los operadores de ransomware, se aprovechan y utilizan cada vez más la IA para mejorar la eficiencia y el impacto de sus ataques, como por ejemplo para la automatización de campañas de ransomware, lo que permite a los grupos delincuenciales generar correos electrónicos phishing más convincentes, identificar vulnerabilidades en los sistemas de forma más eficiente y optimizar los ataques de ransomware[97].

3.4. PHISHING E INGENIERÍA SOCIAL

Los delincuentes están utilizando modelos avanzados de IA basados en ChatGPT y chatbots similares para generar correos electrónicos y mensajes de texto de phishing más convincentes y a gran escala lo que permite que los ataques de ingeniería social dirigidos a víctimas sean más sofisticados y difíciles de detectar. De acuerdo con la empresa de seguridad Keeper se ha observado un aumento de 51% en ataques de phishing impulsados por la IA[98]. Ejemplos se pueden encontrar en FraudGTP y Metasploit.

FraudGPT es un producto vendido en la web oscura y Telegram que funciona de manera similar a ChatGPT pero crea contenido para facilitar los ataques cibernéticos. Miembros del equipo de investigación de amenazas de Netenrich lo identificó por primera vez y lo vio anunciado en julio de 2023[99]

94 El Economista, «Nueva plataforma de inversiones de Slim es falsa; utilizan IA para defraudar», 26 de noviembre de 2023, en: https://www.eleconomista.com.mx/finanzaspersonales/Nueva-plataforma-de-inversiones-de-Slim-es-falsa-utilizan-IA-para-defraudar-20231126-0023.html

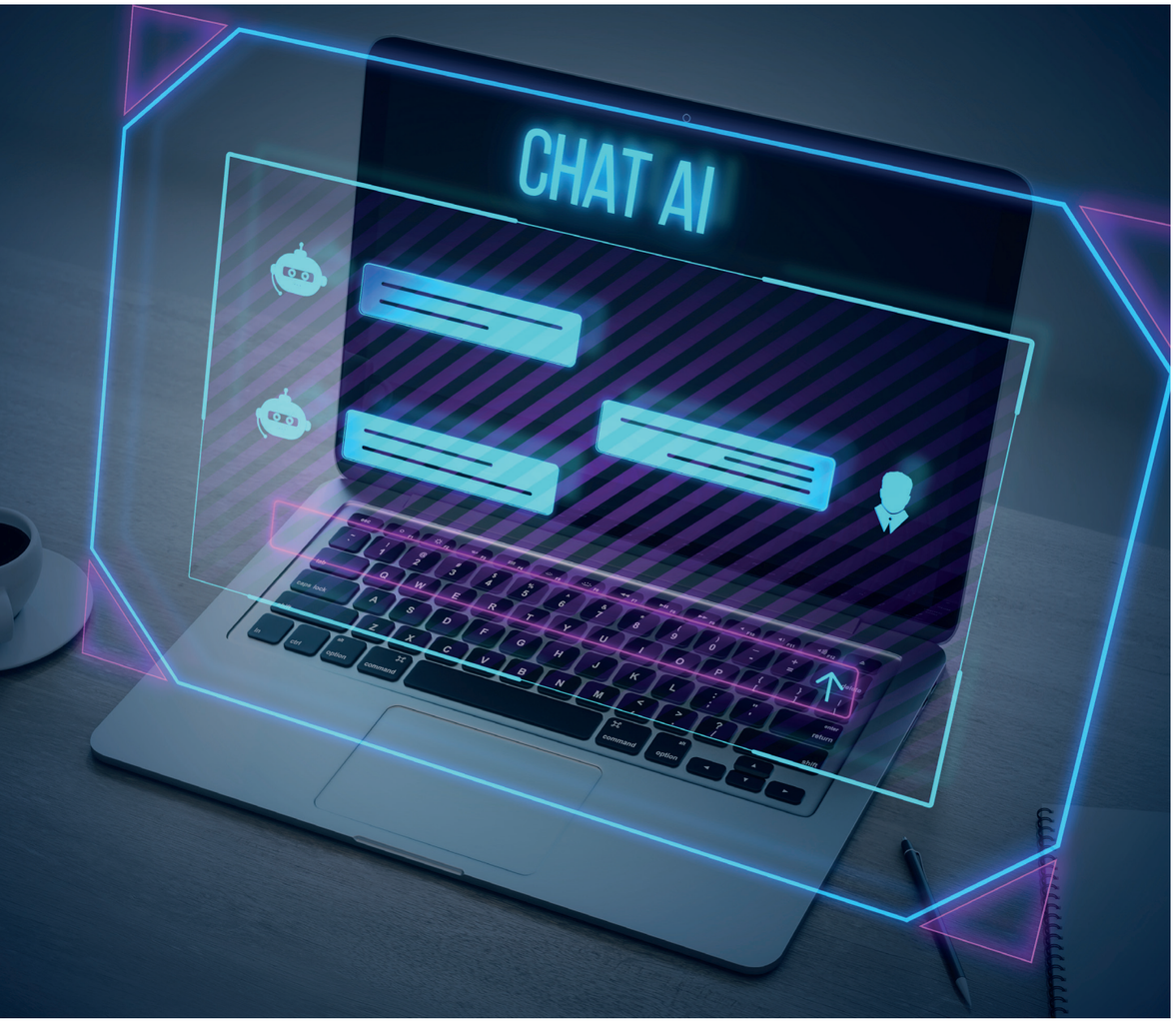
95 TechInformed, «Deepfake cybercrime tool threatens crypto exchanges» 15 de octubre de 2024 en: https://techinformed.com/deepfake-cybercrime-tool-threatens-crypto-exchanges/

96 Center for Internet Security (CIS), «Breaking down the BlackCat Ransomware Operation» en: https://www.cisecurity.org/insights/blog/breaking-down-the-black-cat-ransomware-operation

97 TRM, «Ransomware in 2024: Latest Trends, Mounting Threats, and the Government Response», 11 de octubre de 2024, en: https://www.trmlabs.com/post/ransomware-in-2024-latest-trends-mounting-threats-and-the-government-response

98 Keeper, «Cómo hace la IA para que los ataques de phishing sean más peligrosos», 13 de septiembre de 2024 en:https://www.keepersecurity.com/blog/es/2024/09/13/how-ai-is-making-phishing-attacks-more-dangerous/

99 AMOS, Zac. (2023/08/11) ¿Qué es FraudGPT? HackerNoon. https://hackernoon.com/lang/es/que-es-fraudegpt



Grupos delincuenciales están utilizando sistemas de IA para optimizar la selección y el reclutamiento de jóvenes y menores de edad para realizar actividades ilícitas como estafas románticas, estafas de inversión en criptomonedas, extorsiones y secuestros.

De igual manera, con la inteligencia artificial junto con la ingeniería social se puede intentar influir en un proceso electoral mediante ejércitos de bots.

Otra posibilidad para la comisión de delitos utilizando LLMs puede ser el denominado *prompt injection.* Según Kosinski y Forrest[100], una inyección de prompt es un tipo de ciberataque contra grandes modelos de lenguaje (LLM). Los hackers disfrazan entradas maliciosas de prompts legítimos, manipulando los sistemas de IA generativa (GenAI) para que filtren datos confidenciales, difundan información errónea o cosas peores.

Dentro de la modalidad de estafas, una de ellas posible mediante estas tecnologías como el *deepfake,* sería la estafa procesal, de talmanera que se intente engañar al juez con pruebas falsas en un juicio.

100 KOSINSK M. Y FORREST A. (26 de marzo de 2024. Prompt injection. IBM Research. Recuperado de https://www.ibm.com/es-es/topics/prompt-injection

3.5. TRÁFICO DE PERSONAS: RECLUTAMIENTO Y EXPLOTACIÓN ONLINE

La IA también está revolucionando el tráfico de personas, particularmente en el reclutamiento y explotación de víctimas a través de plataformas digitales. Las redes criminales utilizan IA para identificar a personas vulnerables y dirigirse a ellas con mayor precisión, explotando patrones de comportamiento online tal y como se manifiesta en varios informes recientes sobre trata de personas en Europa de UNODC.

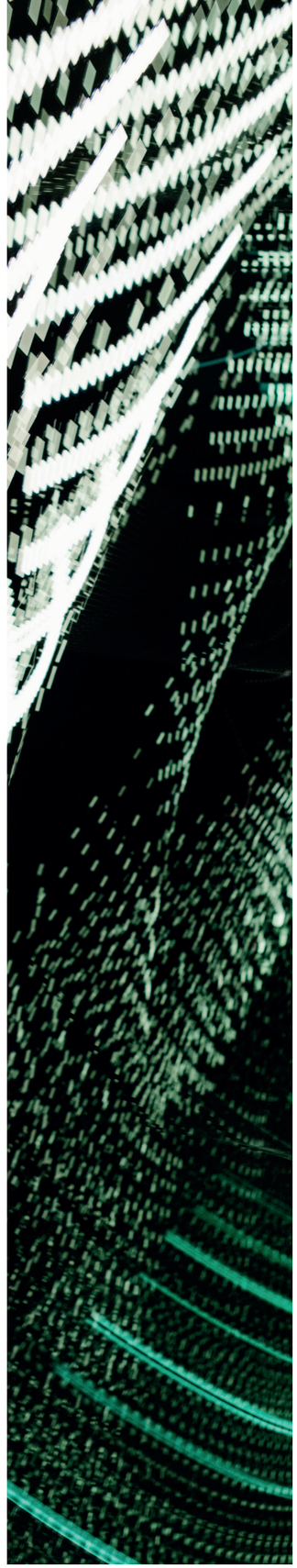
Un ejemplo es el ocurrido en 2022, donde la Policía Nacional de España y la Gendarmería Nacional de Francia, en una operación conjunta con Europol, desmantelaron una sofisticada red de tráfico de personas que utilizaba inteligencia artificial para reclutar y explotar a mujeres jóvenes de Europa del Este y el norte de África. Esta red criminal, especializada en la explotación sexual, utilizaba algoritmos avanzados de IA para identificar, manipular y atraer a víctimas a través de redes sociales y plataformas de citas online.

Los algoritmos de IA empleados por esta red criminal analizaban los perfiles psicológicos y socioeconómicos de las posibles víctimas, segmentando a mujeres jóvenes con situaciones vulnerables, como pobreza, desempleo o problemas familiares. Los sistemas de IA podían filtrar estos datos a partir de interacciones online, comentarios en redes sociales y la actividad de búsqueda de las víctimas. A través de esta información, los criminales personalizaban mensajes y ofrecían falsas oportunidades de trabajo en Europa Occidental, como empleos en la industria de la moda, hostelería o trabajos domésticos.

Una vez establecida la confianza, las víctimas eran convencidas de viajar a países como España y Francia, donde, al llegar, se les confiscaban sus documentos y eran obligadas a trabajar en redes de prostitución forzada. Las víctimas provenían mayoritariamente de Rumanía, Bulgaria, Marruecos y Argelia, países con altos niveles de vulnerabilidad para este tipo de explotación.

La red también empleaba IA para eludir los sistemas de moderación de contenido de las redes sociales y plataformas de publicidad, ocultando sus anuncios ilegales de servicios sexuales. Estos algoritmos podían modificar las palabras clave y el contenido de los anuncios para evitar ser detectados por los filtros automatizados de las plataformas. Publicaban en páginas de clasificados y redes sociales, como Facebook y Instagram, ofreciendo servicios que, en realidad, eran portadas para la explotación sexual.

Los anuncios eran cambiados con frecuencia, adaptando el lenguaje a las regulaciones locales y utilizando imágenes manipuladas para evitar la detección por los sistemas de control de contenido, lo que hacía extremadamente difícil para las autoridades rastrear la actividad criminal en tiempo real.



3.6. DELITOS DE ABUSO Y EXPLOTACIÓN SEXUAL

Los sistemas de IAG tienen la capacidad de generar y transformar texto e imágenes de niños y adolescentes con contenidos de abuso sexual y explotación a gran escala y precisión que resultan muy difícil de identificar y diferenciar[101]. De acuerdo con NCMEC, la prevalencia de contenidos de abuso y explotación sexual de menores generados a través de IAG resulta sumamente difícil de estimar y desde 2023 esa organización ha empezado a incluir una clasificación con estadísticas respecto al número de reportes facilitados con contenido de abuso sexual y explotación generados por sistemas de IAG a través de CyberTipline[102]. Además, se ha empezado a detectar la utilización de robots sexuales para el abuso a menores, ya sean de forma remota o bajo control de la inteligencia artificial o sistemas autónomos, o de forma más mecánica. No obstante, su construcción puede ser realizada mediante impresoras 3D y planos desarrollados por IA u personas tanto con fines lícitos o ilícitos[103] [104].

3.7. CONDUCTAS DE CIBERVIOLENCIA

Existen aplicaciones de IA gratuita y de pago que permiten generar, reproducir y editar imágenes de mujeres mostrándolas con cuerpos desnudos mediante el uso de deepfakes[105]. De acuerdo con WIRED, el contenido explícito no consensual creado por deepkafe bots a través de Telegram ha aumentado exponencialmente. Dicha editorial reporta que existen al menos 50 bots actualmente activos que atienden

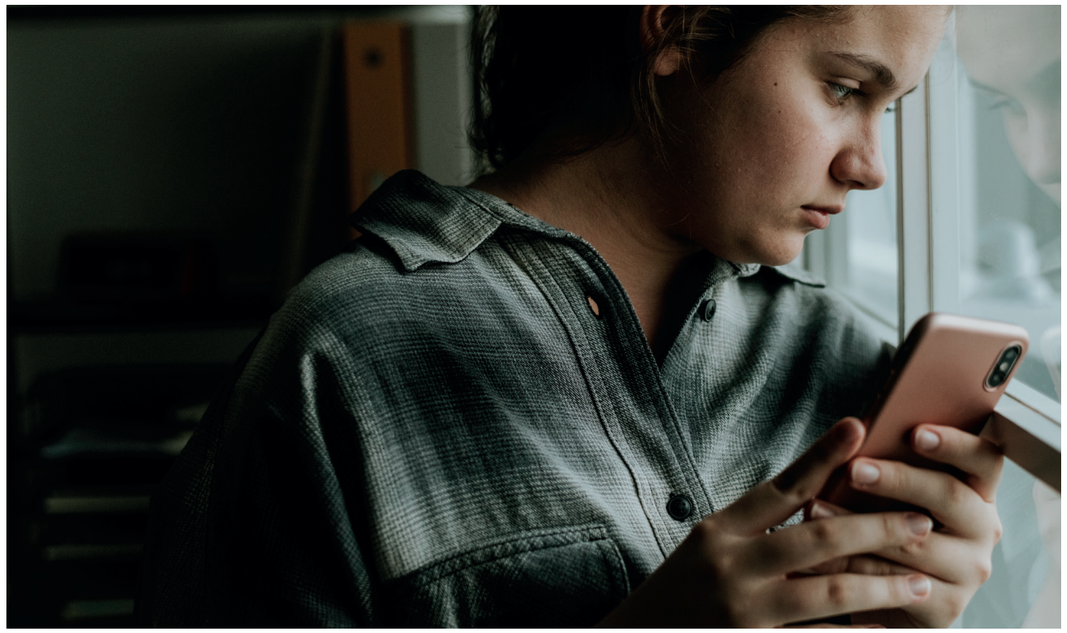
101 Para una descripción de los tipos de imágenes y contenidos de abuso y explotación generados a través de IA, ver: Centre for Artificial Intelligence and Robotics at the United Nations Interregional Crime and Justice Reserach Institute (UNICRI), «Generative AI. A New Threat for Online Child Sexual Exploitation and Abuse», 2024, pp. 9-13 en: https://unicri.it/News-Generative-AI-Threat-Child-Sexual-Exploitation-Abuse

102 «Generative AI. A New Threat for Online Child Sexual Exploitation and Abuse», *op. cit*, p. 13.

103 El pedófilo que estaba construyendo un robot sexual de un menor Disponible em: https://www.elpatagonico.com/el-pedofilo-que-estaba-construyendo-un-robot-sexual-un-menor-n5992054

104 Durán San Juan, Isabela. (2024). ¿Cómo los robots y la inteligencia artificial transformarán las relaciones sexuales del futuro? Infobae. Disponible en: https://www.infobae.com/tecno/2024/04/12/como-los-robots-y-la-inteligencia-artificial-transformaran-las-relaciones-sexuales-del-futuro/

105 Ver: https://nudify.info/download-apps-like-deepnude-alternatives/



a más de 4 millones de solicitudes en forma mensual en la plataforma Telegram y que son utilizados para atacar a miles de mujeres y niñas en todo el mundo[106].

Si bien las conductas de ciberviolencia son cometidas generalmente por círculos sociales cercanos a las víctimas, tienen el potencial de crear un grave daño e impacto psicológico en mujeres, niñas y grupos vulnerables pertenecientes al género LGBT. Esto es lo que principalmente se denomina Violencia facilitada por la tecnología contra las mujeres y las niñas y que ya tiene su propia denominación debido a su actual impacto (Technology-facilitated violence against women and girls (TFVaWG).

En Perú, se dio un caso en agosto de 2023 en un colegio privado donde los alumnos generaron y editaron imágenes deepfake de sus compañeras con cuerpos desnudos y las compartieron y comercializaron fuera del círculo escolar[107]. Este tipo de actividades puede tener repercusiones futuras de revictimización ya que las imágenes pueden terminar circulando en la DarkNet y posteriormente pueden ser utilizadas con propósitos delictivos por acosadores, extorsionadores y pedófilos.

En Argentina, se reporta un caso en octubre de 2024 en un colegio de la localidad de San Andrés en el partido bonaerense de San Martín en donde un alumno menor de edad utilizó aplicaciones de IA para manipular fotos de sus compañeras obtenidas de redes sociales para convertirlas en desnudos y posteriormente comercializar las imágenes editadas a través de la aplicación Discord. Se reporta que fueron 22 alumnas víctimas de este delito, todas ellas menores de 18 años[108]. El Caso Almendralejo en España (2023) es oto caso de manipulación de fotografías y vídeos mostrando menores desnudos utlizando aplicaciones y herramientas de IA.

Este tipo de actividades puede tener graves repercusiones futuras de revictimización ya que las imágenes pueden terminar circulando en la DarkNet y posteriormente pueden ser utilizadas con propósitos delictivos por acosadores, extorsionadores y pedófilos.

106 WIRED, «Millions of People Are Using Abusive AI 'Nudify' Bots on Telegram» 15 de Octubre de 2024, en: https://www.wired.com/story/ai-deepfake-nudify-bots-telegram/#intcid=_wired-right-rail_c0163b39-a6b8-486b-8f60-b788683ebd84_popular4-1-reranked-by-vidi

107 Infobae, «Chorrillos: Escolares que alteraron fotos de compañeras con IA y las comercializaron no fueron expulsados, 29 de agosto de 2023 en: https://www.infobae.com/peru/2023/08/29/chorrillos-escolares-que-alteraron-fotos-de-sus-companeras-con-ia-para-venderlas-no-fueron-expulsados/

108 Clarín, «Escándalo en un colegio de San Martín: denuncian a un alumno que vendía fotos manipuladas con IA de sus compañeras desnudas» 14 de octubre de 2024, en: https://www.clarin.com/policiales/escandalo-colegio-san-martin-denuncian-alumno-vendia-fotos-manipuladas-ia-companeras-desnudas_0_hqEbfd1Plm.html?srsltid=AfmBOoqd3x3x-znnG_za1bJ_yPAkS9QnoX6HxZjyMUm7TUIahZCxAckV

BLOQUE 4: UTILIZACIÓN DE HERRAMIENTAS DE IA POR PARTE DE LAS INSTITUCIONES DE JUSTICIA Y SEGURIDAD

4.1. LA IA EN LAS INSTITUCIONES DE JUSTICIA

La IA ha surgido como una herramienta de apoyo clave para las instituciones de justicia, ofreciendo soluciones que permiten no solo el análisis más rápido de grandes volúmenes de información, sino también la optimización de la toma de decisiones judiciales. Este apartado del estudio se centra en cómo los tribunales, jueces, fiscales y otros actores jurídicos están utilizando la IA en la lucha contra el crimen organizado, particularmente en los países de América Latina, donde las redes delictivas son altamente sofisticadas. Por otra parte, la IA también tiene el potencial de reducir la carga de trabajo de los tribunales al automatizar tareas repetitivas y facilitar el acceso a información legal, lo que también es clave en la lucha contra el crimen organizado, ya que permite una gestión más eficaz de los casos complejos que estas redes criminales presentan.

Tal es la relevancia que, en el ámbito de la Unión Europea, la Comisión Europea para la Eficacia de la Justicia (CEPEJ) está avanzado en sus esfuerzos por integrar la IA en los sistemas judiciales, con énfasis en la creación de mecanismos de certificación para herramientas y servicios de IA. Durante la 42ª reunión plenaria de junio de 2024, se destacaron varios proyectos relacionados con la IA y la ciberjusticia, incluidas las contribuciones de los grupos de trabajo de la CEPEJ sobre la calidad de la justicia (CEPEJ-GT-QUAL) y la ciberjusticia (CEPEJ-GT-CYBERJUST).

En relación con lo anterior, destacamos dos puntos importantes. Por un lado, CEPEJ ha desarrollado la Carta Ética sobre la IA en los Sistemas Judiciales, que establece principios clave para la integración de la IA en los sistemas judiciales. Estos principios incluyen la transparencia, la no discriminación y el respeto a los derechos fundamentales, asegurando que las aplicaciones de IA respeten la equidad en los procesos judiciales y que las decisiones automatizadas sean auditables y comprensibles. Es decir, tiene como objetivo guiar tanto a los desarrolladores de tecnología como a los legisladores y profesionales de la justicia, promoviendo la implementación de estos principios en varios países europeos, organizando capacitaciones y eventos para crear conciencia sobre los riesgos y beneficios del uso de IA en la justicia.

Por el otro, está trabajando en un mecanismo de evaluación de impacto para los productos de IA utilizados en los sistemas judiciales europeos, que tiene como objetivo garantizar que las herramientas desarrolladas, tanto por el sector público como por el privado, cumplan precisamente con las directrices éticas establecidas en la Carta Ética sobre la IA en la Justicia. Esta iniciativa busca asegurar que los algoritmos y herramientas tecnológicas respeten los principios de transparencia, imparcialidad y control humano, fundamentales en el ámbito judicial.

Además de lo anterior, entre los proyectos actuales, la CEPEJ ha trabajado en la creación de un *Centro de Recursos sobre Ciberjusticia e IA*, que proporciona información sobre herramientas tecnológicas aplicadas en el ámbito judicial. Este centro no solo facilita el intercambio de buenas prácticas entre los Estados miembros, sino que también proporciona orientación sobre los riesgos y beneficios del uso de la IA en los tribunales. El Centro de Recursos se centra en los

sistemas del sector público, aplicados por el poder judicial o de importancia para él. Esto puede incluir modelos académicos disponibles públicamente o sistemas de IA de propósito general (por ejemplo, ChatGPT, Co-Pilot), pero no incluye sistemas orientados a abogados o bufetes de abogados (LegalTech). Las entradas se recogen a través de los miembros de la Red Europea de Ciberjusticia (ECN) del CEPEJ. La red está formada por personas de casi todos los Estados miembros del Consejo de Europa y observadores, encargados de la digitalización de los sistemas judiciales nacionales. La información recopilada es discutida y categorizada por el Consejo Asesor de Inteligencia Artificial (AIAB) del CEPEJ.

También ha establecido el *Consejo Asesor de IA* (AIAB), que proporciona orientación experta sobre la operacionalización de los principios de la Carta. Actualmente, este consejo está trabajando en una herramienta de evaluación que permitirá a las autoridades judiciales evaluar la conformidad de sus sistemas de IA con los principios éticos establecidos. El proceso de evaluación para la operacionalización, tiene como objetivo hacer operativa la Carta de la CEPEJ al proporcionar un conjunto de verificaciones, medidas clave y salvaguardias que los tomadores de decisiones dentro del sistema judicial deben seguir al comprar, diseñar, desarrollar, implementar y/o utilizar IA en los sistemas judiciales. Tiene como base evaluar productos de IA según criterios relacionados con la transparencia, la calidad y la ética. Las herramientas que pasen este filtro supone que cumplir voluntariamente un estándar ético superior, lo que proporciona confianza tanto a los operadores de justicia como a los ciudadanos que interactúan con estas tecnologías. La CEPEJ considera que la creación de un sistema de evaluación independiente y estandarizado para certificar estos productos permitirá garantizar que los sistemas de IA sean confiables y que no comprometan la equidad o los derechos fundamentales de los individuos involucrados en procesos judiciales. El consejo asesor de IA igualmente está trabajando actualmente en un informe anual que ofrecerá un resumen conciso de los resultados del seguimiento continuo de la inteligencia artificial emergente u otras herramientas clave de ciberjusticia aplicadas en los sistemas de justicia pública (en adelante, conjuntamente como herramientas de ciberjusticia), realizado a través del Centro de Recursos sobre Ciberjusticia e Inteligencia Artificial del CEPEJ (en adelante, «el Centro de Recursos» o «el Centro»). Además se encuentran elaborado un documento sobre IA y eficiencia en la justicia. El CEPEJ encargó al CEPEJ-SATURN y al CEPEJ-GT-CYBERJUST la tarea de estudiar los posibles efectos del uso de sistemas de IA en la eficiencia de los tribunales (véase el Programa de Actividades 2024-2025 del CEPEJ) así como se encuentran desarrollando otro sobre IA Generativa en tribunales.

En definitiva, estos esfuerzos buscan garantizar que las soluciones tecnológicas implementadas respeten los derechos fundamentales, conforme a la *Carta Ética Europea sobre el Uso de la IA en los Sistemas Judiciales*, adoptada en 2018[109], así como que subrayan el compromiso de la CEPEJ con la transformación digital del sistema judicial europeo, asegurando que la IA se utilice de manera segura y eficiente en consonancia con los derechos humanos.

Este pasado junio se aprobó la POLITICA DE USO DE LA INTELIGENCIA ARTIFICIAL EN LA ADMINISTRACIÓN DE JUSTICIA en la Secretaría General del CTEAJE[110] destinada a todos los trabajadores de la Administración de Justicia pero que no obstante, tiene que ser ratificada su adopción de manera específica por parte del Consejo General del Poder Judicial (CGPJ), Fiscalía General del Estado (FGE), Comunidades Autónomas con competencias en materia de Justicia y Ministerio de la Presidencia, Justicia y Relaciones con las Cortes (MPJRC).

Asimismo, la UNESCO ha lanzado una consulta abierta sobre nuevas directrices para el uso de la IA en los sistemas judiciales[111], y que tienen por objeto garantizar que las tecnologías de IA se integren en los sistemas judiciales de una manera que defienda la justicia, los derechos humanos y el estado de derecho. Un proyecto de directrices elaborado a raíz de la encuesta de la UNESCO sobre el uso de la inteligencia artificial por parte de los operadores judiciales.

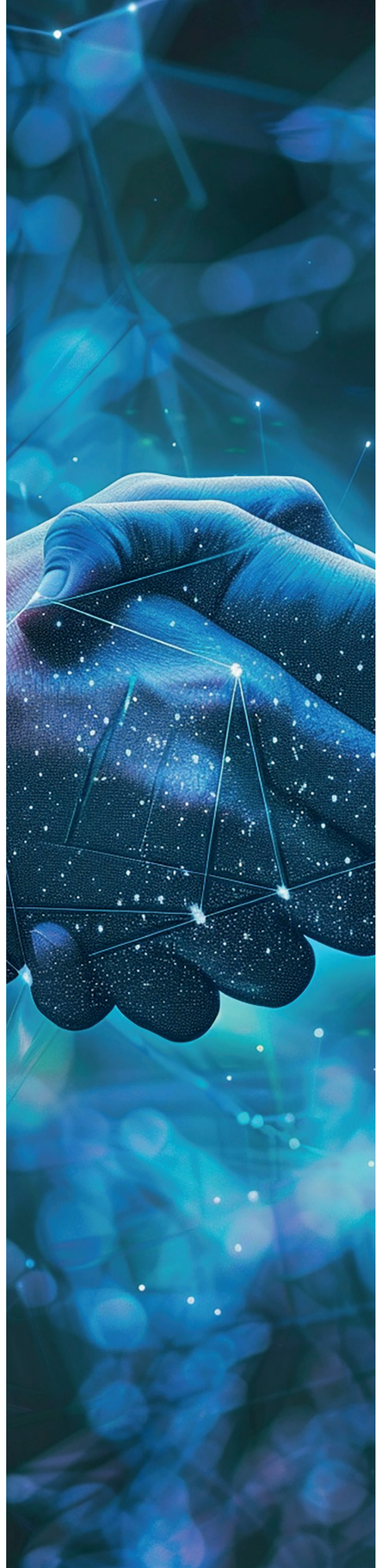
Uno de los trabajos más completos para operadores judiciales desarrollados hasta la fecha es el realizado por UNESCO en su *Kit de herramientas global sobre IA y el Estado de derecho para el poder judicial*[112]. Esta caja de herramientas responde a estas necesidades y proporciona a los actores judiciales (jueces, fiscales, fiscales estatales, abogados públicos, universidades de derecho e instituciones de formación judicial) el

109 Véase, https://protecciondata.es/wp-content/uploads/2021/12/Carta-Etica-Europea-sobre-el-uso-de-la-Inteligencia-Artificial-en-los-sistemas-judiciales-y-su-entorno.pdf

110 https://www.administraciondejusticia.gob.es/documents/7557301/7558184/CTEAJE-NOR-Politica+de+uso+de+la+IA+en+la+AJ+v1.0.pdf/ddc0eda1-950b-e926-b367-be511b16f2f9?t=1721386535984

111 https://unesdoc.unesco.org/ark:/48223/pf0000390781

112 https://unesdoc.unesco.org/ark:/48223/pf0000387331_spa



conocimiento y las herramientas necesarias para comprender los beneficios y riesgos de la IA en su trabajo. El kit de herramientas ayudará a los actores judiciales a mitigar los posibles riesgos de la IA para los derechos humanos al brindar orientación sobre las leyes, principios, normas y jurisprudencia internacional de derechos humanos relevantes que sustentan el uso ético de la IA.

PROYECTOS E INICIATIVAS DE COOPERACIÓN JURÍDICA INTERNACIONAL

En este apartado conviene destacar el **Proyecto INSPECTr** (Intelligence Network and Secure Platform for Evidence Correlation and Transfer), que se desarrolló entre septiembre de 2019 y agosto de 2022. Fue una iniciativa financiada por la Unión Europea en el marco de su programa Horizon 2020 y tuvo como objetivo principal mejorar las capacidades digitales y forenses de las fuerzas de seguridad y las instituciones judiciales mediante el uso de tecnología avanzada.

INSPECTr abordó los desafíos que enfrentan los sistemas de justicia penal en la lucha contra el crimen transfronterizo y organizado, integrando una gama de herramientas tecnológicas de vanguardia para la recopilación, el análisis y el intercambio de información entre diferentes jurisdicciones y que podemos resumir de la siguiente manera:

Análisis de macrodatos (big data): INSPECTr utilizó el análisis de grandes volúmenes de datos para identificar patrones y correlaciones en los flujos de información, facilitando la detección temprana de actividades criminales complejas y organizadas. Las capacidades de procesamiento de datos en tiempo real permitieron a las agencias judiciales y policiales extraer información clave de manera eficiente, incluso en escenarios con grandes cantidades de datos no estructurados, como redes sociales, comunicaciones cifradas o transacciones financieras internacionales. Esto mejoró la capacidad para identificar redes criminales transnacionales, permitiendo una respuesta más rápida y coordinada.

Aprendizaje automático cognitivo: El proyecto también incorporó algoritmos de machine learning que permitieron a las plataformas de inteligencia automatizar la detección de amenazas y patrones de comportamiento delictivo. Al utilizar el aprendizaje automático cognitivo, INSPECTr fue capaz de mejorar el análisis predictivo, ayudando a las instituciones de justicia a prever posibles escenarios delictivos y a optimizar la asignación de recursos. Estas tecnologías permitieron una clasificación más precisa de la información, facilitando la toma de decisiones judiciales más informadas.

Blockchain para la seguridad de la información: La tecnología de cadena de bloques (blockchain) fue clave para garantizar la integridad y trazabilidad de las pruebas en el proceso judicial. Al integrar blockchain en la plataforma, INSPECTr logró crear un sistema que permitía rastrear cada interacción y transferencia de datos entre diferentes actores, asegurando que las pruebas no se alteraran durante su almacenamiento o transmisión. Esto es especialmente relevante en casos de crimen organizado, donde la manipulación de pruebas puede comprometer gravemente los procesos judiciales.

Gracias a esta tecnología, INSPECTr ha permitido a las autoridades mejorar su capacidad para gestionar investigaciones complejas,

mejorar la seguridad jurídica en la transmisión de datos y asegurar que los procedimientos judiciales sean más eficientes y efectivos.

Uno de los aspectos más innovadores del proyecto fue el desarrollo de una plataforma de inteligencia compartida que facilitaba la colaboración transfronteriza entre agencias de justicia penal. El intercambio de datos a nivel internacional se ha considerado fundamental para combatir el crimen organizado, que frecuentemente opera en múltiples países. Sin embargo, este intercambio presentaba desafíos como la compatibilidad entre los sistemas de diferentes jurisdicciones y la protección de la privacidad de los individuos involucrados. La plataforma INSPECTr logró reducir la complejidad y el costo de los intercambios de información, proporcionando un espacio seguro donde las fuerzas de seguridad y los sistemas judiciales pudieran colaborar en la correlación y análisis de pruebas. Un aspecto crucial fue la interoperabilidad de los sistemas judiciales y de seguridad de diferentes países, permitiendo el acceso mutuo a bases de datos, lo que aceleró el tiempo de respuesta ante delitos internacionales. Además, la plataforma permitió el almacenamiento seguro de la información en la nube, minimizando el riesgo de pérdida o manipulación de datos.

En el ámbito de la cooperación jurídica internacional también cobra mucha importancia la traducción e interpretación. En relación con esto, la UE viene financiando diversos proyectos de investigación que ciertamente van a contribuir o están contribuyendo en el empleo de la videoconferencia en materia de cooperación jurídica internacional. Entre ellos, podemos destacar **AVIDICUS (Assessment of Video-Mediated Interpreting in the Criminal Justice System)**, "evaluación de la interpretación mediada por vídeo en el sistema de justicia penal" en español). Se trató de un proyecto en tres etapas centrado en evaluar la fiabilidad y calidad de la interpretación mediada por vídeo (VMI) en procedimientos penales a fin de mejorar la cooperación judicial en Europa, que son las siguientes:

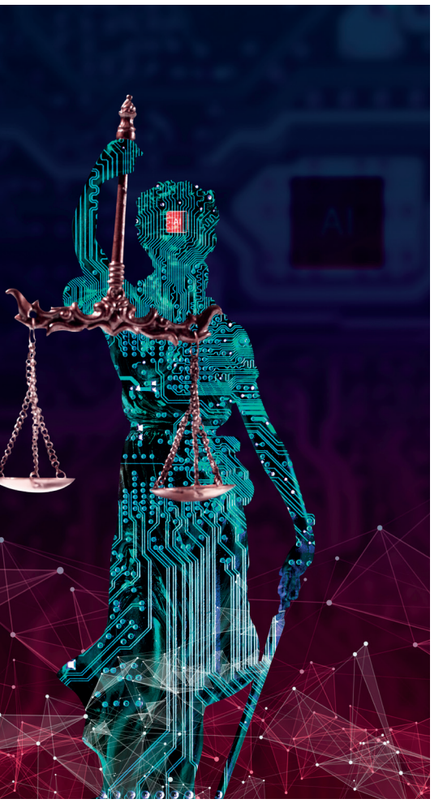
AVIDICUS I (2008-2011): Se centró en investigar los efectos del uso de la videoconferencia en la interpretación judicial y cómo esto impacta la calidad y precisión de la comunicación.

AVIDICUS II (2011-2013): Enfocado en expandir el análisis con estudios comparativos a nivel europeo, explorando el uso de la videoconferencia para la interpretación en múltiples contextos judiciales.

AVIDICUS III (2014-2016): Abordó la implementación práctica de la interpretación mediada por vídeo, proporcionando capacitación específica tanto para intérpretes como para profesionales del derecho, mejorando así las habilidades y la comprensión mutua durante los procedimientos judiciales.

Además, se está trabajando en otros proyectos de inteligencias artificiales que permiten la traducción automática. Es decir, que la traducción se llevaría a cabo por un programa informático sin intervención humana, como es el caso del **programa eTranslation**, impulsado por la Comisión Europea bajo el programa Connecting Europe Facility (CEF). Se trata de una herramienta clave para la traducción automática en procedimientos judiciales transfronterizos, incluidos los juicios penales mediante videoconferencia. Su objetivo principal es eliminar barreras lingüísticas en el mercado digital único de la UE, facilitando la comunicación y la cooperación judicial a través de traducciones automáticas de alta calidad en todos los idiomas oficiales de la Unión. Esta tecnología se está integrando en plataformas judiciales digitales como el Portal Europeo de Justicia para garantizar que las partes involucradas en juicios transfronterizos, incluyendo testigos y acusados, puedan presentar declaraciones en su idioma nativo, con traducciones automáticas inmediatas durante las videoconferencias. Esto es esencial para juicios penales donde las barreras lingüísticas pueden ser un obstáculo en la toma de evidencia o testimonios y así, contribuir a que las causas criminales complejas sean enjuiciadas.

El **proyecto MARCELL** (Multilingual Resources for CEF.AT in the Legal Domain) es una iniciativa que busca mejorar las traducciones automáticas en el ámbito jurídico, también dentro del marco del programa Connecting Europe Facility (CEF). Su objetivo principal es recopilar y estructurar grandes cantidades de datos jurídicos multilingües, provenientes de legislaciones nacionales de varios países europeos, para optimizar la precisión de las traducciones automáticas en contextos legales, utilizando la herramienta eTranslation. MARCELL se enfoca en procesar documentos legislativos y normativos en varios idiomas, asegurando que los términos legales específicos sean traducidos correctamente y mantengan su contexto jurídico. Esto es especialmente útil para juicios transfronterizos, donde las barreras lingüísticas pueden complicar el proceso judicial. El proyecto cubre 7 idiomas: búlgaro, croata, húngaro, letón, rumano, eslovaco y esloveno, y trabaja en estrecha colaboración con instituciones públicas de estos países para crear recursos lingüísticos confiables. Por lo tanto, también apoya a eTranslation proporcionando datos lingüísticos en sectores legales específicos de varios países europeos, lo que mejora la capacidad de la IA para traducir términos legales complejos, lo que resulta en una mejor accesibilidad en procedimientos transfronterizos.



GESTIÓN DE CASOS JUDICIALES

Se sabe, y cada vez más, que la IA permite una mayor eficiencia en la tramitación de expedientes, facilita la detección de patrones y mejora la cooperación transfronteriza entre agencias judiciales. En Europa, varios Estados miembros han adoptado soluciones tecnológicas avanzadas para enfrentar estas problemáticas, entre los que destacan Alemania y otros países con iniciativas similares.

En Alemania, destaca **OLGA (Online-Strafverfahrensregister für Organisierte Kriminalität und Geldwäsche)**, en alemán, "Registro en línea de procesos penales para la lucha contra la delincuencia organizada y el blanqueo de capitales"). Se trata de un sistema digital desarrollado para centralizar y gestionar los casos de crimen organizado que permite a los fiscales y jueces acceder a una plataforma donde se integran todos los datos de investigaciones relacionadas, facilitando una visión global y unificada de las actividades ilícitas y optimizando así la tramitación de procedimientos largos y complejos, ayudando a reducir el tiempo de respuesta judicial.

También en Alemania encontramos **Frauke (Fraud Analysis Using Knowledge Extraction)**. Un proyecto de IA destinado a detectar patrones de fraude y lavado de dinero en grandes bases de datos financieras. Utilizando algoritmos de machine learning, Frauke analiza transacciones y comportamientos sospechosos, proporcionando a las autoridades informes detallados que ayudan a priorizar investigaciones. Este enfoque es especialmente útil en casos que involucran crimen organizado transnacional, donde los flujos financieros suelen ser opacos y dispersos.

En el Reino Unido, cuentan con un sistema Judicial Inteligente que utiliza herramientas de IA para gestionar el flujo de casos y mejorar la eficiencia. Esto incluye la implementación de tecnologías que facilitan la recopilación y análisis de datos, permitiendo a los abogados y jueces acceder rápidamente a la información relevante de los casos. Este enfoque también ha permitido identificar patrones de crimen organizado, facilitando la coordinación entre diferentes agencias judiciales

Francia ha introducido TAJ (*Traitement d'Antécédents Judiciaires*), una base de datos nacional que almacena datos de antecedentes judiciales y está conectada con sistemas de IA para analizar patrones de comportamiento delictivo. Este sistema ayuda a las autoridades judiciales a rastrear y gestionar múltiples investigaciones simultáneas.

Por otra parte, la IA específicamente en el ámbito del procesamiento del lenguaje natural (PLN), está emergiendo como una herramienta clave para la gestión de procedimientos judiciales complejos en Europa. Esta tecnología permite automatizar tareas que históricamente han requerido gran cantidad de recursos humanos y tiempo, como la revisión de documentos legales, la búsqueda de jurisprudencia relevante y la traducción de textos entre diferentes idiomas. La Unión Europea ha reconocido el potencial del PLN en la modernización de los sistemas judiciales y ha impulsado diversas iniciativas para fomentar su desarrollo y aplicación.

En relación con esto encontramos Horizonte 2020. Se trata de un programa marco de investigación e innovación de la Unión Europea (2014-2020) que ha financiado proyectos orientados a la creación de tecnologías de IA aplicables a sectores como la justicia y la gobernanza pública. En el ámbito del procesamiento del lenguaje natural, Horizonte 2020 ha impulsado investigaciones para desarrollar herramientas capaces de analizar grandes volúmenes de documentos legales y agilizar la toma de decisiones. Esto incluye proyectos que permiten a las autoridades judiciales y legales realizar búsquedas de precedentes y normativas en múltiples lenguas y en tiempo real, lo que mejora la eficiencia en procedimientos judiciales que involucran diferentes jurisdicciones o contextos internacionales.

En este mismo sentido, el Parlamento Europeo, a través de su Panel para el Futuro de la Ciencia y la Tecnología (STOA), ha realizado estudios sobre el impacto de las tecnologías de IA, incluido el procesamiento del lenguaje natural, en la justicia y otros sectores clave. Estos estudios se centran en cómo la IA puede asistir a jueces, abogados y personal administrativo al gestionar procedimientos judiciales largos y complejos. STOA ha analizado las ventajas del uso de IA para reducir los errores humanos en la interpretación de normativas, mejorar la coherencia en las decisiones judiciales, y ofrecer herramientas más accesibles para el manejo de grandes cantidades de información.

Un ejemplo de esto a nivel procesal lo encontramos en Estonia. Estonia ha sido pionera en la digitalización judicial, y actualmente están desarrollando un sistema que incorpora inteligencia artificial para la gestión de casos penales. Este sistema busca optimizar la asignación de penas, el seguimiento del cumplimiento de medidas alternativas, y el análisis de grandes volúmenes de datos procesales. Al usar IA, el sistema puede identificar patrones en las decisiones judiciales y sugerir acciones correctivas, lo que facilita el seguimiento de los condenados en todo el país.

4.2. LA IA EN LAS INSTITUCIONES DE SEGURIDAD

Al igual que en otras múltiples disciplinas y ámbitos, la IA y sus posibles aplicaciones es un asunto candente en el ámbito de la investigación criminal. En un reciente artículo, Olowe et al. (2023)[113] analiza la literatura existente sobre este asunto, poniendo de manifiesto que los estudios hasta la fecha están enfocados de manera predominante en el objetivo de la generación de medidas proactivas de disuasión y prevención del crimen gracias a estimaciones predictivas. La siguiente tabla recoge los parámetros de búsqueda bibliográfica empleados por el autor para su estudio.

Criteria	Metrics
Keywords	"crime" OR "criminal investigations" OR "police" OR "policing" AND "AI" OR "artificial intelligence" OR"facial recognition" OR "deep learning" OR "machine learning" OR "neural network" OR "natural language processing" OR "computer vision"
Year range	2012-2022
Subject area	Computer Science, Arts & Humanities, Psychology, Business, Management and Accounting, Decision Sciences, Social Sciences and Multidisciplinary
Document type Source Type	Journal Articles and IS Conference Papers Journal and Conference Proccedings
Publication stage Language	Final English

Table 1. Research protocol for bibliometric review

113 Olowe, O., Kawalek, P., & Odusanya, K. (2023). *Artificial Intelligence Adoption in Criminal Investigations: Challenges and Opportunities for Research*. https://aisel.aisnet.org/ukais2023

Como puede verse en la tabla a continuación, en este mismo estudio[114] se analiza las palabras clave de toda la bibliografía objeto de su estudio sobre la aplicación de la IA a la investigación criminal, concluyendo que los términos más recurrentemente empleados son Machine Learning, Detección de Fraude, Aprendizaje Profundo, Malware, Ciber-seguridad y redes sociales.

Rank	Keyword	Occurrences
1	Machine Learning	27
2	Fraud Detection	19
3	Deep Learning	12
4	Malware	6
5	Cyber Security	5
6	Social Media	5
7	Intrusion Detection System	4
8	Cyberbullying	4
9	Artificial Intelligence	4
10	Class Imbalance	4
11	Crime Prediction	3
12	Twitter	3
13	Malware Detection	3
14	Anomaly Detection	3
15	Text Mining	3

Table 3. Top 15 most occurring author keywords

Las máquinas dotadas de IA, al igual que los programas o softwares, pueden ayudar a los investigadores a acortar el tiempo empleado en diversas tareas a lo largo de diferentes estadios del proceso investigativo[115]. Desde la gestión y dirección de casos hasta la gestión de personal hasta la generación de informes, los sistemas con capacidades de IA mejoran la automatización de los flujos de trabajo, pudiendo ofrecer garantías de que los procedimientos de gestión son ejecutados de manera coherente y transparente. Esta eficiencia libera recursos de las operaciones y potencia la efectividad general de la organización[116].

114 Idem

115 Singh Sankhla, M., Kumar, R., & Jadhav, E. B. (2020). Artificial Intelligence: Advancing Automation in Forensic Science & Criminal Investigation. *Seybold Report*, *15*(8). https://www.researchgate.net/publication/343826071

116 Varma Microsoft, P. (s. f.). *Transforming Law Enforcement Policies and Governance Procedures: The Benefits of AI Integration*. https://doi.org/10.5678/ijai.2020.12345

A continuación, se describen algunas de las aplicaciones de la IA a la investigación criminal que están siendo tenidas en cuenta de manera relevante.

APLICACIONES DE LA IA A LA INVESTIGACIÓN CRIMINAL

ANÁLISIS DE RIESGO Y POLICÍA PREDICTIVA

El análisis de riesgos está presente en muchos estadios del sistema de justicia criminal en Estados Unidos, aportando valoraciones en relación con la práctica de detenciones, emisión de sentencias, medidas correctivas o reingresos en prisión [117]. Es razonable pensar que en el estricto ámbito de la actividad de investigación, valoraciones de naturaleza análoga pueden ser de interés para el investigador en relación con el perfilamiento de criminales o la priorización de líneas de investigación.

Efectivamente, los algoritmos pueden ser desplegados para contribuir en tareas de asesoramiento legal y de apoyo a la decisión en el ámbito de la investigación criminal. En Estonia ha sido desplegada y testeada una IA capaz de escuchar y decidir sobre pequeñas demandas. En esos entornos de justicia criminal, los algoritmos de IA son empleados principalmente para evaluar los perfiles de los investigados y la previsibilidad de reincidencia, afectando principalmente en la emisión de sentencias[118] [119] [120].

En Estados Unidos, donde mayor despliegue de IA existe en el sistema de justicia criminal, es de destacar la decisión adoptada por la Corte Suprema de Wisconsin sobre un algoritmo de análisis de riesgo empleado el enjuiciamiento del caso Loomis v. Wisconsin (2016). Dicho algoritmo es conocido como Correctional Offender Management Profiling for Alternative Sanctions (COMPAS), y en la sentencia en cuestión, el tribunal resolvió que el uso del algoritmo no supuso una violación del derecho del procesado a un juicio justo, pues se emplearon otros argumentos para fundamentar la decisión a parte del resultado del algoritmo[121].

Las fuerzas y cuerpos de seguridad en todo el mundo llevan mucho tiempo haciendo predicciones sobre la incidencia de la actividad delictiva, y despliegan agentes y recursos policiales en función de esas evaluaciones de riesgo[122], pero, como ya se ha mencionado, la IA también puede tener un papel muy relevante en la gestión de un caso de investigación mediante el análisis de datos y el apoyo a la decisión, por ejemplo, contribuyendo a la definición y priorización de líneas de investigación.

En el ámbito policial, el concepto de análisis de riesgo está directamente relacionado con el término *predictive policing, el cuál cubre diferentes métodos de predicción de la actividad criminal, basados en el cálculo de probabilidades. Estos métodos asumen que la actividad criminal está sujeta a las reglas de la probabilidad y que, por tanto, pueden hacerse predicciones sobre la base de los datos pasados. Pero, efectivamente, también puede aplicarse este método enfocando a la identificación de individuos potencialmente peligrosos. Esta predictive policing, a diferencia de otras disciplinas forenses, tiene vocación de tener efectos antes de que el crimen tenga lugar*[123] [124].

Efectivamente, el papel de la IA en la prevención criminal está resultando cada vez más relevante, ofre-

117 Garrett, B. L., Rudin, C., Williams, L. N., Beale, S. S., Benjamin, S., Boyle, J., Buccafusco, C., Dellinger, W., Fergusen, A., Griffin, L., Grunwald, B., Helfer, L., Kang, S., Lemos, M., Meltzer, A., Park, H., Purdy, J., Rai, A., Siegel, N., … Weiner, J. (2023). *The right to a glass box: Rethinking the use of artificial intelligence in criminal justice*.

118 Simmler, M., Brunner, S., Canova, G., & Schedler, K. (2023). Smart criminal justice: exploring the use of algorithms in the Swiss criminal justice system. *Artificial Intelligence and Law*, *31*(2), 213-237. https://doi.org/10.1007/s10506-022-09310-1

119 Faqir, R. S. A. (2023). Digital Criminal Investigations in the Era of Artificial Intelligence: A Comprehensive Overview. *International Journal of Cyber Criminology*, *17*(2), 77-94. https://doi.org/10.5281/zenodo.4766706

120 Kanwel, S., Imran Khan, M., & Usman, M. (2023). From Bytes to Bars: The Transformative Influence of Artificial Intelligence on Criminal Justice. *Qlantic Journal of Social Sciences*, *4*(4), 84-89. https://doi.org/10.55737/qjss.059046443

121 Idem.

122 Garrett, B. L., Rudin, C., Williams, L. N., Beale, S. S., Benjamin, S., Boyle, J., Buccafusco, C., Dellinger, W., Fergusen, A., Griffin, L., Grunwald, B., Helfer, L., Kang, S., Lemos, M., Meltzer, A., Park, H., Purdy, J., Rai, A., Siegel, N., … Weiner, J. (2023). *The right to a glass box: Rethinking the use of artificial intelligence in criminal justice*.

123 Simmler, M., Brunner, S., Canova, G., & Schedler, K. (2023). Smart criminal justice: exploring the use of algorithms in the Swiss criminal justice system. *Artificial Intelligence and Law*, *31*(2), 213-237. https://doi.org/10.1007/s10506-022-09310-1

124 Singh Sankhla, M., Kumar, R., & Jadhav, E. B. (2020). Artificial Intelligence: Advancing Automation in Forensic Science & Criminal Investigation. *Seybold Report*, *15*(8). https://www.researchgate.net/publication/343826071

ciendo soluciones innovadoras para mejorar las medidas de seguridad y reducir la actividad criminal[125]. El Big Data en el ámbito criminal permite el desarrollo del *data-driven policing, es decir, el desarrollo de la actividad policial mediante decisiones basadas en el análisis de datos y no únicamente en la intuición o experiencias pasadas*[126]. En una escala mayor de sofisticación, el *AI-driven data análisis, es decir, el análisis de datos basado en el uso de IA, refuerza a las agencias de seguridad en el desarrollo de estrategias basadas en la evidencia para enfrentarse a las necesidades sociales y a las amenazas emergentes. Mediante el análisis de grandes volúmenes de datos, incluyendo estadísticas criminales, el feedback de las comunidades sociales, y las tendencias demográficas, los algoritmos de IA ofrecen información sobre problemas subyacentes que puedan pasar desapercibidos, haciendo posible a los responsables de definir las estrategias de seguridad arbitrar intervenciones específicas y distribuir de una manera eficiente los recursos*[127]. Un ejemplo de esta aplicación podría ser un algoritmo basado en IA que evaluara el riesgo de que una agresión de violencia de género tenga lugar en un caso dado de alta en sistema VioGen.

DETECCIÓN DE CRÍMENES EN SERIE

Por detección de crímenes en serie, nos referimos a la actividad encaminada a determinar qué crímenes fueron cometidos por la misma persona o por el mismo grupo de personas. Ante este desafío, la IA puede analizar los datos existentes para determinar conjuntos de crímenes similares en modus operandi [128].

En el caso de Italia, el primer software de análisis predictivo implementado en el país, KeyCrime, fue desarrollado en 2007 en el cuartel general de la Policía de Milán por el entonces asistente de la Policía estatal, Mario Venturi. Este programa integra una actividad de cálculo de big data que es capaz de detectar crímenes en serie y predecir dónde, cuándo y cómo podría tener lugar el próximo hecho delictivo. A su vez, elaboraría sus premisas basándose en cuatro elementos fundamentales de cada crimen: su tipología, el objetivo, el modus operandi (incluidos los objetos, las armas y los medios de transporte empleados) y las características psicofísicas del autor (incluyendo gestos, ropa, tatuajes, piercings, cicatrices o cualquier objeto visible que pudiera identificarlo). Así, el proceso de análisis cuenta con dos fases diferentes: una primera inductiva, en la cual se analiza con detalle un crimen concreto para identificar los elementos comunes con otros de similares características y así relacionarlos con un único autor; y una segunda deductiva, la cual, tras observar los elementos clave identificados en la serie criminal, se podría predecir el cuándo, dónde y cómo se cometerá el delito futuro, esto es, las principales "w" (where, when y how). KeyCrime, por lo tanto, permitiría una protección policial más eficaz. Según una auditoría, la aplicación del software en el entorno de Milán favoreció un aumento de ocho puntos porcentuales la probabilidad de resolver un crimen en serie, disminuyendo el número de robos que los grupos criminales son capaces de perpetrar antes de su detención.[129]

En España se ha desarrollado el programa Eurocop PredCrime[130], un proyecto del lejano año 2011 en colaboración con la Universidad Jaume I de Castellón. Se trata de un sistema para la predicción y prevención del delito cuya finalidad es elaborar un mapa de previsión de riesgo, en lugares concretos de una ciudad y en determinados horarios. Vendría a ser la "versión española de PredPol". En este caso el sistema no marca cuadrículas, sino mapas de calor con zonas proclives a la realización de un delito. El programa ha sido testado por la Policía Local de Castellón y de Rivas-Vaciamadrid.

Dentro de este tipo de aplicaciones, podría destacar el uso de esta técnica para la identificación de asesinos en serie mediante el análisis de los datos de homicidios resueltos y sin resolver en todo un país o territorio, la atribución de distintas actividades ilegales de tráfico de drogas a una misma organización criminal o la detección y atribución de campañas masivas de fraudes por internet con un mismo origen.

125 Kanwel, S., Imran Khan, M., & Usman, M. (2023). From Bytes to Bars: The Transformative Influence of Artificial Intelligence on Criminal Justice. *Qlantic Journal of Social Sciences*, *4*(4), 84-89. https://doi.org/10.55737/qjss.059046443

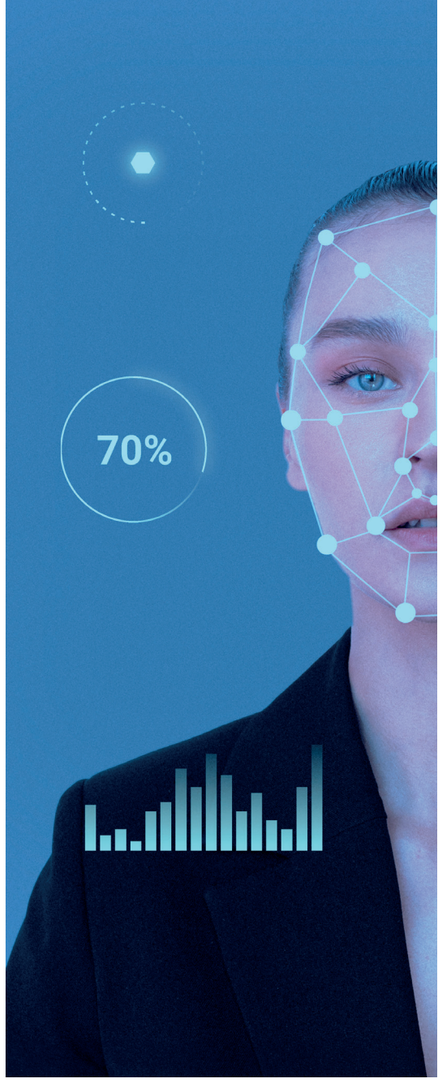
126 Abusamadov, K. (2024). Revolutionizing Crime Prevention: The Role of AI and Big Data in Modern Law Enforcement. *Journal of law, market & innovation*, *I*(2), 21-25. https://doi.org/10.5281/zenodo.11928015

127 Varma Microsoft, P. (s. f.). *Transforming Law Enforcement Policies and Governance Procedures: The Benefits of AI Integration*. https://doi.org/10.5678/ijai.2020.12345

128 Garrett, B. L., Rudin, C., Williams, L. N., Beale, S. S., Benjamin, S., Boyle, J., Buccafusco, C., Dellinger, W., Fergusen, A., Griffin, L., Grunwald, B., Helfer, L., Kang, S., Lemos, M., Meltzer, A., Park, H., Purdy, J., Rai, A., Siegel, N., … Weiner, J. (2023). *The right to a glass box: Rethinking the use of artificial intelligence in criminal justice*.

129 VV.AA., HERRERA TRIGUERO, Francisco; PERALTA GUTIÉRREZ, Alfonso; TORRES LÓPEZ, Leopoldo Salvador. Capítulo "Uso policial de sistemas de inteligencia artificial en el ámbito comparado" pág. 453 y ss. El derecho y la inteligencia artificial. 2022. 24/10/2022. Editorial Universidad de Granada. 978-84-338-7049-0

130 https://www.eurocop.com/catedra-eurocop/proyectos-en-marcha/eurocop-pred-crime-sistemas-para-la-prediccion-y-prevencion-del-delito/#:~:text=El%20Proyecto%20Eurocop%20Pred%2DCrime,un%20crimen%20a%C3%BAn%20no%20producido.



RECONOCIMIENTO FACIAL

El uso de la IA aplicada a las técnicas de reconocimiento facial en el marco de la investigación criminal es uno de los más recurrentemente referidos en la literatura científica [131][132]. El reconocimiento facial automático es una técnica empleada desde hace muchos años por los investigadores para, por ejemplo, identificar personas o sospechosos anónimos de un crimen a partir de una imagen de su rostro. Muy potentes soluciones informáticas son capaces de detectar coincidencias cotejando una imagen anónima con miles de otras imágenes de personas identificadas, todo ello en cuestión de segundos. Generalmente estas capacidades han estado limitadas por la necesidad disponer de una imagen anónima de cierta calidad, que no siempre es algo posible. También ha sido una limitación para estas soluciones tradicionales contar, por un lado, con una base de datos de imágenes identificadas de calidad y, por otro lado, con alta capacidad de cómputo. Sin duda, la IA puede aportar a estas soluciones una mayor eficacia, permitiendo cotejar imágenes de menos calidad, mejorando la precisión y reduciendo a la vez el número de falsos positivos. La IA no presenta fatiga como el ser humano y ciertos desarrollos en curso tratan de hacer aprender a al IA a identificar los rostros tal y como lo haría el ser humano[133].

El Ministerio de Interior está entrenando un algoritmo para su nuevo sistema ABIS[134], siglas en inglés que responden a "sistema automático de identificación biométrica". La IA ayudará a la policía a identificar a los presuntos culpables de un delito. Este sistema solo se activará ante ilícitos graves y lo está desarrollando la firma francesa Thales.

El algoritmo en sí de este sistema ABIS recibe el nombre de Cogent y el funcionamiento no se trata de un modelo de IA que emplee una IA de reconocimiento facial en tiempo real y en remoto para identificar en directo a todos los ciudadanos que paseen ante unas cámaras. En su lugar es "un sistema científico criminalístico que permite la identificación de personas detenidas o a las que se atribuya la comisión de ilícitos penales en base a información previa recogida por la obligación legal y el desarrollo de las misiones asignadas a las Fuerzas y Cuerpos de Seguridad del Estado". Es un programa informático que la policía española podrá aplicar sobre grabaciones y fotografías de la escena de un posible crimen.

Según pregunta escrita en el Congreso[135] la herramienta tiene por finalidad identificar a personas implicadas en las actuaciones llevadas a cabo en ilícitos penales que hagan necesaria la aplicación de la medida cautelar de detención o identificación para trasladarla a la autoridad judicial. Para ello, se utilizarán como marco de consulta las diferentes grabaciones obtenidas y tratadas, por cámaras tanto públicas como privadas, que hayan podido captar la imagen de la

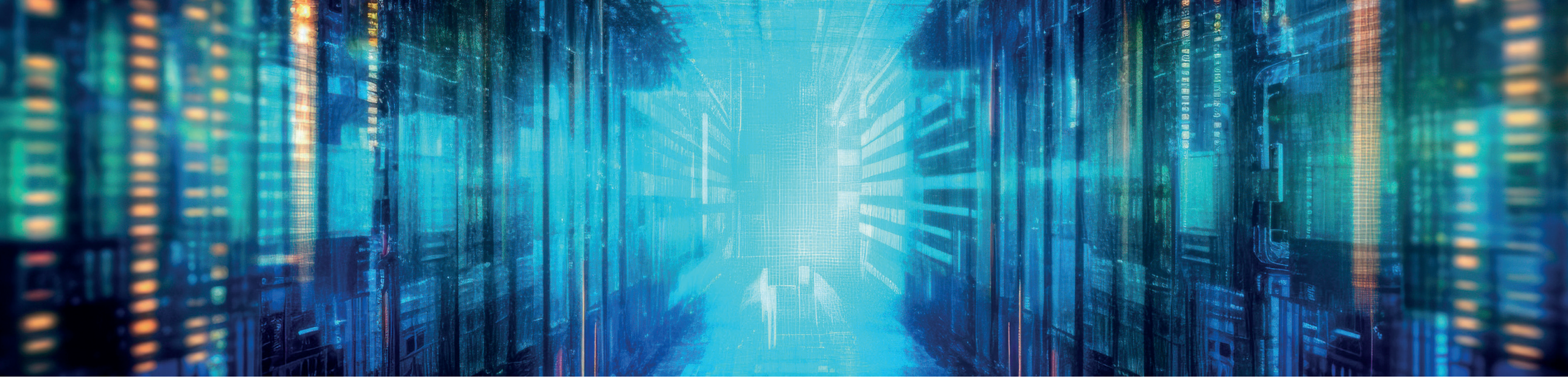
131 Idem.

132 Kanwel, S., Imran Khan, M., & Usman, M. (2023). From Bytes to Bars: The Transformative Influence of Artificial Intelligence on Criminal Justice. *Qlantic Journal of Social Sciences*, *4*(4), 84-89. https://doi.org/10.55737/qjss.059046443

133 idem

134 AGUILAR, Alberto R. Interior reconoce que no ha consultado a la AEPD sobre el algoritmo de reconocimiento facial que está entrenando para la policía. Business Insider. 26 de diciembre de 2022. Consultado el 1 de julio de 2023. Disponible en: https://www.businessinsider.es/interior-no-ha-consultado-aepd-antes-construir-ia-policial-1173474

135 Pregunta escrita Congreso 184/99831 10/01/2023 251055 AUTOR/A: CORTÉS GÓMEZ, Ismael (GCUP-ECP-GC); SANTIAGO ROMERO, Enrique Fernando (GCUP-ECP-GC). Consultado 28 de junio de 2023 y disponible en: https://www.congreso.es/entradap/l14p/e25/e_0255721_n_000.pdf



persona responsable del ilícito penal. El nuevo sistema FRS de ABIS, realiza una evaluación, que consiste en buscar a partir de una imagen (fotografía o fotograma) captada en la comisión de un ilícito penal y comparar la misma con las imágenes indubitadas correspondientes a la reseña policial de detenidos para determinar la posible existencia de uno o varios candidatos. Una vez se cuente con este resultado, los especialistas en reconocimiento facial deberán realizar sobre cada imagen ofertada, el mismo proceso de comparación manual realizado hasta la fecha para confirmar o descartar la correspondiente identificación.

Asimismo, Interior sostiene que el sistema permite el ejercicio de todos los derechos de protección de datos, que cumple con la normativa de privacidad y que ha superado el juicio de proporcionalidad en relación con otros medios que se puedan implementar o que se estén implementando (ADN, huellas dactilares, etc.).

De hecho, los beneficios de la persona interesada quedan acreditados durante todo el ciclo de vida de los datos, incidiéndose en que no se aplican decisiones automatizadas por la herramienta sin intervención y supervisión humana.

Al igual que del resto de derechos que tiene una persona detenida por la comisión de un ilícito tipificado en el código penal, cuando se utilice el sistema, se dejará constancia en las correspondientes diligencias y en la información que se les facilita a las personas implicadas en actuaciones procesales penales.

Se explican unos plazos de conservación y de revisión que puede llegar hasta 20 años y que la tasa de error o falsos negativos es de un 3%

Es decir, se trata de un sistema de alto riesgo, cuyo uso se permite a posteriori mediante aprobación por autorización judicial.

DETECCIÓN E IDENTIFICACIÓN DE DISPAROS

Esta aplicación de la IA se basaría en la identificación de patrones, por ejemplo, en una grabación de audio, a fin de reconocer un disparo de arma de fuego, e ir más allá, tratando de aportar información sobre el tipo de arma o munición empleada e, incluso, sobre el posible punto de origen del disparo y su trayectoria. Estas aplicaciones podrían incluso embarcase en teléfonos inteligentes u otros dispositivos portátiles que puedan asistir a los agentes sobre el terreno en una situación de uso de armas de fuego[136]. Por supuesto, otra aplicación de esta técnica es la forense, que resultaría del postprocesado de grabaciones de audio o vídeo de eventos violentos en los que hayan tenido lugar disparos de armas de fuego, como en casos de homicidio o de un atentado terrorista.

También, en relación con la balística forense, la IA puede ser de ayuda para la detección de patrones en las marcas dejadas por las armas en los cartuchos o proyectiles, permitiendo automatizar y acelerar los procesos de cotejo forense[137].

ANÁLISIS MASIVO DE DATOS

La investigación criminal, con toda lógica, siempre se ha apoyado de manera significativa en la recolección de datos y en su análisis. En el caso de las ciencias forenses, también se descansa en el filtrado y evaluación de los elementos de interés recolectados para un caso y su cotejo con grandes cantidades de datos. En los últimos tiempos, se han producido avances disruptivos para el desarrollo de etas áreas, todo ello gracias a nuevas técnicas, propias de la era digital en la que vivimos. Efectivamente, los datos recolectados precisan ser analizados con el mínimo esfuerzo posible, pero con la máxima precisión alcanzable. Las grandes cantidades de datos a los que se tiene acceso en el marco de las investigaciones son descomunales y cada vez es mayor[138], esto hace acuciante contar con capacidades de *minería de datos hasta ahora no empleadas en este dominio. A todo lo anterior hay que añadir, en muchas ocasiones, que la heterogeneidad de los datos con los que es necesario trabajar es mayúscula. Un ejemplo de ello es el caso de evidencias electrónicas obtenidas de la interceptación legal de las comunicaciones electrónicas de un investigado, las cuales pueden contener archivos digitales de vídeo o audio en conjunto con textos y otros muchos formatos. En estos casos, contar con soluciones de minería de datos basadas en IA que cuenten con capacidades de visión computacional o de transcripción de voz a texto, traducción en distintos idiomas incluida, es un avance de un impacto descomunal en la labor de los investigadores*[139].

En concreto, las agencias policiales están empleando soluciones IA de esta naturaleza, tanto para la detección de las fases preparatorias de la comisión de hechos criminales, como para el escrutinio y análisis de crímenes ya cometidos, incluyendo la identificación de los responsables en ambos casos[140]. Un ejemplo de esta detección de fases preparatorias de un crimen podría consistir en el análisis de grandes cantidades de datos bancarios y tributarios para la detección de complejas operativas fiscales, orientadas a la comisión

136 Singh Sankhla, M., Kumar, R., & Jadhav, E. B. (2020). Artificial Intelligence: Advancing Automation in Forensic Science & Criminal Investigation. *Seybold Report*, *15*(8). https://www.researchgate.net/publication/343826071

137 Idem

138 Un ejemplo de ello es la gran cantidad de información digital que puede albergar un solo teléfono inteligente o la resultante de una interceptación de comunicaciones electrónicas correspondiente a una línea de fibra óptica.

139 Simmler, M., Brunner, S., Canova, G., & Schedler, K. (2023). Smart criminal justice: exploring the use of algorithms in the Swiss criminal justice system. *Artificial Intelligence and Law*, *31*(2), 213-237. https://doi.org/10.1007/s10506-022-09310-1

140 Kanwel, S., Imran Khan, M., & Usman, M. (2023). From Bytes to Bars: The Transformative Influence of Artificial Intelligence on Criminal Justice. *Qlantic Journal of Social Sciences*, *4*(4), 84-89. https://doi.org/10.55737/qjss.059046443

de fraudes a gran escala. Otro ejemplo es su posible empleo para el análisis en tiempo de real de grandes cantidades de datos de la red de comunicaciones electrónicas de una organización, para la detección preventiva de posibles TTP[141] empleadas por ciberdelincuentes para perpetrar un ataque o intrusión a la red.

Por ejemplo, IBM Security i2 Analyst's Notebook es un producto de software de IBM utilizado para el análisis de datos e investigación. Es una herramienta de software basada en la metodología ELP (entidad-relación-propiedad), que ofrece al usuario la posibilidad de conocer las relaciones entre las entidades de datos para descubrir patrones e información de los datos. Analyst's Notebook es una herramienta de uso común para los analistas digitales de las fuerzas del orden, el ejército y otros organismos gubernamentales de inteligencia, o en los departamentos de fraude de las entidades financieras, los organismos reguladores, etc. Forma parte del Human Terrain System (HTS), un programa del Ejército de los Estados Unidos que integra a científicos sociales con brigadas de combate.

De igual manera, dentro del proyecto COPKIT en su fase de extracción de datos nos encontramos con el desarrollo Reconocimiento de entidades nombradas – CKNER, herramienta desarrollada por el Instituto Austriaco de Tecnología (AIT). Un servicio que integra varios reconocedores de entidades con nombre de última generación, cada uno de ellos con modelos estándar y específicos de dominio entrenados en corpus genéricos, que se centran en datos de texto adquiridos a través del rastreo de mercados de la darknet que ofrecen armas y drogas, centrándose en textos cortos y mal escritos.

Y por ejemplo en organizaciones criminales familiares puede ser muy útil dentro del mismo proyecto la Extracción de relaciones – CKREL-EXT, una herramienta desarrollada por el Instituto Austriaco de Tecnología (AIT). El Servicio REST para reconocer relaciones entre entidades (drogas, armas, nombres de usuario, ubicaciones, etc.). Toma un texto (por ejemplo, uno o varios párrafos de texto tomados de un anuncio del mercado de la red oscura) como entrada y produce un gráfico de entidades con nombre como resultado. El componente depende de las entidades reconocidas por el CKNER.

Del mismo proyecto en igual sentido, el Buscador de conexiones – CF, una herramienta desarrollada por Legind Technologies (LTA). Busca conexiones en grafos de entidades (por ejemplo, no limitadas a "personas" o identidades, sino también heterogéneas) con vínculos inciertos. Puede utilizarse en investigaciones en línea (incluida la red oscura) para encontrar relaciones entre varios elementos de identidades digitales (por ejemplo, nombres de usuario, billeteras de moneda digital, etc.), en relación con bases de inteligencia internas si están disponibles (y si es posible la fusión de grafos).

En cuanto a redes sociales, Haternet[142] identifica y monitoriza la evolución del discurso del odio en Twitter (actualmente X), visualizándolo utilizando técnicas de análisis de redes, introduce una base de datos sobre el discurso del odio en español, consistente en 6000 tuits etiquetados, y compara la clasificación visualizando su estado y evolución.

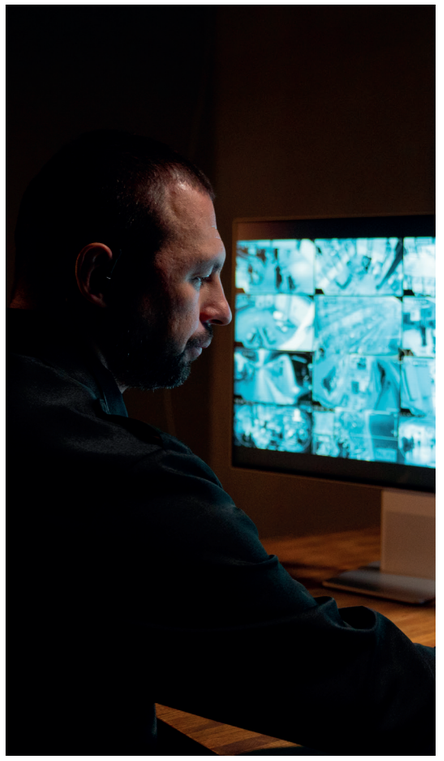
141 Técnicas, Tácticas y Procedmientos.

142 Pereira-Kohatsu JC, Quijano-Sánchez L, Liberatore F, Camacho-Collados M. Detecting and Monitoring Hate Speech in Twitter. Sensors (Basel, Switzerland). 2019 Oct;19(21):E4654. DOI: 10.3390/s19214654. PMID: 31717760; PMCID: PMC6864473.

Efectivamente, el análisis de *big data es esencial para detectar patrones y anomalías que podrían indicar posibles brechas de ciberseguridad o actividades fraudulentas. Al supervisar el tráfico de datos en tiempo real, los equipos de ciberseguridad, equipados con estas capacidades de IA, pueden identificar patrones inusuales que caracterizan un ciberataque en sus fases preparatorias, como intentos inusuales de inicio de sesión, picos en las solicitudes de acceso a datos o anomalías en las transacciones financieras*[143].

DIRECCIÓN DE INVESTIGACIONES

En la actualidad, la ejecución de la mayoría de investigaciones criminales exige a los investigadores el manejo masivo de datos que, con frecuencia, pueden ser heterogéneos y desestructurados. Hemos comprobado como el DL es especialmente efectivo cuando se trata de aprender, tratar y analizar datos no estructurados, por lo que esta tecnología y las aplicaciones de IA basadas en ella, pueden resultar especialmente interesantes para asistir en la toma de decisiones a los investigadores. Soluciones dotadas de este tipo de aprendizajes automáticos, basadas en redes neuronales profundas, pueden mejorar el análisis, organización y gestión de un caso, dando lugar a decisiones y recomendaciones sobre el mismo y sobre asuntos legales que tengan relación con él[144].



VIGILANCIA ENCUBIERTA AVANZADA

La posibilidad de explotar en tiempo real, y prácticamente sin latencias perceptibles al ser humano, modelos y algoritmos de IA ya entrenados para multiplicidad de tareas, algunas de ellas ya mencionadas, es una realidad indiscutible. Esta inmediatez expande un universo de nuevas capacidades de IA a las actividades investigativas de obtención de inteligencia basadas en la vigilancia encubierta de los investigados y sus actividades ilícitas. Los operativos de vigilancia física y electrónica llevados a cabo durante las investigaciones por los funcionarios de las agencias de IC, pueden beneficiarse ahora de capacidades insospechadas hasta hace muy poco tiempo. Por ejemplo, las capacidades de visión artificial hacen posible el análisis automático en tiempo real de las imágenes obtenidas por medio de medios de captación de imagen encubiertos, permitiendo detectar patrones de forma inmediata, predecir actividades potencialmente criminales antes de que se materialicen y facilitar, de esta forma, una intervención a tiempo o una toma de decisiones acertada y oportuna en el curso de una vigilancia. Efectivamente, la visión artificial puede detectar sutiles indicadores en el comportamiento de los investigados que sean relevantes para la toma de cualquier decisión[145].

COTEJO FORENSE DE EVIDENCIAS

La IA aplicada a la ciencia forense ha sido empleada en procesos judiciales en el contexto de las mezclas complejas de ADN. En los casos de mezclas de ADN de múltiples y, en ocasiones anónimos, contribuyentes, los algoritmos de IA han sido diseñados para determinar si un sujeto puede, o no, haber contribuido a una muestra tomada en la escena del crimen[146] [147] [148].

143 Abusamadov, K. (2024). Revolutionizing Crime Prevention: The Role of AI and Big Data in Modern Law Enforcement. *Journal of law, market & innovation*, *I*(2), 21-25. https://doi.org/10.5281/zenodo.11928015

144 Faqir, R. S. A. (2023). Digital Criminal Investigations in the Era of Artificial Intelligence: A Comprehensive Overview. *International Journal of Cyber Criminology*, *17*(2), 77-94. https://doi.org/10.5281/zenodo.4766706

145 idem.

146 Singh Sankhla, M., Kumar, R., & Jadhav, E. B. (2020). Artificial Intelligence: Advancing Automation in Forensic Science & Criminal Investigation. *Seybold Report*, *15*(8). https://www.researchgate.net/publication/343826071

147 Garrett, B. L., Rudin, C., Williams, L. N., Beale, S. S., Benjamin, S., Boyle, J., Buccafusco, C., Dellinger, W., Fergusen, A., Griffin, L., Grunwald, B., Helfer, L., Kang, S., Lemos, M., Meltzer, A., Park, H., Purdy, J., Rai, A., Siegel, N., … Weiner, J. (2023). *The right to a glass box: Rethinking the use of artificial intelligence in criminal justice*.

148 Kanwel, S., Imran Khan, M., & Usman, M. (2023). From Bytes to Bars: The Transformative Influence of Artificial Intelligence on Criminal Justice. *Qlantic Journal of Social Sciences*, *4*(4), 84-89. https://doi.org/10.55737/qjss.059046443

En el ámbito de la lofoscopia, la IA también tiene algo que aportar. El DL ha tenido un éxito considerable en el campo de la visión artificial y en el reconocimiento de patrones, dado que hace innecesaria la intervención humana para la identificación de características de cualquier muestra. El DL aprende automáticamente a hacer estas tareas mediante el entrenamiento con suficientes datos con los que se le provee. Estas ventajas de la IA la hacen especialmente interesante para la mejora en la eficacia de la ejecución de diversas tareas relacionadas con la identificación y clasificación automática de huellas dactilares. Esta tecnología puede reducir sustancialmente el número de comparaciones necesarias hasta conseguir un cotejo positivo, sumando precisión a la vez[149].

En el ámbito de la grafística también pueden aplicarse estas técnicas, que en algunos casos están ayudando a determinar el sexo del autor de un texto manuscrito. Otros ámbitos forenses de interés para la aplicación de la IA son el de la balística, la datación de la muerte por análisis de sangre o la identificación odontológica[150]

Investigadores del Instituto DaSCI (Universidad de Granada) y el centro CITIC (Universidad de A Coruña), en colaboración con la empresa Panacea Coop, han publicado un estudio que demuestra una mejora en el reconocimiento de restos humanos por superposición craneofacial, haciendo mucho más objetiva la toma de decisiones por parte de los expertos forenses.

La superposición craneofacial es una técnica forense que apoya la toma de decisiones cuando se trata de identificar restos óseos. Concretamente, se basa en el análisis de la superposición de un cráneo encontrado y sin identificar (post mortem) con fotografías faciales (ante mortem) de personas desaparecidas[151].

ANÁLISIS FORENSE DIGITAL

En la actualidad, el análisis forense de evidencias digitales representa un desafío descomunal para los investigadores. Como ya se ha mencionado en epígrafes anteriores, las cantidades de datos digitales obtenidos de los investigados, procedentes de registros en sus dispositivos electrónicos, pueden ser descomunales. Al problema de la cantidad, se suma el de la complejidad, pues la heterogeneidad de los datos y las soluciones de cifrado que son comúnmente empleadas por los dispositivos electrónicos, muchas veces de manera trasparente para los propios investigados, dificultan enormemente la interpretación de la información. Las soluciones de IA en este campo se enfocan en automatizar el análisis y la correlación de los datos adquiridos durante la investigación y, en base a su aprendizaje, presentar al investigador la información que es de su interés. Así, estos sistemas acaban reduciendo la cantidad de datos que hay que analizar personalmente por los investigadores, y ofrecen el subconjunto del total de la evidencia que es más probable que constituya la información de interés para la investigación.

Otra rama del análisis forense digital es el análisis forense de redes informáticas. Generalmente, el objeto de estos análisis es el estudio de la actividad de red para detectar el origen de infracciones de la política de seguridad de la red o brechas de seguridad de la información[152].

Asimismo, deben citarse otras aplicaciones de inteligencia artificial españolas, tales como, en primer lugar, 4nseek.es[153], que es una herramienta destinada a las fuerzas y cuerpos de seguridad. Su propósito es ayudar a los agentes a luchar contra el abuso sexual a menores. Analiza el contenido de discos o particiones y, a través de distintas técnicas de inteligencia artificial, es capaz de detectar la aparición de indicios de abuso sexual a menores en imágenes o vídeos, ofreciendo al usuario un listado de resultados priorizado

A todo lo anterior se suma el desafío que suponen los usos de la IA para propósitos criminales disruptivos por parte de los criminales, insospechados hasta la aparición de esta nueva tecnología. Si se añade a ello la complejidad de reunir pruebas transfronterizas para llevar a cabo investigaciones nacionales cuando un sistema de IA ha participado en la comisión o perpetración de una conducta ilícita, la dificultad para los investigadores se incrementa exponencialmente[154].

RECONSTRUCCIÓN DE ESCENA DEL CRIMEN Y REALIDAD VIRTUAL

La aplicación combinada de tecnologías digitales 3D e IA puede emplearse para mejorar la ejecución de ciertas fases de las técnicas de visualización forense. Esta combinación de tecnologías puede crear modelos gráficos 3D de objetos y personas, sobre la base de mediciones e imágenes, y animar dichos modelos para recrear la escena del crimen y los acontecimientos relacionados. Un ejemplo de esta técnica es el empleo del reconocimiento visual de patrones para el análisis del tamaño, forma y distribución de manchas de sangre, todo ello con el mencionado objeto de reconstruir los hechos acaecidos[155].

El proyecto VALCRI (Visual Analytics for Sense-Making in Criminal Intelligence Analysis) permite analizar la escena del crimen escaneando en segundos millones de fuentes de información en distintos formatos (Comisión Europea, 2018; Gallego, 2018): registros, víctimas, imágenes, lugar del delito, interrogatorios, etc. Detecta todos los patrones sospechosos y es capaz de reconstruir escenas y se presenta en pantallas táctiles interactivas, con la finalidad de compararlo con datos anteriores. De igual manera, combina la inteligencia artificial y el análisis visual, y se sirve de software de reconocimiento facial para detectar e identificar a personas concretas a partir de fuentes. Este sistema busca reconocer aquellos detalles que los humanos podemos pasar por alto. El proyecto está financiado por la Comisión Europea y en él han participado la Policía de West Midlands (Reino Unido) y la de Amberes (Bélgica) y está coordinado por la Universidad de Middlesex[156] [157].

149 Singh Sankhla, M., Kumar, R., & Jadhav, E. B. (2020). Artificial Intelligence: Advancing Automation in Forensic Science & Criminal Investigation. *Seybold Report*, *15*(8). https://www.researchgate.net/publication/343826071

150 idem.

151 Práxedes Martínez-Moreno, Andrea Valsecchi, Pablo Mesejo, Óscar Ibañez, Sergio Damas. Evidence evaluation in craniofacial superimposition using likelihood ratios. Information Fusion (2024). https://doi.org/10.1016/j.inffus.2024.102489

152 idem.

153 https://www.incibe.es/incibe/informacion-corporativa/con-quien-trabajamos/proyectos europeos/4nseek/herramienta

154 Faqir, R. S. A. (2023). Digital Criminal Investigations in the Era of Artificial Intelligence: A Comprehensive Overview. *International Journal of Cyber Criminology*, *17*(2), 77-94. https://doi.org/10.5281/zenodo.4766706

155 Singh Sankhla, M., Kumar, R., & Jadhav, E. B. (2020). Artificial Intelligence: Advancing Automation in Forensic Science & Criminal Investigation. *Seybold Report*, *15*(8). https://www.researchgate.net/publication/343826071

156 GALLEGO, Paloma. ¿Inteligencia artificial para casos de crímenes sin resolver?. En: Big Data Magazine. 30 mayo 2018. Disponible en: https://bigdatamagazine.es/inteligencia-artificial-para-casosde-crimenes-sin-resolver.

157 Comisión Europea. Visual analytics for brighter criminal intelligence analysis. En: Cordis. 7 febrero 2018.

PROYECTOS E INICIATIVAS PARA FORTALECER LA COOPERACIÓN EN MATERIA DE SEGURIDAD Y LAS INVESTIGACIONES DE DELITOS

El **Proyecto TRACE** financiado con fondos del programa de investigación e innovación Horizon[158] de la Unión Europea está conformado por un consorcio amplio de organizaciones cuyo propósito final es dotar a las autoridades policiales europeas de las herramientas y los recursos necesarios para identificar, rastrear, documentar y desmantelar flujos de dinero ilícitos de manera oportuna y eficaz en seis áreas: (i) financiación del terrorismo, (ii) análisis forense web, (iii) extorsión cibernética, (iv) uso de criptomonedas en transacciones del mercado inmobiliario, (v) lavado de dinero en arte y antigüedades, y (vi) juegos de azar en línea. El proyecto incluye el desarrollo de herramientas de IA para el análisis y visualización de datos financieros, la identificación de patrones de actividad financiera sospechosa y la colaboración con otras agencias para compartir información. El proyecto incluye once componentes entre los cuales se incluye uno sobre aspectos éticos, legales e impacto social[159].

El **proyecto *iBorderCtrl*** es una iniciativa financiada por la Unión Europea para mejorar el control fronterizo mediante el uso de tecnologías avanzadas, incluyendo el reconocimiento facial y un sistema de evaluación de riesgos. Comenzó en 2016 y finalizó en 2019, con pruebas en Grecia, Hungría y Letonia. Su objetivo principal es acelerar el cruce de fronteras para los viajeros de terceros países, utilizando un sistema de pre-registro y tecnologías biométricas, como el escaneo facial y de las venas de la palma, para verificar la identidad de los viajeros antes de que lleguen a la frontera.

A pesar del interés que generó, uno de los elementos más controvertidos del proyecto fue un detector de mentiras impulsado por IA, que evalúa las expresiones faciales de los viajeros durante una entrevista virtual para determinar si están mintiendo sobre su propósito de viaje o duración de la estancia. Si el sistema detecta irregularidades, se realizan verificaciones adicionales por agentes humanos. A pesar de su innovación, el sistema ha sido criticado por problemas éticos y de fiabilidad, como su posible discriminación contra personas con discapacidades o ansiedad, y por preocupaciones sobre la invasión de los derechos fundamentales. Y recientemente, en 2023, el Tribunal de Justicia de la Unión Europea falló a favor de la transparencia, permitiendo que algunos documentos relacionados con el proyecto sean accesibles al público, aunque se mantiene la protección de ciertos intereses comerciales del consorcio que lo desarrolló

La Guardia Civil de España ha adoptado un sistema llamado **ATLAS**, una herramienta basada en IA que permite analizar grandes volúmenes de información obtenidos de dispositivos electrónicos incautados, como teléfonos móviles y computadoras. ATLAS utiliza técnicas de minería de datos y aprendizaje automático para identificar pruebas ocultas y relacionar comunicaciones entre sospechosos, lo que es particularmente útil en casos de crimen organizado y narcotráfico. Además, el Cuerpo Nacional de Policía de España ha desarrollado **Centinela**, que aplica algoritmos de inteligencia artificial para detectar vínculos entre distintas organizaciones criminales, tanto a nivel nacional como internacional. Centinela es especialmente eficaz en la lucha contra el narcotráfico y la trata de personas.

158 El Proyecto TRACE en: https://trace-illicit-money-flows.eu/

159 Los 11 componentes del proyecto se describen en: https://trace-illicit-money-flows.eu/project-outcomes/

4.3. HERRAMIENTAS DE IA PARA TEMÁTICAS U ÁREAS ESPECÍFICAS

ANÁLISIS Y EVALUACIÓN DE PRUEBAS

En el ámbito del análisis y evaluación de pruebas se pueden destacar los siguientes proyectos e iniciativas que están enmarcados tanto en el ámbito judicial como policial.

- **AVENUE (Analysis of Video Evidence with Novel Enhanced Understanding Engine)**. Este proyecto, financiado por la Comisión Europea dentro del marco del programa Horizonte 2020, utiliza IA para el análisis avanzado de pruebas en formato de video, que son cada vez más comunes en investigaciones penales. AVENUE aplica algoritmos de visión por computadora y reconocimiento de objetos para automatizar la revisión de grandes cantidades de grabaciones de cámaras de seguridad o teléfonos móviles, ayudando a identificar sospechosos y comportamientos relevantes para el caso. Se utiliza en investigaciones de delitos graves, como terrorismo y crimen organizado, donde la evidencia de video es clave.

- **TENSOR (Retrieval and analysis of heterogeneous data for predicting and mitigating violent actions)**. TENSOR es un proyecto de IA desarrollado también en el marco de Horizonte 2020, que tiene como objetivo detectar y analizar pruebas obtenidas en línea relacionadas con actividades terroristas o del crimen organizado. Esta herramienta puede analizar pruebas digitales heterogéneas, como publicaciones en redes sociales, videos, o correos electrónicos, identificando patrones de radicalización o posibles amenazas violentas. La IA de TENSOR analiza datos semiestructurados y no estructurados, lo que permite una identificación rápida de amenazas en tiempo real.

- **COPKIT**. De igual manera, cabe mencionar el proyecto COPKIT que se ha centrado en el problema de analizar, investigar, mitigar y prevenir el uso de las nuevas tecnologías de la información y la comunicación por parte del crimen organizado y los grupos terroristas. Para ello, COPKIT propone un sistema de alerta temprana (EA) y acción temprana (EA) basado en inteligencia, tanto a nivel estratégico como operativo. De acuerdo con su funcionalidad, cada componente desarrollado durante el proyecto se describe en una de las seis fases del ecosistema de alerta temprana/acción temprana de COPKIT: recopilación de datos, extracción de información, enriquecimiento de la información, descubrimiento de conocimientos, evaluación y previsión. La duración del proyecto fue de 40 meses (de 2018 a 2021). Está coordinado por Isdefe , una empresa pública propiedad del Ministerio de Defensa español. El consorcio está formado por 18 socios de 13 países con diferentes perfiles y experiencia en varios ámbitos, incluidos organismos de seguridad pública, industria, academia y organizaciones de investigación y tecnología. Además, el proyecto COPKIT otorga un papel clave al " Consejo Asesor de Usuarios Finales y Partes Interesadas " (EUSAB). Este consejo externo e independiente está dirigido por EUROPOL e incluye usuarios finales y expertos de diferentes campos que asesoran al consorcio sobre la implementación práctica de los objetivos del proyecto, así como sobre cuestiones relacionadas con su dirección, progreso, resultados y entregables.[160]

- **ROXANNE (Real-time network, text, and speech analytics for combating organized crime and terrorism)**. Este proyecto, que también forma parte del programa Horizonte 2020, integra IA para analizar datos de redes sociales, textos y pruebas de audio obtenidas en investigaciones criminales. ROXANNE emplea tecnologías de reconocimiento de voz y procesamiento del lenguaje natural (NLP) para identificar redes de comunicación entre sospechosos de crimen organizado o terrorismo. Además, ofrece una plataforma colaborativa para compartir pruebas y análisis entre agencias policiales y judiciales de diferentes países de la Unión Europea.

160 https://copkit.eu/

- **iCOP (Identifying and Catching Online Predators)**. Este proyecto de IA, desarrollado para combatir el abuso sexual infantil en línea, analiza grandes volúmenes de datos obtenidos en investigaciones sobre explotación sexual infantil. La herramienta iCOP utiliza algoritmos de IA para identificar patrones de comportamiento sospechoso en redes de intercambio de archivos y comunidades en línea. iCOP puede escanear archivos multimedia en busca de contenido ilegal y proporcionar pruebas clave en investigaciones penales sobre este tipo de delitos.

- Asimismo, en 2018, se desarrolló una herramienta que era capaz de detectar qué denuncias son falsas en casos de robos con violencia e intimidación o tirones[161]: **Veripol**. Está basada en un algoritmo y un modelo matemático que, gracias a la inteligencia artificial y al procesamiento del lenguaje natural, es capaz de localizar las palabras que las presuntas víctimas más utilizan cuando mienten a la hora de denunciar unos hechos. Según el estudio de sus creadores, entre los que destaca Miguel Camacho, inspector de Policía y actual coordinador de Innovación Tecnológica y Ciberseguridad del Consejo de Estado, su porcentaje de acierto a la hora de determinar qué denuncias son falsas es del 91 %, cifra muy superior al 75 % que obtuvo un policía experto

ASISTENCIA EN LA TOMA DE DECISIONES Y RESOLUCIONES JUDICIALES ASISTIDAS POR IA

*E*n este apartado es conveniente destacar el proyecto ***JuLIA*** **(Justice, fundamentaL rIghts and Artificial Intelligence Applications)**, que se trata de una iniciativa financiada por la Unión Europea que se centra en proporcionar formación sobre IA a jueces europeos, específicamente en relación con los derechos fundamentales y la toma de decisiones automatizadas (ADM). El proyecto busca mejorar la comprensión de los efectos de la IA en los sistemas judiciales y su impacto en los derechos fundamentales, como el derecho a un juicio justo y la protección contra la discriminación.

JuLIA ofrece cursos en línea y presenciales, incluyendo talleres transnacionales para jueces, con el objetivo de capacitarles en el uso de herramientas de IA dentro del marco legal europeo. Además, el proyecto contempla la creación de un "Casebook" sobre los límites del gobierno algorítmico, enfocado en el uso de la IA en la administración pública. Esta formación busca garantizar que los jueces estén equipados para enfrentar los desafíos que la IA presenta en la protección de derechos fundamentales y el proceso judicial.

Colaboran varias instituciones, entre ellas la Facultad de Derecho de la Universidad de Groningen, junto con otros seis socios internacionales, y sigue el trabajo previo del **proyecto *FRICoRe* (Fundamental Rights in Courts and Regulation)** que concluyó en 2022. No obstante, este proyecto se encuentra en fase de implementación y prevé un taller transnacional en 2024.

SEGUIMIENTO EN LA EJECUCIÓN DE LAS PENAS IMPUESTAS EN SENTENCIA

La IA ha comenzado a desempeñar un rol importante en el monitoreo y seguimiento del cumplimiento de las penas impuestas por los juzgados y tribunales penales en varios países de la Unión Europea, aunque su implementación varía en función del nivel de adopción tecnológica de cada estado miembro y de sus sistemas judiciales. La IA no solo mejora la eficiencia en la gestión de casos, sino que también proporciona herramientas para un control más preciso y eficaz del cumplimiento de las sentencias, particularmente en relación con la libertad condicional, las medidas alternativas a la prisión, y el seguimiento de delincuentes reincidentes.

Uno de los aspectos clave donde ha mostrado su utilidad es en la supervisión y gestión de medidas alternativas a la privación de libertad, como la libertad condicional, el arresto domiciliario o localización permanente con dispositivos electrónicos y otras sanciones que permiten la reintegración social de los infractores sin recurrir al encarcelamiento. Es el ejemplo de Países Bajos, donde los algoritmos de IA ayudan a predecir la probabilidad de reincidencia o incumplimiento de las condiciones establecidas, analizando patrones de comportamiento previos. De esta forma, mejoran la interoperabilidad entre diferentes agencias y permiten un control más eficaz sobre el comportamiento de los condenados bajo medidas de libertad condicional, basándose en datos históricos y en variables sociodemográficas o contextuales que ofrecen una visión más detallada del riesgo que representa cada condenado. Por su parte, su sistema penitenciario ha integrado tecnologías de IA para monitorizar y analizar el comportamiento de los reclusos en tiempo real a través de dispositivos portátiles y sistemas de vigilancia. Estos datos son analizados mediante algoritmos que detectan patrones de comportamiento anómalo que podrían señalar el riesgo de fuga o violencia. Los avances en análisis predictivo permiten a las autoridades intervenir antes de que se produzcan incidentes.

En España cuentan con el **Sistema VioGén** (Sistema de Seguimiento Integral en los casos de Violencia de Género), que es una herramienta desarrollada por el Ministerio del Interior de España para mejorar el seguimiento y la protección de las víctimas de violencia de género. Creado en 2007, este sistema integra información sobre las víctimas y los agresores, lo que permite una evaluación constante de los riesgos y una mejor coordinación entre las autoridades encargadas de la protección y el cumplimiento de las medidas y penas impuestas.

El sistema envía alertas automáticas cuando se detectan situaciones de riesgo o cuando el agresor incumple las medidas cautelares o penas impuestas por el tribunal, como las órdenes de alejamiento o la prohibición de comunicación con la víctima. El sistema permite que las fuerzas de seguridad reciban alertas en tiempo real sobre posibles incumplimientos. Esto permite que las fuerzas de seguridad actúen de manera proactiva, protegiendo a la víctima antes de que se produzca un incidente violento. La coordinación entre las policías locales, la Policía Nacional, la Guardia Civil y otras autoridades es fundamental en este aspecto.

VioGén ha sido utilizado en más de 800.000 casos desde su creación, y según el Ministerio del Interior, en más del 90% de los casos en los que se ha activado una alerta de riesgo extremo o alto, se ha conseguido evitar un nuevo episodio violento. La plataforma está en constante evolución, incorporando mejoras basadas en las experiencias de los usuarios y en los avances tecnológicos, como el uso de big data y machine learning para mejorar la precisión de las evaluaciones de riesgo. El éxito de VioGén ha sido tal que otros países han mostrado interés en desarrollar sistemas similares, y la Unión Europea ha destacado el modelo español como un ejemplo de buenas prácticas en la lucha contra la violencia de género y el seguimiento del cumplimiento de las penas y medidas de protección. No obstante es importante tener en cuenta elementos clave de transparencia y aplicabilidad cuando se desarrollan aplicaciones o sistemas como VioGén y RisCanvi.

Eurojust, la agencia de cooperación judicial de la UE, también está desarrollando tecnologías de IA para mejorar la cooperación transfronteriza en la implementación y seguimiento de sentencias penales, facilitando la interoperabilidad entre diferentes sistemas judiciales. Estas tecnologías están ayudando a asegurar que las sentencias impuestas en un país miembro sean debidamente ejecutadas o reconocidas en otro, reduciendo el tiempo de respuesta y mejorando la eficiencia en la administración de la justicia penal en la región.

161 Quijano-Sánchez, L., Liberatore, F., Camacho-Collados, J., & Camacho-Collados, M. (2018). Aplicación automática de detección de lenguaje engañoso a informes policiales: extracción de patrones de comportamiento de un modelo de clasificación de varios pasos para comprender cómo mentimos a la policía. Knowledge-Based Systems , 149, 155-168.

4.4. HERRAMIENTAS DE IA UTILIZADAS EN PAÍSES DE AMÉRICA LATINA Y EL CARIBE

A continuación, se enlistan algunas herramientas de IA que actualmente son o que fueron utilizadas en algunos países de la región. Es importante mencionar que la lista de herramientas o sistemas desarrollados puede estar incompleta tanto en los ámbitos de seguridad como de justicia, así como en los países que disponen de ellas.

ARGENTINA

PROMETEA es un sistema de IA desarrollado con el propósito de mejorar la eficiencia y la automatización en la toma de decisiones judiciales y administrativas. Fue creado por el Ministerio Público Fiscal de la Ciudad Autónoma de Buenos Aires y el Laboratorio de Innovación e Inteligencia Artificial de la Facultad de Derecho de la Universidad de Buenos Aires[162]. Entre las principales bondades de PROMETEA se encuentran:

- Capacidad de automatizar la redacción de documentos judiciales y administrativos, como dictámenes, resoluciones y sentencias con gran rapidez y precisión.
- Entender y procesar textos normativos y judiciales complejos, extrayendo la información relevante para emitir una recomendación o redactar una resolución.
- Ayuda a clasificar, priorizar y gestionar grandes volúmenes de casos, identificando aquellos que requieren atención urgente o que pueden resolverse rápidamente mediante procedimientos automatizados
- Garantiza que cada decisión o recomendación generada por el sistema sea explicable.

BRASIL

VICTOR. Es una herramienta de IA utilizada por el Tribunal Supremo Federal del Brasil que analiza y clasifica los recursos de apelación presentados ante dicho tribunal y tiene la capacidad de predecir si el recurso solicitado tendrá un amplio impacto social y de relevancia para que merezca ser sometido al estudio y análisis por los Magistrados que integran ese tribunal. El proyecto fue nombrado como tributo a Víctor Nunes Leal, un Magistrado brasileño encargado de sistematizar la jurisprudencia del TSF y de facilitar la aplicación de los precedentes judiciales aplicables a los juicios de apelación en Brasil[163].

162 BID, "PROMETEA, Transformando la administración de justicia con herramientas de inteligencia artificial" en: https://publications.iadb.org/es/publications/spanish/viewer/PROMETEA-Transformando-la-administracion-de-justicia-con-herramientas-de-inteligencia-artificial.pdf

163 Habib Lantyer, Victor, «The Era of Artificial Intelligence in Law: Brazil in a Global Context » (December 1, 2023), p.10, disponible en SSRN: https://ssrn.com/abstract=4650117

CHILE

En el área de predicción de delitos, destaca el Sistema Predictivo del Delito Urbano que fue desarrollado en 2017 por la Universidad de Chile junto con el Departamento de Análisis Criminal de Carabineros de Chile, cuyo objetivo es predecir zonas de mayor riesgo de ocurrencia de delitos con la finalidad de realizar patrullajes focalizados y de reforzar la eficacia del sistema de persecución penal[164].

COLOMBIA

PRETORIA es un sistema de IA que apoya y optimiza el proceso de selección, análisis y estructuración de las sentencias de tutela (acción constitucional para proteger derechos fundamentales) de la Corte Constitucional de Colombia. Fue desarrollado conjuntamente por el Laboratorio de Innovación e Inteligencia Artificial de la Universidad de Buenos Aires, la Universidad del Rosario y otras instituciones públicas y privadas de Colombia en 2020[165].

Este sistema lleva a cabo actividades de selección, análisis y estructuración de las sentencias de tutela para revisión de la Corte Constitucional a través de tres funciones:

- (i) Búsqueda. Permite identificar información de interés para la selección de las sentencias;
- (ii) Categorización. Permite categorizar y seleccionar la información conforme a criterios relevantes para la Corte Constitucional; y
- (iii) Estadísticas. Permite la creación de líneas de tiempo y gráficos que ayudan a tener una visión más completa e integral sobre la tutela.

PRISMA (Perfil de Riesgo de Reincidencia para Solicitud de Medidas de Aseguramiento). Es una herramienta de IA utilizada para predicciones de riesgo de reincidencia a partir del perfil del investigado que tiene como propósito apoyar la labor de los fiscales y jueces para determinar una medida de aseguramiento como por ejemplo la detención preventiva de un individuo que está siendo investigado por las autoridades colombianas. La herramienta puede mostrar la predicción de riesgo de reincidencia criminal que pueda llegar a tener el investigado durante el proceso penal. Esta herramienta busca optimizar la gestión de los cupos carcelarios y apoya a los jueces en materia penal mostrando información de personas con una mayor probabilidad de reincidencia criminal[166].

164 Fair Trials, «Inteligencia artificial en la seguridad pública y en el sistema penal en América Latina. Análisis basado en el debido proceso», diciembre 2024, p.17, disponible en: https://www.fairtrials.org/app/uploads/2024/08/Inteligencia-artificial-en-la-seguridad-publica-y-en-el-sistema-penal-en-America-Latina.pdf

165 Corte Constitucional Republica de Colombia, "PRETORIA, sistema inteligente de la Corte Constitucional para apoyar la selección de tutelas, es premiada como mejor herramienta de modernización en materia de justicia por la CEJ » Boletin No. 187, Bogota 15 de diciembre de 2020, en : https://www.corteconstitucional.gov.co/noticia.php?PRETORIA,-sistema-inteligente-de-la-Corte-Constitucional-para-apoyar-la-selecci%C3%B3n-de-tutelas,-es-premiada-como-mejor-herramienta-de-modernizaci%C3%B3n-en-materia-de-justicia-por-la-CEJ-9031

166 Una explicación y ejemplos sobre el Sistema PRISMA, ver: Fiscalía General de la Nación. Dirección de Políticas Públicas y Estrategia. "Herramienta PRISMA: Perfil de Riesgo de Reincidencia para la Solicitud de Medidas de Aseguramiento" en: https://www.fiscalia.gov.co/colombia/wp-content/uploads/Perfil-de-riesgo-de-reincidencia-para-solicitudes-de-medida-de-aseguramiento.pdf

La OCDE reporta acerca de una herramienta de IA relacionada con la predicción de sentencias en juicios contra el Estado en Colombia desarrollada por la Agencia Nacional de Defensa Jurídica del Estado (ANDJE) y Quantil, una empresa privada que consiste en una herramienta matemática para estimar la probabilidad de una sentencia desfavorable en un proceso litigioso contra la nación, y recomendar el monto óptimo de un acuerdo basado en las condiciones vigentes del caso. De acuerdo con la OCDE, el componente predictivo del modelo se basa en técnicas de aprendizaje automático, mientras que la optimización del arreglo conciliatorio se basa en fundamentos financieros y de teoría de los juegos[167].

En el área de vigilancia y protección ciudadana, la OCDE reporta que en la región de ALC se está haciendo un mayor uso experimental de la IA para analizar imágenes de rostros juntamente con otros videos, imágenes y audio con la finalidad de detectar actividades criminales e identificar delincuentes y hace referencia a los casos del Centro de Comando, Control, Comunicaciones y Cómputo (C4) de Bogotá, Colombia, y del ECU 911 en Ecuador[168].

TALION se trató de un prototipo de sistema experto de lógica difusa para la determinación de la pena según los criterios y parámetros establecidos en la ley y que se presentó en el lejano 2009 en un concurso del Ministerio de Tecnologías de la Información y Comunicación de Colombia[169]

COSTA RICA

El Poder Judicial, a través de la Comisión de Protección de Datos implementó en marzo de 2024 una herramienta de IA que permite realizar los procesos de despersonalización de documentos con el objeto de proteger la información sensible contenida en las sentencias que se encuentran dentro del sistema Nexus. PJ., que utiliza el Poder Judicial de Costa Rica con el fin de tutelar los derechos que le asisten a las personas usuarios conforme a la Ley No. 8968 de Protección de la Persona frente al tratamiento de sus datos[170].

MÉXICO

SORJUANA. Es una herramienta de IA construida con base en herramientas de programación de terceras empresas como Streamlit, Google y Pinecone que fue desarrollada por la ponencia de la Ministra Ana Margarita Ríos Farjat de la Suprema Corte de Justicia de la Nación (SCJN) y cuyo principal propósito es facilitar la revisión, comprensión y socialización del contenido de las versiones públicas y sentencias de la SCJN. La herramienta se encuentra en fase de prueba y solo considera los casos de la ministra Ríos Farjat, más no la totalidad de los casos que están siendo analizados por la SCJN. El sistema entrega respuestas resumidas sobre el contenido de las sentencias y permite a los usuarios solicitar explicaciones sobre capítulos específicos de leyes y reglamentos vigentes relacionados con las sentencias[171].

En México también se desarrolló EXPERTIUS, un prototipo de sistema experto basado en el paradigma simbólico o top-down. Ha sido desarrollado en el Instituto de Investigaciones Jurídicas y el Centro de Ciencia Aplicada y Desarrollo Tecnológico (ambos de la UNAM), bajo el auspicio del Consejo Nacional de Ciencia y Tecnología, en colaboración del Tribunal Superior de Justicia del Estado de Tabasco. Su finalidad es ayudar a la toma de decisiones y a la homogenización del conocimiento colectivo de la comunidad judicial, en el juicio especial de alimentos para determinar pensión económica. Expertius II es una evolución con un modelo más complejo que pretende simular el razonamiento lógico del juez con redes neuronales en sistemas tipo Deep y machine learning[172].

URUGUAY

La Fundación World Wide Web reporta acerca de la implementación del sistema predictivo de delito PredPol entre 2014 y 2017 en Uruguay cuya licencia fue adquirida por el Ministerio de Interior con el objeto de ofrecer información y modelos estadísticos sobre predicción de delitos en determinadas áreas geográficas. Se reporta que su uso fue descontinuado por no haberse reducido los índices delincuenciales en este país y por no haber cumplido sus objetivos generales [173].

167 OCDE/CAF, «Casos Prácticos de uso de la IA en los Gobiernos de América Latina y el Caribe» en Uso estratégico y responsable de la inteligencia artificial en el sector público de América Latina y el Caribe, p.41, en: https://www.oecd-ilibrary.org/docserver/6150ef8b-es.pdf?expires=1728901099&id=id&accname=guest&checksum=0E79838E53C73FBED198E22AC70437A8

168 OCDE/CAF, «Casos Prácticos de uso de la IA en los Gobiernos de América Latina y el Caribe» *Op. cit.*, nota 61, pp. 46-47.

169 https://www.calameo.com/read/000099861d3cfe698d294

170 Poder Judicial de Costa Rica, «Novedosa herramienta de Inteligencia Artificial se aplica en mejora de la protección de datos », comunicado de prensa sin fecha, en: https://pj.poder-judicial.go.cr/index.php/prensa/1186-novedosa-herramienta-de-inteligencia-artificial-se-aplica-en-mejora-de-la-proteccion-de-datos

171 El portal del sistema Sor Juana se encuentra en: https://ponenciamamrfgpt.streamlit.app/

172 CACERES NIETO, E. (2023). La inteligencia artificial aplicada al derecho como una nueva rama de la teoría jurídica. Anales de la Cátedra Francisco Suárez, 57, 63–89. https://doi.org/10.30827/acfs.v57i.26281

173 World Wide Web Foundation, «Algoritmos e Inteligencia Artificial en Latinoamérica. Un Estudio de implementaciones por parte de gobiernos en Argentina y Uruguay », septiembre 2018, pp. 27-30, en: https://webfoundation.org/docs/2018/09/WF_AI-in-LA_Report_Spanish_Screen_AW.pdf

Ver también OCDE/CAF, «Casos Prácticos de uso de la IA en los Gobiernos de América Latina y el Caribe» en Uso estratégico y responsable de la inteligencia artificial en el sector público de América Latina y el Caribe, pp. 45-46.

BLOQUE 5: RECOMENDACIONES DE ACTUACIÓN Y CONCLUSIONES

5.1. RECOMENDACIONES DE ACTUACIÓN

La creciente adopción y expansión de la inteligencia artificial en América Latina y el Caribe ha transformado profundamente el panorama del crimen organizado en la región. Las organizaciones criminales están aprovechando las capacidades avanzadas de la IA para desarrollar nuevas formas de delitos que desafían las estrategias tradicionales de las autoridades. La capacidad de generar identidades sintéticas o "personas falsas" mediante tecnologías como los deepfakes representa una amenaza significativa que socava la confianza en las instituciones y en las interacciones sociales. Al igual que el dinero falsificado puede desestabilizar economías y erosionar la confianza en los sistemas financieros, las "personas falsificadas" tienen el potencial de desestabilizar sistemas sociales y políticos al facilitar fraudes, extorsiones y campañas de desinformación.

Ante este desafío, es imperativo que los países de América Latina y el Caribe adopten un enfoque regional y colaborativo para desarrollar estrategias efectivas que combatan el uso indebido de la IA por parte del crimen organizado. A continuación, se presentan recomendaciones específicas dirigidas a abordar estos desafíos desde una perspectiva regional.

1. **Definición y comprensión de la IA.** Se recomienda publicar una definición clara de lo que la IA implica para cada organismo. Fomentar un entorno de aprendizaje e investigación sobre la IA y su posible impacto en la actividad policial.

2. **Creación de un marco de cooperación regional.** Dado que el crimen organizado opera a través de fronteras nacionales, es esencial establecer mecanismos de cooperación regional que permitan el intercambio de información, buenas prácticas y recursos técnicos. La creación de un organismo regional dedicado a la ciberseguridad y al combate del uso indebido de la IA puede facilitar la coordinación entre países y fortalecer la capacidad colectiva para enfrentar estas amenazas. Para reducir esta laguna jurídica de la impunidad, un instrumento internacional con nuevos delitos o circunstancias agravantes relacionadas con la inteligencia artificial facilitaría abolir los controles de doble incriminación y puede garantizar una confianza mutua más rápida y una cooperación internacional mucho más eficaz entre los tribunales. También ayuda a perseguir nuevos delitos y a reducir la impunidad de los mismos, e intenta aumentar la indemnización por daños y perjuicios a las víctimas.

3. **Actualización de marcos regulatorios de protección de datos.** Fortalecer las leyes de protección de datos personales puede limitar la disponibilidad de información que las organizaciones criminales necesitan para generar identidades falsas. Una regulación más estricta sobre el acceso y manejo de datos sensibles puede reducir el riesgo de que esta información sea explotada por actores maliciosos.

4. **Estrategias nacionales sobre IA y sub-estrategias para autoridades de seguridad, justicia y el poder judicial**. Diseñar estrategias conjuntas sobre IA que contengan objetivos y metas específicas que sean medibles en la implementación y el uso de sistemas de IA por parte de las autoridades de justicia y el poder judicial. Estos objetivos pueden ir desde mejorar la seguridad pública, reducir los índices de delincuencia, aumentar la eficiencia de la gestión de casos investigados o mejorar la imparcialidad de los procesos judiciales. Los objetivos deberán será medibles en el corto y mediano plazo y las estrategias deben ser revisadas y actualizadas periódicamente.

5. **Transparencia en los procesos de integración de la IA en los procedimientos de la lucha contra el Crimen Organizado.** Es fundamental que las instituciones públicas competentes y las agencias de investigación criminal generen un entorno y un ambiente de trasparencia, en el que, desde la responsabilidad, no solo legal sino también ética, y mediante la cooperación con la comunidad, se alcance el uso más eficiente de la IA en la investigación criminal, generando a la vez confianza en la ciudadanía. Comprometerse con la comunidad para explicar cómo se utilizará la IA en la actuación policial y recabar sus opiniones. Establecer canales para que la comunidad realice aportaciones continuas y haya transparencia en relación con el uso de la IA.

6. **Prohibición y regulación de la creación de identidades sintéticas.** Se recomienda que los países de la región establezcan marcos legales que prohíban explícitamente la creación y distribución de "personas falsas" o identidades sintéticas con fines ilícitos. Esta regulación debe equiparar la fabricación de identidades digitales falsas con la falsificación de moneda, reconociendo el potencial dañino que ambas prácticas tienen sobre la sociedad y la economía. Además, es fundamental que estas leyes contemplen sanciones penales y civiles adecuadas para disuadir y castigar estas actividades.

7. **Fortalecimiento de la legislación sobre delitos informáticos.** Los códigos penales tanto sustantivos como procedimentales de los países latinoamericanos deben actualizarse para incluir delitos o circunstancias agravantes relacionadas con el uso malicioso de la IA, como la generación y difusión de deepfakes, la manipulación algorítmica y el fraude automatizado. Esto permitirá a las autoridades perseguir y sancionar eficazmente a quienes utilicen estas tecnologías para actividades delictivas.

8. **Implementación de estándares éticos en el desarrollo de IA.** Promover la adopción de principios éticos en el desarrollo y despliegue de tecnologías de IA es esencial para prevenir su uso malicioso. Esto incluye la incorporación de medidas de seguridad que dificulten la creación de identidades sintéticas con fines ilícitos y la promoción de prácticas responsables entre desarrolladores y científicos de datos y establecer una estructura de gobernanza para supervisar el uso ético y responsable de la IA, designando funciones y responsabilidades, y asegurando la rendición de cuentas.

9. **Establecimiento de unidades especializadas en cibercrimen.** Crear y fortalecer unidades especializadas dentro de las fuerzas del orden dedicadas exclusivamente al cibercrimen y al uso indebido de la IA permitirá una respuesta más rápida y efectiva. Estas unidades deben contar con personal capacitado y recursos tecnológicos adecuados para enfrentar amenazas altamente sofisticadas cometidas a través de sistemas y aplicaciones de IA.

10. **Promoción de la investigación y desarrollo en seguridad de IA.** Fomentar la investigación académica y científica en seguridad de IA ayudará a desarrollar nuevas técnicas y herramientas para contrarrestar las amenazas emergentes. El apoyo a centros de investigación regionales puede impulsar soluciones innovadoras adaptadas al contexto latinoamericano. De igual

manera deben fomentarse e impulsarse las certificaciones y estándares internacionales de calidad de seguridad de la información como ISO 27001, Esquema Nacional de Seguridad español o Common Criteria.

11 **Desarrollo de campañas de concientización pública.** Es fundamental educar a la población sobre los riesgos asociados con las "personas falsas" y otras formas de manipulación mediante IA. Campañas de concientización pueden ayudar a los ciudadanos a identificar señales de fraude o desinformación, promoviendo una cultura de escepticismo saludable y precaución en interacciones digitales.

12 **Fortalecimiento de las capacidades de monitoreo en los centros penitenciarios.** Considerando el potencial de los sistemas de IA para mejorar la gestión, eficiencia y seguridad en los centros y comunidades penitenciarias, se recomienda una mayor difusión y uso de este tipo de herramientas para medir y predecir posibles amenazas, identificar perfiles criminales de alto riesgo para la población penitenciaria, estableciendo mecanismos de supervisión y auditoría y garantizando el cumplimiento de principios éticos y responsables tales como la transparencia, explicabilidad y supervisión humana.

13 **Realizar una evaluación exhaustiva de los riesgos para comprender los posibles retos y amenazas asociados a la IA.** Establecer un mecanismo de evaluación de impacto de derechos fundamentales y análisis y gestión de riesgos con adopción de medidas técnicas, organizativas y legales de reducción y mitigación continua de los riesgos a medida que evolucionan las tecnologías de IA.

14 **Fomento de la colaboración público-privada.** Las empresas tecnológicas y proveedores de servicios digitales juegan un papel crucial en la detección y prevención del uso indebido de la IA. Se recomienda establecer alianzas entre el sector público y privado para desarrollar soluciones tecnológicas que puedan identificar y bloquear contenidos maliciosos, así como compartir información sobre amenazas emergentes e implementar mecanismos agiles para obtener la preservación de datos informáticos almacenados (datos de abonado y datos de tráfico) de los proveedores de servicios que puedan ser útiles para las autoridades de justicia en las investigaciones penales relacionadas con cualquier tipo de delito.

15 **Planificar las adquisiciones de IA.** Comprender los problemas que pretende resolver con la IA y adaptar el enfoque de contratación para garantizar que las soluciones elegidas satisfacen las necesidades y cumplen las políticas establecidas.

16 **Inversión en capacitación y recursos tecnológicos.** Establecer programas de formación para policía, fiscalía y poder judicial diseñados para entender el funcionamiento de las aplicaciones de IA mayormente utilizadas con propósitos delictivos por el crimen organizado. Esto incluye la adquisición de herramientas avanzadas para la detección y análisis de actividades delictivas que emplean IA, así como la formación especializada para interpretar y utilizar eficazmente estos recursos en investigaciones y procesos judiciales.

17 **Implementar programas de formación para desarrollar los conocimientos y habilidades necesarios para aprovechar la IA.** Fomentar la formación interdepartamental para garantizar un enfoque unificado de la adopción de la IA.

18 **Establecer procesos de evaluación rápida de viabilidad de los proyectos de integración de soluciones de IA en los procedimientos de investigación criminal.** Por ejemplo, en línea con lo exigido por el Parlamento Europeo[174] y lo sugerido por la *Carta ética europea sobre el uso de la inteligencia artificial en los sistemas judiciales y su entorno*[175] , podría diseñarse una lista de verificación que ayuda a hacer una primera y rápida evaluación de riesgos y beneficios de cada potencial proyecto, incluyendo aspectos relevantes tanto para su viabilidad legal como para el valor procesal de productos que pueda generar la solución IA planteada. Esta lista rápida verificación ayudaría a priorizar, o incluso descartar, de una manera ágil, los proyectos y la dedicación de esfuerzos a cada uno de ellos. Los extremos objeto de chequeo incluidos en dicha lista podrían ser los que se enumeran a continuación, y el resultado de su valoración solo puede ser una de dos opciones: FAVORABLE/DESFAVORABLE:

- Está dirigido a la protección y beneficio de todos los miembros de la sociedad.
- Mejora los métodos de trabajo y las capacidades de investigación criminal de la Institución.
- Puede producir discriminación y contener sesgos.
- Evaluación de impacto de afectación a DDFF.
  - Afecta a DDFF como el derecho a la intimidad, presunción de inocencia, tutela judicial efectiva y un juez imparcial.
  - Implica el tratamiento de datos personales.
  - Implica la vigilancia masiva.
  - Implica análisis automatizados o el reconocimiento en espacios accesibles al público de características humanas, como los andares, las huellas dactilares, el ADN, la voz y otras señales biométricas y de comportamiento.
- Ofrece un estándar de calidad y seguridad aceptable.
  - Es fiable.
  - Es consistente.
  - Es auditable.
  - Garantiza la atribución de responsabilidades legales en caso de efectos nocivos.
- Es accesible, trasparente, inteligible y explicable.
- Respeta la autonomía humana. Es decir, garantiza que la decisión final será tomada por un humano.

174 Resolución del Parlamento Europeo, de 6 de octubre de 2021, sobre la inteligencia artificial en el Derecho penal y su utilización por las autoridades policiales y judiciales en asuntos penales (2020/2016(INI)

175 Carta ética europea sobre el uso de la inteligencia artificial en los sistemas judiciales y su entorno (2018). https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52021IP0405

5.2. CONCLUSIONES

Abordar el papel de la inteligencia artificial no solo como una herramienta poderosa para el desarrollo y la innovación, así como para agilizar procesos en la administración de justicia, seguridad y del sistema penitenciario o la investigación criminal, particularmente en el análisis de grandes volúmenes de datos y la identificación de patrones complejos, sino también como un factor que, si no se gestiona adecuadamente, puede potenciar actividades delictivas complejas y de alto impacto. Actividades que ya estamos viendo en la vida cotidiana con técnicas de phising, deepfakes y fraudes, entre otros.

Es crucial que tanto los gobiernos como las organizaciones de seguridad se adapten a esta nueva realidad, desarrollando y aplicando tecnología responsable, ética y objetiva que permita contrarrestar el uso indebido de la IA. Para ello será necesario desarrollar marcos regulatorios sólidos que permitan, por un lado, el florecimiento de empresas tecnológicas de IA y su utilización por parte del conjunto de la sociedad y de los ecosistemas empresariales y gubernamentales; y, por otro, que tipifiquen y persigan su uso indebido tanto por las organizaciones criminales como la mala utilización de instituciones estatales.

El presente estudio realizado desde un punto de vista holístico y multisectorial concluye los siguientes puntos clave:

- **Necesidad de un marco regulatorio sólido y ético**: La implementación de inteligencia artificial en la lucha contra la delincuencia organizada plantea importantes preguntas éticas, de privacidad y de sesgo de género, además de protección de datos. Existe una delgada línea entre la vigilancia para la seguridad y la violación de derechos fundamentales. Por lo tanto, se recomienda que los países de la región generen reformas al marco jurídico sustantivo y procesal penal para regular el uso de la IA con fines y propósitos delictivos. En caso de falta de legislación, la creación de un marco regulatorio robusto que supervise tanto el uso de IA por parte de las autoridades como su posible uso indebido. Este marco debe equilibrar la protección de los derechos de los ciudadanos con la seguridad nacional y el orden público. El Convenio Marco del Consejo de Europa sobre Inteligencia Artificial y derechos humanos, democracia y Estado de derecho (CETS No. 225) y el Reglamento Europeo de IA son dos ejemplos que podrían ser utilizadas tanto para ser ratificados por los países interesados, en el caso del Convenio Marco del Consejo de Europa, como para la redacción de unas normas mínimas sobre IA en América Latina y el Caribe. Asimismo, se confía y habrá que estar atento a posibles resultados relacionados con un futuro instrumento del Consejo de Europa sobre inteligencia artificial y derecho penal encargado por el Comité de Ministros del Consejo de Europea al CDPC.

- En el **ámbito procesal**, la inclusión de obligaciones de colaboración con los proveedores de tecnología y de sistemas de IA en la investigación de delitos graves por las autoridades de justicia resultaría de gran relevancia. Dichas colaboraciones se llevan haciendo con cierta fluidez en aspectos vinculados al abuso sexual a menores. No obstante, en temas de IA la colaboración debe ser estandarizada, regulada y forzada si fuere necesario dado el impacto que tiene la IA en multitud de delitos. De igual manera, debe establecerse tipificado un delito de desobediencia expreso para en aquellos casos que no se respondan los requerimientos judiciales.

- **Estrategias regionales, nacionales y lineamientos específicos**. La creación de estrategias nacionales y lineamientos sobre IA para las autoridades del sistema de justicia de la región requieren ser priorizadas. Las estrategias deben ser desarrolladas en conjunto por las autoridades investigadoras del sistema de justicia y el poder judicial y establecer mecanismos de coordinación y colaboración con los proveedores de tecnología y de sistemas de IA. Las estrategias deben incluir mecanismos de participación de expertos en IA de la comunidad científica y académica que puedan ayudar a contribuir en la labor de las actividades al combate al crimen organizado, y en particular en la prevención de la protección de las víctimas de delitos cometidos a través de sistemas de IA.

- **Desarrollo de herramientas, digitalización de procesos y obtención de datos de calidad**. Dotar a las autoridades de seguridad, justicia y al poder judicial en la región con herramientas y sistemas de IA que faciliten y optimicen la labor de sus actividades y que puedan a ayudar a combatir de mejor forma al crimen organizado y a prevenir el delito en sus diversas modalidades. Las herramientas de IA deben ser previamente valoradas y probadas por las autoridades, deberán cumplir con principios éticos y responsables, haberse sometido a evaluaciones de impacto, facilitar información sobre al menos las reglas, parámetros, lógica aplicada, importancia y consecuencias de los códigos fuente conocidos y contar con índices de medición que ayuden a alcanzar objetivos y metas concretas durante su implementación en el ámbito de cada organización. Además, se recomienda el desarrollo de herramientas de IA porque su desarrollo interno favorece la digitalización de procesos para poder nutrir a la herramienta de información y datos. Sin datos correctamente obtenidos y de buena calidad se correría el riesgo de desarrollar herramientas con sesgos y tendencias no previstas.

- **Capacitación y desarrollo de habilidades para combatir la tecnología criminal**: La eficacia de la lucha contra la delincuencia organizada que emplea IA depende, en gran medida, de la capacitación adecuada de los cuerpos y fuerzas de seguridad, de las autoridades de justicia y del conjunto de operadores de la cadena penal, así como del desarrollo de equipos de expertos en ciberseguridad, ciberdelincuencia e inteligencia artificial.

- **Aumento de la complejidad y sofisticación de las actividades delictivas. Dificultad en la detección y prevención**: La capacidad de los algoritmos de IA para analizar grandes volúmenes de datos y aprender patrones complejos también es aprovechada por organizaciones delictivas para realizar ciberataques más precisos, suplantaciones de identidad mediante deepfakes, realizar operaciones de lavado de dinero de forma más ágil y evadir la detección, aumentando la sofisticación de los grupos criminales. Además, desde otro punto de vista, la IA facilita la automatización de diversas actividades ilícitas, como fraudes financieros y ciberataques coordinados. El uso de la IA para anonimizar comunicaciones y transacciones permite a los delincuentes operar con mayor seguridad y menos riesgo de detección. Esto representa una barrera significativa para la aplicación de la ley, que se ve superada en velocidad y complejidad tecnológica. Esta dinámica hace urgente la necesidad de desarrollar herramientas de IA en el campo de la seguridad para contrarrestar estas tácticas.

- **Desarrollo de alianzas estratégicas público-privadas.** La colaboración público-privada en el desarrollo de herramientas de IA para los sectores de justicia, seguridad y penitenciario es una oportunidad única para aprovechar el potencial tecnológico y transformarlo en soluciones que impacten positivamente en la vida de los ciudadanos, contribuyendo de modo directo e indirecto a prevenir el crimen y luchar contra los grupos de delincuencia organizada nacionales y transnacionales. Esta alianza estratégica permitirá a ambos sectores construir un futuro más seguro, eficiente y transparente, además de promover el desarrollo de ecosistemas tecnológicos que tendrán retornos positivos para las economías nacionales y locales.

BIBLOGRAFÍA

Abusamadov, K. (2024). Revolutionizing crime prevention: The role of AI and big data in modern law enforcement. *Journal of Law, Market & Innovation, 1*(2), 21-25.

AGUILAR, Alberto R. (2023). Interior reconoce que no ha consultado a la AEPD sobre el algoritmo de reconocimiento facial que está entrenando para la policía. Business Insider. 26 de diciembre de 2022. Consultado el 1 de julio de 2023.

AIplusInfo. (2023). *How will artificial intelligence affect policing and law enforcement?* **Artificial Intelligence +.**

Amnesty International. (2018). *Toxic Twitter: Violence and abuse against women online*.

AMOS, Zac. (2023/08/11) ¿Qué es FraudGPT? HackerNoon.

ASMANN, P. (2018, Agosto 15). Are armed drones the weapon of the future for Mexico's cartels? *InSight Crime*

Blauth, T. F., Gstrein, O. J., & Zwitter, A. (2022). Artificial intelligence crime: An overview of malicious use and abuse of AI. *IEEE Access, 10*, **77110–77122.**

BID. (n.d.). *PROMETEA: Transformando la administración de justicia con herramientas de inteligencia artificial*.

CACERES NIETO, E. (2023). La inteligencia artificial aplicada al derecho como una nueva rama de la teoría jurídica. Anales de la cátedra Francisco Suárez, 57, 63–89

Caldwell, M., Andrews, J. T. A., Tanay, T., & Griffin, L. D. (2020). AI-enabled future crime. *Crime Science, 9*(1).

Centre for AI and Digital Policy (CAIDP). (2023). *Artificial Intelligence and Democratic Values Index 2023*, **pp. 56–77.**

Center for Internet Security (CIS). (n.d.). Breaking down the BlackCat ransomware operation.

CEPAL. (2024, September 24). *Latin American Artificial Intelligence Index (ILIA) keeps Chile, Brazil, and Uruguay as regional leaders*.

CEPAL/CENIA. (2023, August 7). *Latin American Artificial Intelligence Index (ILIA)*. **National Library of Congress of Chile, Department of Studies, Extension and Publications.**

Clarín. (2024, October 14). Escándalo en un colegio de San Martín: denuncian a un alumno que vendía fotos manipuladas con IA de sus compañeras desnudas.

Comisión Europea. (2018, April 25). *Communication from the Commission: Artificial Intelligence for Europe* **(COM(2018) 237 final).**

Comisión Europea. (7 febrero 2018). Visual analytics for brighter criminal intelligence analysis. En: Cordis.

Comisión Europea. (2023, March 14). *Global Gateway: EU, Latin American, and Caribbean partners launch the EU-LAC Digital Alliance in Colombia* **[Press release].**

Corte Constitucional República de Colombia. (2020, December 15). PRETORIA, sistema inteligente de la Corte Constitucional para apoyar la selección de tutelas, es premiada como mejor herramienta de modernización en materia de justicia por la CEJ [Boletín No. 187].

Corte Constitucional de la República de Colombia. (2024). Sentencia T-323 de 2024. Sala Segunda de Revisión.

Council of Europe. (2001), 'Convention on Cybercrime', ETS No.185**, 2001**

Council of Europe. (2018). *Algorithms and human rights: Study on the human rights dimensions of automated data processing techniques and possible regulatory implications*.

Council of Europe. (2020, April 8). *Recommendation CM/Rec(2020)1 of the Committee of Ministers to member states on the human rights impacts of algorithmic systems*.

Council of Europe. (2020). European Committee on Crime Problems (CPDC), '*Feasibility study on a future Council of Europe instrument on artificial intelligence and criminal law***'.**

Council of Europe. (2024). *Framework Convention on Artificial Intelligence, Human Rights and the Rule of Law*, **CETS No. 225, Vilnius 5.IX.2024.**

Deloitte. (2021). *Surveillance and predictive policing through AI*. **Deloitte Insights. https://www.deloitte.com/global/en/Industries/government-public/perspectives/urban-future-with-a-purpose/surveillance-and-predictive-policing-through-ai.html**

Durán San Juan, Isabela. (2024). ¿Cómo los robots y la inteligencia artificial transformarán las relaciones sexuales del futuro? Infobae

El Comercio. (2023, July 16). Clonan voces de personas con IA para estafar o fingir secuestros: al menos 55 casos en Perú.

ENACT. (2023). *AI and organised crime in Africa*. **Sigsworth, R. ENACT Observer.**

European Commission. (2019, April 8). *Ethics guidelines for trustworthy AI*.

European Commission. (2022). *AI Watch: National strategies on artificial intelligence - A European perspective (2022 Edition)*.

European Parliament. (2020, July). *Artificial intelligence and law enforcement: Impact on fundamental rights* **[Study requested by the LIBE Committee of the European Parliament]. González Fuster, G.**

European Parliament. (2021, October 6). *Resolución del Parlamento Europeo sobre la inteligencia artificial en el Derecho penal y su utilización por las autoridades policiales y judiciales en asuntos penales (2020/2016(INI))*.

European Union Agency for Fundamental Rights (FRA). (2022, December 8). *Bias in algorithms: Artificial intelligence and discrimination*.

European Union. (2024). *Regulation (EU) 2024/1689 of 13 June 2023 establishing harmonised rules on artificial intelligence (AI)*.

EUROPOL. (2017). *European Union serious and organised crime threat assessment (SOCTA)*. **European Union Agency for Law Enforcement Cooperation.**

EUROPOL. (2020). *Malicious uses and abuses of artificial intelligence*. **Trend Micro Research, European Union Agency for Law Enforcement Cooperation.**

EUROPOL. (2024). *AI and policing: The benefits and challenges of artificial intelligence for law enforcement* **[Observatory report from the Europol Innovation Lab].**

Faqir, R. S. A. (2023). Digital criminal investigations in the era of artificial intelligence: A comprehensive overview. *International Journal of Cyber Criminology, 17***(2), 77-94.**

Fiscalía General de la Nación, Dirección de Políticas Públicas y Estrategia. (n.d.). *Herramienta PRISMA: Perfil de riesgo de reincidencia para la solicitud de medidas de aseguramiento*.

Fortune. (2024, May 17). A deepfake 'CFO' tricked the British design firm behind the Sydney Opera House in $25 million scam.

Habib Lantyer, V. (2023, December 1). *The era of artificial intelligence in law: Brazil in a global context*. **SSRN.**

HAO, Karen trad. MILUTINOVIC Ana (2021, 14 de abril). La IA de Facebook discrimina a las mujeres en los anuncios de trabajo. *Technology Review*.

Hart, R. (2024, mayo 4). El conflicto entre Scarlett Johansson y OpenAI podría generar una guerra de las celebridades contra las empresas de IA. Forbes Argentina. Consultado el 1 de junio de 2024.

Hayward, K. J., & Maas, M. M. (2021). Artificial intelligence and crime: A primer for criminologists. *Crime, Media, Culture, 17***(2), 209–233.**

Infobae. (2023, August 29). Chorrillos: Escolares que alteraron fotos de compañeras con IA y las comercializaron no fueron expulsados.

InSight Crime. (2024, August 26). 4 ways AI is shaping organized crime in Latin America.

INTERPOL. (2023). *Global threat assessment on scams*. **INTERPOL.**

JOSEFINA GARCÍA, JON MARINA. (2024). Uso de inteligencia artificial en el mercado de valores (high-frequency trading). Pérez Llorca Techlaw 2024

Kanwel, S., Imran Khan, M., & Usman, M. (2023). From bytes to bars: The transformative influence of artificial intelligence on criminal justice. *Qlantic Journal of Social Sciences, 4***(4), 84-89.**

Keeper. (2024, September 13). Cómo hace la IA para que los ataques de phishing sean más peligrosos.

King, T. C., Aggarwal, N., Taddeo, M., & Floridi, L. (2020). Artificial intelligence crime: An interdisciplinary analysis of foreseeable threats and solutions. *Science and Engineering Ethics, 26***(1), 89–120.**

KOSINSK M. Y FORREST A. (2024). Prompt injection. IBM Research. 26 de marzo de 2024.

LAMAS LOPEZ, F., & PERALTA GUTIERREZ, A. (2023). Marco de Derecho Internacional Público y Usos Militares de la Inteligencia Artificial en la UE. Revista Electrónica De Estudios Internacionales, (46), 505–525

LÓPEZ B., Joaquín M. (2019). El sistema de inteligencia artificial para adelantarse a crímenes en Bogotá. En: La República. 22 abril 2019

MARTIN, Nacho. (2024). El Independiente. (2024, November 9). WormGPT: El ChatGPT sin restricciones que usan los ciberdelincuentes.

National School of Political and Administrative Studies. (n.d.). *Artificial intelligence – a double-edged sword: Organized crime's AI vs. law enforcement's AI.*

Naciones Unidas. (2024, March 7). *Artificial intelligence already reproduces gender stereotypes*.

OECD.AI Policy Observatory. *Catalogue of tools & metrics for trustworthy AI*.

OECD.AI Policy Observatory. *OECD AI Incidents Monitor (AIM)*.

OECD. (2022). *OECD framework for the classification of AI systems*. *OECD Digital Economy Papers* **No. 323. OECD Publishing.**

OECD. (2024, April 24). *Report of the implementation of the OECD recommendation on artificial intelligence* **(C/MIN(2024)17).**

OECD & CAF Development Bank of Latin America. (2022). *The strategic and responsible use of artificial intelligence in the public sector of Latin America and the Caribbean*. **OECD Publishing.**

Olowe, O., Kawalek, P., & Odusanya, K. (2023). Artificial intelligence adoption in criminal investigations: Challenges and opportunities for research. In *UKAIS 2023 Conference Proceedings*.

Pereira-Kohatsu JC, Quijano-Sánchez L, Liberatore F, Camacho-Collados M. (2019). Detecting and Monitoring Hate Speech in Twitter. Sensors (Basel, Switzerland). Oct;19(21):E4654. DOI: 10.3390/s19214654. PMID: 31717760; PMCID: PMC6864473

Poder Judicial de Costa Rica. (n.d.). Novedosa herramienta de inteligencia artificial se aplica en mejora de la protección de datos. [Comunicado de prensa].

Práxedes Martínez-Moreno, Andrea Valsecchi, Pablo Mesejo, Óscar Ibañez, Sergio Damas. (2024). Evidence evaluation in craniofacial superimposition using likelihood ratios. Information Fusion.

Quijano-Sánchez, L., Liberatore, F., Camacho-Collados, J., & Camacho-Collados, M. (2018). Aplicación automática de detección de lenguaje engañoso a informes policiales: extracción de patrones de comportamiento de un modelo de clasificación de varios pasos para comprender cómo mentimos a la policía. Knowledge-Based Systems , 149, 155-168

RIOS, Juan. Infobae. (2024, 29 de octubre). Cuidado en WhatsApp: copian la voz de tu mamá, usan IA para crear la estafa y roban dinero del banco.

Simmler, M., Brunner, S., Canova, G., & Schedler, K. (2023). Smart criminal justice: Exploring the use of algorithms in the Swiss criminal justice system. *Artificial Intelligence and Law, 31***(2), 213-237.**

Singh Sankhla, M., Kumar, R., & Jadhav, E. B. (2020). Artificial intelligence: Advancing automation in forensic science & criminal investigation. *Seybold Report, 15***(8).**

TechInformed. (2024, October 15). Deepfake cybercrime tool threatens crypto exchanges.

TRM. (2024, October 11). Ransomware in 2024: Latest trends, mounting threats, and the government response.

United Nations. (2024, March 11). *General Assembly Resolution 78/L.49: Harnessing opportunities of secure, protected and reliable AI systems for sustainable development***.**

UNESCO. (2019). *I'd Blush if I Could***.**

UNESCO. (2021). *Recommendation on the ethics of artificial intelligence* **(SHS/BIO/PI/2021/1).**

UNESCO. (2023). *Global AI toolkit on AI and rule of law for the judiciary***.**

UNESCO. (2024). *Artificial intelligence and gender equality: Key findings of UNESCO's Global Dialogue***.**

UNESCO. (2024). *Challenging systematic prejudices: An investigation into bias against women and girls in large language models***.**

UNICRI. (2024). *Generative AI: A new threat for online child sexual exploitation and abuse***.**

UNICRI & INTERPOL. (2024, February). *Responsible AI innovation in law enforcement: AI toolkit***.**

UNODC. (2024). *Casino underground banking report 2024***. UNODC Publications.**

Varma Microsoft, P. (n.d.). Transforming law enforcement policies and governance procedures: The benefits of AI integration.

VV.AA., HERRERA TRIGUERO, Francisco; PERALTA GUTIÉRREZ, Alfonso; TORRES LÓPEZ, Leopoldo Salvador. (2022). Capítulo "Uso policial de sistemas de inteligencia artificial en el ámbito comparado" pág. 453 y ss. El derecho y la inteligencia artificial. 2022. 24/10/2022. Editorial Universidad de Granada. 978-84-338-7049-0.

West, S. M., Whittaker, M., & Crawford, K. (2019). *Discriminating systems: Gender, race, and power in AI***. AI Now Institute.**

WIRED. (2024, October 15). Millions of people are using abusive AI 'Nudify' bots on Telegram.

World Wide Web Foundation. (2018, September). *Algoritmos e inteligencia artificial en Latinoamérica: Un estudio de implementaciones por parte de gobiernos en Argentina y Uruguay***.**

EL PACCTO 2.0

EU-LAC Partnership on justice and security

Financiado por
la Unión Europea