Cuadernos de RES PUBLICA en derecho y criminología
Nº 8 / 2026 ISSN: 2990-0697 DOI: 10.46661/respublica.13665.

Aplicación de la inteligencia artificial a la prevención, detección e investigación de la ciberdelincuencia

Application of artificial intelligence to the prevention, detection and investigation of cybercrime

María Luisa Romero Romero
Universidad Pablo de Olavide. Sevilla (España)
mlromrom@alu.upo.es
ORCID: 0009-0001-9883-0402

Resumen

Este artículo analiza la afectación de la inteligencia artificial (en adelante, IA) en la ciberseguridad y en la ciberdelincuencia desde una perspectiva jurídico-penal. A partir del examen del Convenio de Budapest, del Reglamento General de Protección de Datos (en adelante, RGPD), del Reglamento de Inteligencia Artificial de la Unión Europea (en adelante, AI Act), del Código Penal (en adelante, CP) y de la Ley de Enjuiciamiento Criminal (en adelante, LECrim), se concluye que la IA cumple una función ambivalente: refuerza tanto las capacidades ofensivas como las defensivas. Su incorporación exige, no obstante, un marco garantista basado en la proporcionalidad, la supervisión humana, la trazabilidad, la auditabilidad y el control judicial, especialmente por su impacto sobre colectivos vulnerables y sobre la confianza en la información digital.

Palabras clave: Inteligencia artificial; ciberseguridad; ciberdelincuencia; prueba digital; garantías procesales.

Abstract

This article examines the the impact of artificial intelligence (hereinafter, AI) on cybersecurity and cybercrime from a criminal-law perspective. Through the analysis of the Budapest Convention, the General Data Protection Regulation (hereinafter, GDPR), the Regulation on Artificial Intelligence of the European Union (hereinafter, AI Act), the Spanish Criminal Code (hereinafter, CC) and the Spanish Criminal Procedure Act (hereinafter, SCPA), it concludes that AI has a dual role, strengthening both offensive and defensive capacities. However, its use requires a guarantee-based framework grounded in proportionality, human oversight, traceability, auditability and judicial control, particularly in view of its impact on vulnerable groups and on trust in digital information.

Keywords: Artificial intelligence; cybersecurity; cybercrime; digital evidence; procedural safeguards.

Cómo citar este trabajo: Romero Romero María Luisa, (2026). Aplicación de la inteligencia artificial a la prevención, detección e investigación de la ciberdelincuencia. *Cuadernos de RES PUBLICA en derecho y criminología, (8), 01–22.* https://doi.org/10.46661/respublica.13665.

Recepción: 17.06.2026 **Aceptación:** 19.06.2026 **Publicación:** 23.06.2026

1 Introducción

La IA constituye una realidad tecnológica de difícil delimitación unívoca, en la medida en que su significado depende tanto del enfoque técnico desde el que se la examine como del contexto funcional en el que se despliegue. No se trata de una única tecnología, sino de un conjunto heterogéneo de sistemas y técnicas orientados a reproducir, asistir o automatizar operaciones que, tradicionalmente, se han vinculado a la inteligencia humana, tales como el aprendizaje, la inferencia, la predicción o la toma de decisiones.[1]

Desde una perspectiva clásica, Russell y Norvig han sistematizado la IA a partir de cuatro grandes modelos: sistemas que piensan como humanos, sistemas que actúan como humanos, sistemas que piensan racionalmente y sistemas que actúan racionalmente. Esta clasificación sigue siendo útil porque permite comprender que la IA no se define únicamente por su complejidad técnica, sino también por la finalidad perseguida y por el tipo de comportamiento que el sistema es capaz de desplegar.[2]

Ahora bien, en el ámbito jurídico no basta con una definición puramente técnica. Lo verdaderamente relevante es que estos sistemas pueden incidir sobre bienes jurídicos, derechos fundamentales y procesos de decisión con consecuencias jurídicas directas o indirectas. La IA adquiere así relevancia normativa no solo por lo que es, sino por lo que hace y por el tipo de riesgos que introduce en contextos socialmente sensibles.[3]

Además, la expansión de sistemas basados en IA ha intensificado la necesidad de distinguir entre usos de bajo impacto y usos que pueden generar consecuencias especialmente gravosas. La preocupación por la explicabilidad, la supervisión humana, la documentación del sistema y la gestión del riesgo no surge solo del progreso técnico, sino de la constatación de que estos instrumentos pueden condicionar el acceso a información, la identificación de amenazas, la clasificación de comportamientos y, en determinados contextos, la adopción de decisiones con relevancia jurídica.[4]

1.1. Ciberseguridad, ciberdelincuencia y ciberdelito: distinciones necesarias

Una adecuada aproximación jurídico-penal al fenómeno exige diferenciar con claridad tres nociones que con frecuencia se emplean de forma imprecisa: ciberseguridad, ciberdelincuencia y ciberdelito. Aunque se trata de conceptos conectados, no son equivalentes ni cumplen la misma función analítica.

La ciberseguridad puede entenderse como el conjunto de políticas, procedimientos, herramientas y medidas dirigidas a proteger redes, sistemas, dispositivos, servicios y datos frente a accesos no autorizados, alteraciones, usos indebidos, interrupciones o destrucciones ilegítimas. Se trata, por tanto, de una noción funcional y preventiva, vinculada a la protección de la confidencialidad, integridad y disponibilidad de los activos digitales.[5]

La ciberdelincuencia, en cambio, designa un fenómeno criminológico más amplio: engloba el conjunto de conductas ilícitas cometidas en

[1] LÓPEZ DE MÁNTARAS, R. y MESEGUER GONZÁLEZ, P.: *Inteligencia artificial*. Edit. Los Libros de la Catarata, Madrid, 2017, págs. 17-18.

[2] RUSSELL, S. y NORVIG, P.: *Inteligencia Artificial: un enfoque moderno*. 2.ª ed., Pearson Educación, S.A., Madrid, 2013, pág. 30.

[3] LLEDÓ BENITO, I.: *La robótica y la inteligencia artificial en la nueva era de la revolución industrial 4.0. Los desafíos jurídicos, éticos y tecnológicos de los robots inteligentes*. Edit. Dykinson, Madrid, 2021, pág. 81.

[4] Reglamento (UE) 2024/1689, de 13 de junio de 2024; ISO/IEC 42001:2023.

[5] NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY: *Framework for Improving Critical Infrastructure Cybersecurity: version 1.1*, 2018, págs. 4 y 6.

el ciberespacio o mediante tecnologías digitales, así como las dinámicas sociales, económicas y organizativas que rodean tales conductas. La ciberdelincuencia no se agota en la mera suma de tipos penales, sino que remite a una forma específica de criminalidad caracterizada por la deslocalización, la facilidad de escalado, la rapidez de ejecución, el anonimato relativo y la frecuente dimensión transnacional del ataque.[6]

Por su parte, el ciberdelito remite ya al plano dogmático-penal. Con esta expresión se alude a las conductas tipificadas penalmente en las que los sistemas informáticos, las redes de comunicación o los datos digitales constituyen bien el objeto del ataque, bien el medio comisivo empleado para lesionar otros bienes jurídicos.[7]

1.2. Ciberdelitos en sentido estricto y delitos comunes cometidos en entorno digital

Desde una perspectiva jurídico-penal, una de las clasificaciones más útiles consiste en distinguir entre ciberdelitos en sentido estricto y delitos comunes cometidos mediante tecnologías digitales. Esta distinción resulta preferible a las clasificaciones puramente descriptivas o casuísticas, porque permite ordenar las conductas atendiendo a la función que desempeña el elemento tecnológico en la estructura del injusto.

Los ciberdelitos en sentido estricto son aquellos en los que el sistema informático, la red, los datos o la infraestructura tecnológica constituyen el objeto directo del ataque. Aquí la lesión o puesta en peligro recae inmediatamente sobre la confidencialidad, integridad o disponibilidad de sistemas y datos, así como sobre el correcto funcionamiento de los entornos digitales. El Convenio de Budapest ofrece un marco particularmente útil para la sistematización del fenómeno, al identificar un núcleo de conductas que merecen una respuesta penal coordinada por parte de los Estados.[8]

Junto a ellos se encuentran los delitos comunes cometidos en entorno digital, esto es, aquellas infracciones que tutelan bienes jurídicos clásicos -patrimonio, intimidad, honor, libertad sexual, propiedad intelectual o indemnidad de menores, entre otros- pero que se ejecutan o amplifican mediante el uso de tecnologías de la información. En estos casos, lo digital no constituye necesariamente el objeto del ataque, pero sí el medio comisivo que facilita la conducta, multiplica su alcance o reduce los costes de ejecución.[9]

La utilidad de esta clasificación es doble. En primer lugar, permite comprender que no toda conducta cometida por internet pertenece al mismo bloque dogmático. En segundo término, muestra que la transformación digital no siempre exige la creación de nuevos tipos penales, pues en ocasiones basta con reinterpretar o adaptar figuras ya existentes a nuevas modalidades comisivas.[10]

1.3. Bienes jurídicos protegidos y problemas de política criminal

La expansión de la ciberdelincuencia y el impacto creciente de la IA obligan a examinar qué bienes jurídicos resultan comprometidos y si la respuesta penal vigente mantiene un equilibrio adecuado entre protección eficaz y respeto a los principios limitadores del ius puniendi.

[6] SUBIJANA ZUNZUNEGUI, I. J.: "El ciberterrorismo: Una perspectiva legal y judicial". En *Eguzkilore*, núm. 22, San Sebastián, 2008, págs. 169-187.

[7] Consejo de Europa: Convenio sobre la Ciberdelincuencia, hecho en Budapest el 23 de noviembre de 2001.

[8] Consejo de Europa: Convenio sobre la Ciberdelincuencia, cit., arts. 2 a 5.

[9] Consejo de Europa: Convenio sobre la Ciberdelincuencia, cit., arts. 7 a 9; LLEDÓ BENITO, I.: Op. Cit. (*La robótica y la inteligencia artificial en la nueva era de la revolución industrial 4.0. Los desafíos jurídicos, éticos y tecnológicos de los robots inteligentes*), pág. 81.

[10] SUBIJANA ZUNZUNEGUI, I. J.: Op. Cit. ("El ciberterrorismo: Una perspectiva legal y judicial"), págs. 169-187.

En los ciberdelitos en sentido estricto, los bienes jurídicos protegidos se proyectan principalmente sobre la confidencialidad, integridad y disponibilidad de los sistemas y datos, así como sobre la seguridad y fiabilidad del tráfico digital. En los delitos comunes cometidos en entorno tecnológico, en cambio, la afectación se desplaza hacia bienes más tradicionales, como el patrimonio, la intimidad, el honor, la indemnidad sexual o la libertad.[11]

A ello se añade una dificultad propia de la criminalidad digital: la misma herramienta tecnológica puede proyectarse simultáneamente sobre varios bienes jurídicos. Un ataque automatizado de phishing potenciado mediante IA, por ejemplo, puede comprometer al mismo tiempo el patrimonio de la víctima, la confidencialidad de sus datos, su identidad digital e incluso la seguridad de los sistemas a través de accesos ulteriores.

Desde la perspectiva de la política criminal, el principal desafío consiste en evitar dos riesgos contrapuestos: la infrarespuesta frente a formas de criminalidad cada vez más automatizadas, deslocalizadas y técnicamente sofisticadas, y la sobrerreacción punitiva, consistente en utilizar el Derecho penal como respuesta expansiva ante cualquier novedad tecnológica.[12]

2. Marco normativo aplicable

2.1. Convenio de Budapest y cooperación internacional

La ciberdelincuencia presenta una dimensión estructuralmente transnacional. A diferencia de otras formas de criminalidad, las conductas cometidas en el ciberespacio pueden ejecutarse desde un Estado, desplegar sus efectos en otro y afectar a víctimas, servidores o infraestructuras situadas en varias jurisdicciones de forma simultánea. Esta realidad hace insuficiente una respuesta puramente interna y explica la necesidad de instrumentos de coordinación internacional que permitan armonizar tipos penales y facilitar la cooperación entre autoridades.

En este marco, el Convenio sobre la Ciberdelincuencia, hecho en Budapest el 23 de noviembre de 2001, constituye el principal referente jurídico internacional. Su relevancia no se limita a la definición de un catálogo básico de conductas típicas, sino que reside también en haber establecido una lógica común para la lucha contra la criminalidad informática, basada en la compatibilidad mínima entre legislaciones nacionales y en la previsión de mecanismos de asistencia mutua.[13]

Desde el punto de vista material, el Convenio identifica un núcleo de comportamientos especialmente vinculados al entorno digital, entre ellos el acceso ilícito, la interceptación ilícita, la interferencia en datos, la interferencia en sistemas, la falsificación informática, el fraude informático y determinadas conductas relacionadas con contenidos ilícitos. Esta sistematización resulta especialmente útil para el análisis jurídico-penal porque permite diferenciar entre ataques dirigidos directamente contra sistemas o datos y delitos tradicionales cometidos a través de medios tecnológicos.[14]

Ahora bien, la importancia del Convenio de Budapest va más allá de la tipificación. Su

[11] Consejo de Europa: Convenio sobre la Ciberdelincuencia, cit.

[12] LLEDÓ BENITO, I.: Op. Cit. (*La robótica y la inteligencia artificial en la nueva era de la revolución industrial 4.0. Los desafíos jurídicos, éticos y tecnológicos de los robots inteligentes*), pág. 81; SUBIJANA ZUNZUNEGUI, I. J.: Op. Cit. ("El ciberterrorismo: Una perspectiva legal y judicial"), págs. 169-187.

[13] Consejo de Europa: Convenio sobre la Ciberdelincuencia, hecho en Budapest el 23 de noviembre de 2001.

[14] Consejo de Europa: Convenio sobre la Ciberdelincuencia, cit., arts. 2 a 10.

auténtico valor reside en haber asumido que la eficacia frente a la ciberdelincuencia depende de la cooperación internacional. La preservación rápida de datos, el intercambio de información, la asistencia judicial y la articulación de respuestas coordinadas son elementos esenciales cuando la prueba digital es volátil, fácilmente desplazable y frecuentemente alojada fuera del territorio nacional.

2.2. Derecho de la Unión Europea: protección de datos, ciberseguridad e inteligencia artificial

Si el Convenio de Budapest constituye el eje internacional de referencia, el Derecho de la Unión Europea representa hoy el espacio normativo más relevante para articular una respuesta equilibrada entre innovación tecnológica, seguridad y tutela de derechos fundamentales. En el ámbito que aquí interesa, ese marco europeo se proyecta principalmente sobre dos grandes pilares: la protección de datos personales y la regulación de los sistemas de IA.

El primero de ellos viene dado por el Reglamento (UE) 2016/679, Reglamento General de Protección de Datos (en adelante, RGPD). Su importancia para este trabajo es evidente, pues gran parte de las herramientas de IA aplicadas a la ciberseguridad operan mediante el tratamiento masivo de información, en ocasiones incluyendo datos personales o datos capaces de afectar significativamente a la esfera privada de los individuos.[15]

El segundo gran pilar normativo es el Reglamento (UE) 2024/1689, conocido como Reglamento de Inteligencia Artificial de la Unión Europea (en adelante, AI Act), que adopta un enfoque basado en el riesgo. Este modelo resulta especialmente pertinente porque permite ordenar jurídicamente los distintos usos de la IA en función de su peligrosidad potencial para la seguridad y los derechos fundamentales.[16]

Desde una perspectiva material, el AI Act introduce exigencias que tienen una incidencia directa en el ámbito de la ciberseguridad y de la eventual detección de conductas delictivas: gestión de riesgos, calidad de los datos, documentación técnica, trazabilidad mediante registros, transparencia suficiente para interpretar los resultados, supervisión humana y requisitos de precisión, solidez y ciberseguridad. Junto a estos dos reglamentos, resulta útil incorporar la ISO/IEC 42001:2023, no como norma penal ni como fuente directamente sancionadora, sino como estándar técnico-organizativo que refuerza la gobernanza de los sistemas de IA.[17] [18]

2.3. Derecho español: Código Penal, LECrim y normativa complementaria

En el ámbito interno, la respuesta jurídica frente a la ciberdelincuencia se articula fundamentalmente a través del Código Penal (en adelante, CP) y de la Ley de Enjuiciamiento Criminal (en adelante, LECrim). El primero proporciona la base sustantiva para la tipificación y sanción de las conductas delictivas, mientras que la segunda resulta esencial para determinar bajo qué condiciones pueden adoptarse diligencias de investigación tecnológica en el curso del proceso penal.[19]

Desde la perspectiva material, el CP no configura una categoría autónoma y cerrada de ciberdelito, sino una pluralidad de figuras dispersas que permiten sancionar tanto

[15] Parlamento Europeo y Consejo de la Unión Europea: Reglamento (UE) 2016/679, de 27 de abril de 2016.

[16] Parlamento Europeo y Consejo de la Unión Europea: Reglamento (UE) 2024/1689, de 13 de junio de 2024.

[17] ISO/IEC 42001:2023, *Information technology - Artificial intelligence - Management system*.

[18] Reglamento (UE) 2024/1689, cit., arts. 8 a 15.

[19] ESPAÑA: Ley Orgánica 10/1995, de 23 de noviembre, del Código Penal; ESPAÑA: Real Decreto de 14 de septiembre de 1882, por el que se aprueba la Ley de Enjuiciamiento Criminal.

ataques dirigidos contra sistemas y datos como delitos tradicionales cometidos mediante tecnologías de la información. Sin embargo, en un trabajo como el presente no basta con atender al plano sustantivo: la transformación tecnológica del delito obliga también a examinar cómo se obtienen, preservan y valoran las evidencias digitales y qué límites procesales rigen la actuación investigadora.

En este punto, la LECrim adquiere una importancia central. Tras la reforma operada por la Ley Orgánica 13/2015, la LECrim incorporó un régimen específico de medidas de investigación tecnológica, sometiéndolas a autorización judicial y a los principios de especialidad, idoneidad, excepcionalidad, necesidad y proporcionalidad. A ello se añaden exigencias concretas sobre la solicitud de la medida, el contenido del auto judicial, su duración, el control judicial de su ejecución y su cese.[20]

La regulación procesal española contempla, además, un abanico de diligencias particularmente relevantes para la investigación de la ciberdelincuencia: la interceptación de comunicaciones telefónicas y telemáticas, la captación y grabación de comunicaciones orales, la utilización de dispositivos de seguimiento, localización y captación de imagen, el registro de dispositivos de almacenamiento masivo de información y los registros remotos sobre equipos informáticos. Esta sistematización resulta de especial interés para este TFM, porque permite conectar el análisis de la IA con el problema jurídico de la obtención de información digital, la injerencia en derechos fundamentales y la exigencia de control judicial reforzado.[21]

Especialmente significativa es la regulación del registro de dispositivos de almacenamiento masivo de información, pues la LECrim exige una motivación individualizada para el acceso al contenido de ordenadores, teléfonos u otros soportes digitales, aclarando además que la mera incautación del dispositivo no legitima por sí sola el acceso a su contenido. Del mismo modo, en materia de registros remotos sobre equipos informáticos, la ley delimita los supuestos en que la medida puede acordarse, exige concreción sobre su alcance y prevé garantías relativas a la integridad, preservación y copia de los datos obtenidos.[22]

2.4. Tensiones entre seguridad, innovación tecnológica y derechos fundamentales

El marco normativo aplicable a la IA y a la ciberdelincuencia no se caracteriza únicamente por la acumulación de normas, sino por la existencia de una tensión constante entre tres exigencias distintas: reforzar la seguridad, permitir la innovación tecnológica y preservar los derechos fundamentales.

La primera de estas exigencias responde a una necesidad real. La ciberdelincuencia evoluciona con rapidez, opera a gran escala y aprovecha con eficacia los avances técnicos disponibles. En ese contexto, la utilización de sistemas automatizados para detectar anomalías, identificar patrones de ataque o reforzar la capacidad de respuesta puede resultar no solo útil, sino en muchos casos necesaria.

Sin embargo, la innovación tecnológica no puede convertirse en una cláusula general de legitimación. El hecho de que una herramienta sea técnicamente eficaz no significa que su utilización sea jurídicamente

[20] ESPAÑA: Ley de Enjuiciamiento Criminal, arts. 588 bis a y ss.; ESPAÑA: Ley Orgánica 13/2015, de 5 de octubre, de modificación de la Ley de Enjuiciamiento Criminal para el fortalecimiento de las garantías procesales y la regulación de las medidas de investigación tecnológica.

[21] ESPAÑA: Ley de Enjuiciamiento Criminal, arts. 588 bis b a 588 bis j; 588 sexies a y ss.; 588 septies a y ss.; 588 octies.

[22] ESPAÑA: Ley de Enjuiciamiento Criminal, arts. 588 sexies a, 588 sexies c, 588 septies a, 588 septies b y 588 septies c.

irreprochable. Precisamente por ello, tanto el RGPD como el AI Act y los marcos de gobernanza técnica insisten en principios de minimización, trazabilidad, transparencia, supervisión humana y control del riesgo.[23]

La verdadera cuestión jurídica no consiste, por tanto, en optar entre seguridad o derechos, sino en determinar qué condiciones normativas hacen posible una utilización legítima de la IA en contextos de ciberseguridad y persecución del delito. Esta será una de las ideas centrales del trabajo: la tecnología puede fortalecer la capacidad de respuesta frente a la ciberdelincuencia, pero solo si se integra en un modelo normativo compatible con la legalidad, la proporcionalidad y el control humano efectivo.[24]

En la misma línea, los debates europeos recientes sobre el uso de la inteligencia artificial por autoridades policiales y judiciales han insistido en que las aplicaciones incompatibles con la dignidad humana, la no discriminación, la presunción de inocencia o el derecho de defensa carecen de encaje en un Estado de Derecho. Ello refuerza la idea de que la gobernanza de la IA en materia penal no puede descansar en la sola promesa de eficiencia, sino en exigencias acumulativas de seguridad, transparencia, explicabilidad y control externo.

3. La inteligencia artificial como factor de transformación de la ciberdelincuencia

3.1. Automatización, escalabilidad y personalización del ataque

Uno de los efectos más relevantes de la IA sobre la ciberdelincuencia consiste en la transformación de la capacidad operativa del autor. La tecnología digital ya había permitido reducir costes, superar barreras territoriales y ampliar el número potencial de víctimas; sin embargo, la incorporación de sistemas de IA añade un elemento cualitativamente distinto: la posibilidad de automatizar tareas complejas, escalar campañas delictivas con gran rapidez y personalizar el ataque con un grado de precisión muy superior al de etapas anteriores.[25]

La automatización no debe entenderse únicamente como sustitución del trabajo humano, sino como multiplicación de la eficacia ofensiva. Allí donde antes era necesario un esfuerzo manual considerable para analizar objetivos, redactar mensajes verosímiles, detectar vulnerabilidades o adaptar un engaño a un perfil concreto, la IA permite ejecutar esas tareas de forma acelerada y, en determinados casos, con aprendizaje continuo. Ello favorece una ciberdelincuencia más eficiente, menos costosa y potencialmente más difícil de contener.[26]

A ello se une la escalabilidad. Los sistemas basados en IA permiten actuar sobre grandes volúmenes de datos, generar múltiples variantes de una misma campaña y adaptar los mensajes o vectores de ataque a distintos entornos sin necesidad de rediseñarlos completamente en cada ocasión. En consecuencia, el ciberataque deja de ser necesariamente una acción artesanal o limitada y pasa a concebirse, en muchos casos, como una actividad reproducible a gran escala.

Igualmente importante es la personalización del ataque. La IA facilita el tratamiento y

[23] Reglamento (UE) 2016/679, cit.; Reglamento (UE) 2024/1689, cit.; ISO/IEC 42001:2023, cit.

[24] LLEDÓ BENITO, I.: Op. Cit. (*La robótica y la inteligencia artificial en la nueva era de la revolución industrial 4.0. Los desafíos jurídicos, éticos y tecnológicos de los robots inteligentes*), pág. 81.

[25] FERRÉ, X.: "Cómo la inteligencia artificial está cambiando la ciberdelincuencia". En *EY*, 25 de mayo de 2023.

[26] CENTRO CRIPTOLÓGICO NACIONAL (CCN-CERT): *Aproximación a la Inteligencia Artificial y la ciberseguridad (BP/30)*, cit.; HITER, S.: "Generative AI and Cybersecurity". En *eWeek*, junio de 2023.

correlación de datos procedentes de múltiples fuentes, permitiendo construir mensajes, perfiles o simulaciones especialmente persuasivos. De este modo, ciertas formas de engaño dejan de apoyarse en comunicaciones burdas o genéricas y pasan a incorporar un alto grado de adecuación al destinatario concreto.[27]

3.2. IA generativa, phishing, deepfakes y fraude

La IA generativa constituye uno de los desarrollos más sensibles en esta materia, porque permite producir textos, imágenes, voces y vídeos con un elevado grado de realismo y a muy bajo coste. Su relevancia para la ciberdelincuencia es evidente: cuanto más verosímil sea la simulación, más eficaces pueden resultar la suplantación, el engaño, la manipulación reputacional o la captación fraudulenta de datos y recursos patrimoniales.[28]

En el ámbito del phishing, la IA permite generar correos electrónicos más convincentes, lingüísticamente correctos y adaptados al contexto de la víctima. La diferencia con modelos anteriores no reside solo en la mejora formal del mensaje, sino en la posibilidad de automatizar campañas masivas de engaño selectivo, ajustando contenido, tono y apariencia a perfiles concretos.[29]

Junto a ello, los deepfakes representan probablemente la manifestación más visible del impacto ofensivo de la IA generativa. La creación de imágenes, audios o vídeos sintéticos con apariencia de autenticidad agrava los riesgos para la intimidad, el honor, la propia imagen y, en determinados supuestos, también para el patrimonio o la seguridad colectiva. El problema no se limita a la falsedad del contenido, sino a su capacidad para generar efectos jurídicamente relevantes.[30]

3.3. Anonimato, atribución de autoría y sofisticación técnica

Otro de los grandes efectos de la IA sobre la ciberdelincuencia se proyecta sobre la atribución de autoría. La criminalidad digital ya presenta, por su propia naturaleza, problemas relevantes de identificación del responsable, debido al uso de intermediarios técnicos, servidores distribuidos, técnicas de ocultación, redes privadas virtuales y estructuras transnacionales. La IA no elimina estos problemas; por el contrario, puede intensificarlos al facilitar un mayor grado de automatización y de separación entre la decisión humana y la ejecución material del ataque.[31]

En la medida en que determinadas herramientas basadas en IA pueden generar contenido, interactuar con víctimas, adaptar mensajes o tomar decisiones operativas preconfiguradas, la intervención humana directa aparece más difuminada. Desde el punto de vista penal, ello no suprime la necesidad de una imputación personal del hecho, pero sí complica la reconstrucción de la cadena de decisión, la determinación del dominio funcional y la prueba del dolo o del conocimiento del resultado.

---

[27] FERRÉ, X.: Op. Cit. ("Cómo la inteligencia artificial está cambiando la ciberdelincuencia"); "FraudGPT, el método de ciberdelincuencia con inteligencia artificial". En *eldiario.com*, 24 de enero de 2024.

[28] WESTERLUND, M.: "The Emergence of Deepfake Technology: A Review". Vol. 9, noviembre de 2019, pág. 41.

[29] "FraudGPT, el método de ciberdelincuencia con inteligencia artificial", cit.; HITER, S.: Op. Cit. ("Generative AI and Cybersecurity").

[30] WESTERLUND, M.: Op. Cit. ("The Emergence of Deepfake Technology: A Review"), pág. 41; SIXMA VAN HEEMSTRA FERNÁNDEZ, A.: *La Desinformación como Amenaza para la Democracia: El Caso del Brexit*. Documento de Opinión IEEE 42/2023, pág. 4.

[31] SUBIJANA ZUNZUNEGUI, I. J.: Op. Cit. ("El ciberterrorismo: Una perspectiva legal y judicial"), págs. 169-187; Consejo de Europa: Convenio sobre la Ciberdelincuencia, cit.

A esta dificultad se suma la creciente sofisticación técnica de los ataques. La IA puede utilizarse para identificar vulnerabilidades, modificar patrones de actuación, sortear mecanismos tradicionales de detección o mejorar la capacidad de evasión frente a sistemas defensivos. Incluso los propios sistemas de IA pueden convertirse en objeto de ataque, mediante la manipulación intencional de entradas, la introducción de datos maliciosos en procesos de entrenamiento o la explotación de debilidades del modelo para obtener resultados erróneos o eludir controles.[32]

3.4. Incidencia en la expansión del riesgo penal

La IA no solo modifica técnicas concretas de ataque, sino que contribuye a una expansión general del riesgo penal en el entorno digital. Ello ocurre porque incrementa la capacidad ofensiva de los autores, reduce barreras de entrada, acelera la ejecución, amplía el número de potenciales víctimas y dificulta la reacción temprana de los mecanismos de control.[33]

Esta expansión del riesgo se manifiesta en varios planos. En primer lugar, en la mayor facilidad para lesionar simultáneamente diversos bienes jurídicos. En segundo término, en la aceleración temporal del daño. Finalmente, en la ampliación del impacto social, pues determinados usos maliciosos de la IA no afectan solo a víctimas individuales, sino también a la confianza general en la autenticidad de la información y en la seguridad del entorno digital.[34]

La gravedad del fenómeno exige una respuesta jurídica eficaz, pero ello no exime de mantener un enfoque garantista, especialmente cuando están en juego derechos fundamentales, exigencias de proporcionalidad y control sobre herramientas tecnológicas cada vez más complejas.[35]

4. Victimización digital y colectivos especialmente vulnerables

La afectación de la IA sobre la ciberdelincuencia no se agota en la sofisticación técnica del ataque ni en la ampliación de su escala. También obliga a analizar quiénes soportan con mayor intensidad sus efectos. La victimización digital no se distribuye de manera uniforme, sino que incide con especial fuerza sobre determinados colectivos cuya vulnerabilidad deriva de factores etarios, cognitivos, sociales o estructurales. En este sentido, la IA no crea por sí sola la vulnerabilidad, pero sí puede amplificarla, al hacer más persuasivos los engaños, más realistas las manipulaciones y más difícil la detección temprana del riesgo.[36]

Desde esta perspectiva, el análisis jurídico no puede limitarse a describir nuevas herramientas delictivas, sino que debe examinar cómo estas tecnologías intensifican situaciones preexistentes de exposición desigual. La ciberdelincuencia impulsada o facilitada por IA produce así una victimización

[32] CENTRO CRIPTOLÓGICO NACIONAL (CCN-CERT): *Aproximación a la Inteligencia Artificial y la ciberseguridad (BP/30)*, págs. 32-33; ISO/IEC 42001:2023; ISO/IEC 27001:2022.

[33] FERRÉ, X.: Op. Cit. ("Cómo la inteligencia artificial está cambiando la ciberdelincuencia"); HITER, S.: Op. Cit. ("Generative AI and Cybersecurity").

[34] SIXMA VAN HEEMSTRA FERNÁNDEZ, A.: Op. Cit. (*La Desinformación como Amenaza para la Democracia: El Caso del Brexit*), pág. 4; WESTERLUND, M.: Op. Cit. ("The Emergence of Deepfake Technology: A Review"), pág. 41.

[35] LLEDÓ BENITO, I.: Op. Cit. (*La robótica y la inteligencia artificial en la nueva era de la revolución industrial 4.0. Los desafíos jurídicos, éticos y tecnológicos de los robots inteligentes*), pág. 81; SUBIJANA ZUNZUNEGUI, I. J.: Op. Cit. ("El ciberterrorismo: Una perspectiva legal y judicial"), págs. 169-187.

[36] LLEDÓ BENITO, I.: Op. Cit. (*La robótica y la inteligencia artificial en la nueva era de la revolución industrial 4.0. Los desafíos jurídicos, éticos y tecnológicos de los robots inteligentes*), pág. 81.

diferenciada, en la que determinados grupos -menores, mujeres, personas mayores o incluso la colectividad cuando se ve afectada la esfera pública informativa- aparecen especialmente expuestos.[37]

4.1. Menores de edad

Los menores de edad constituyen uno de los colectivos más sensibles frente a la ciberdelincuencia en general y frente a sus manifestaciones potenciadas por IA en particular.

Su especial vulnerabilidad responde a varias razones concurrentes: una elevada exposición a entornos digitales, una percepción del riesgo todavía inmadura, una menor capacidad para detectar dinámicas manipulativas complejas y una construcción temprana de la identidad personal en redes y plataformas digitales.[38]

La gravedad del problema se acentúa si se tiene en cuenta que la explotación sexual infantil en entornos digitales ya presentaba, con anterioridad a la expansión de la IA generativa, una dimensión internacional preocupante. Los datos manejados por INTERPOL, ICSE y ECPAT ponen de relieve una presencia significativa de víctimas menores de edad en materiales de explotación sexual, así como una especial gravedad en determinados supuestos de victimización reflejados en esos contenidos.39

A ello se suma el impacto emocional propio de técnicas de clonación de voz o de simulación automatizada, que pueden utilizarse para generar situaciones ficticias de urgencia o peligro y provocar respuestas impulsivas en el entorno familiar del menor. La relevancia de estos supuestos no es solo patrimonial, sino también psicológica: el daño se produce mediante la explotación de vínculos afectivos y mediante la apariencia de autenticidad creada por la tecnología.[40]

4.2. Mujeres y violencia digital de género

Otro de los ámbitos en los que la IA ha mostrado una capacidad particularmente lesiva es el de la violencia digital contra las mujeres. Aquí la tecnología no introduce solo un nuevo medio de agresión, sino una intensificación de dinámicas ya conocidas de cosificación, humillación, hostigamiento y control, ahora amplificadas por la facilidad de creación y difusión de contenidos sintéticos.[41]

La manifestación más evidente de esta problemática es la producción de deepfakes de contenido sexual no consentido. La alteración de fotografías reales o la generación completa de imágenes íntimas falsas permite construir representaciones verosímiles que lesionan gravemente la dignidad, el honor y la propia imagen de la víctima, incluso cuando el contenido es completamente artificial.

Lo jurídicamente relevante no es la autenticidad del archivo, sino su capacidad para provocar descrédito, exposición pública, estigmatización y daño reputacional.[42]

En este contexto, la IA actúa como factor multiplicador de la violencia. Disminuye el coste técnico necesario para producir el ataque, acelera la circulación del contenido y amplía exponencialmente el número de receptores potenciales. La lesión deja de depender de una conducta aislada y pasa a

---

[37] SIXMA VAN HEEMSTRA FERNÁNDEZ, A.: Op. Cit. (*La Desinformación como Amenaza para la Democracia: El Caso del Brexit*), pág. 4.

[38] SAVE THE CHILDREN: *Violencia viral. Análisis de la violencia contra la infancia y la adolescencia en el entorno digital*, 2019, págs. 17 y 32.

[39] ICSE - INTERPOL - ECPAT INTERNATIONAL: *Towards a Global Indicator on Unidentified Victims in Child Sexual Exploitation Material*, 2018, págs. 53 y 58.

[40] AUFERIL, B.: "Auge de la ciberdelincuencia impulsada por inteligencia artificial". En *LinkedIn*, 2024.

[41] LLEDÓ BENITO, I.: Op. Cit. (*La robótica y la inteligencia artificial en la nueva era de la revolución industrial 4.0. Los desafíos jurídicos, éticos y tecnológicos de los robots inteligentes*), pág. 81.

[42] WESTERLUND, M.: Op. Cit. ("The Emergence of Deepfake Technology: A Review"), pág. 41.

proyectarse en un ecosistema digital donde la difusión, replicación y permanencia del contenido aumentan la intensidad del daño.[43]

4.3. Personas mayores y fraudes personalizados

Aunque la victimización general por ciberdelitos no se concentra estadísticamente en las franjas de mayor edad, este colectivo presenta una especial exposición cualitativa frente a fraudes personalizados, engaños basados en la urgencia y técnicas de clonación de voz o simulación afectiva.44

La IA intensifica este riesgo mediante la personalización del fraude. La utilización de datos extraídos de redes sociales, filtraciones previas o información accesible en línea permite construir mensajes o escenarios altamente verosímiles. A ello se añade la clonación de voz, técnica particularmente idónea para simular llamadas de familiares en apuros o situaciones ficticias que exigen una transferencia inmediata de dinero.[45]

4.4. Desinformación, opinión pública y afectación colectiva

La victimización digital no siempre recae únicamente sobre sujetos individualmente identificables. En determinados supuestos, la IA puede generar una afectación de alcance colectivo, al alterar la confianza social en la autenticidad de la información y favorecer dinámicas de desinformación masiva. Desde esta perspectiva, la víctima deja de ser solo la persona concreta engañada y pasa a ser también la opinión pública, en la medida en que se deterioran las condiciones mínimas para una comunicación veraz y para una deliberación democrática no manipulada.[46]

La IA generativa facilita la producción rápida y económica de imágenes, vídeos, voces y textos con apariencia de autenticidad. En combinación con la lógica viral de las plataformas digitales, ello permite multiplicar el impacto de narrativas engañosas, reforzar campañas de manipulación y dificultar la distinción entre contenido real y contenido fabricado.[47]

Desde una perspectiva político-criminal, esta afectación colectiva exige, sin embargo, una respuesta cautelosa. Aunque la desinformación potenciada por IA pueda comprometer el debate público y erosionar la confianza en la información digital, ello no justifica sin más una criminalización expansiva de la mentira ni la instauración de filtros estatales de verdad.

El reto consiste en compatibilizar la reacción frente a los supuestos ya reconducibles a tipos penales vigentes con medidas de alfabetización mediática, verificación transparente y formación crítica de los usuarios, evitando que la lucha contra los bulos derive en censura o en restricciones desproporcionadas de la libertad de expresión e información.

4.5. La vulnerabilidad como categoría criminológica y jurídica

A la vista de lo expuesto, la vulnerabilidad no puede ser entendida como una condición fija o puramente subjetiva, sino como una categoría dinámica que resulta de la

[43] WESTERLUND, M.: Op. Cit. ("The Emergence of Deepfake Technology: A Review"), pág. 41; LLEDÓ BENITO, I.: Op. Cit. (*La robótica y la inteligencia artificial en la nueva era de la revolución industrial 4.0. Los desafíos jurídicos, éticos y tecnológicos de los robots inteligentes*), pág. 81.

[44] MINISTERIO DEL INTERIOR: *Informe sobre la cibercriminalidad en España 2022*. Gobierno de España, 2023; AUFERIL, B.: Op. Cit. ("Auge de la ciberdelincuencia impulsada por inteligencia artificial").

[45] AUFERIL, B.: Op. Cit. ("Auge de la ciberdelincuencia impulsada por inteligencia artificial").

[46] SIXMA VAN HEEMSTRA FERNÁNDEZ, A.: Op. Cit. (*La Desinformación como Amenaza para la Democracia: El Caso del Brexit*), pág. 4.

[47] WESTERLUND, M.: Op. Cit. ("The Emergence of Deepfake Technology: A Review"), pág. 41; SIXMA VAN HEEMSTRA FERNÁNDEZ, A.: Op. Cit. (*La Desinformación como Amenaza para la Democracia: El Caso del Brexit*), pág. 4.

interacción entre características del sujeto, entorno tecnológico y estrategias del autor. La IA no determina automáticamente quién será víctima, pero sí modifica las condiciones del riesgo y amplifica determinadas asimetrías preexistentes.[48]

En menores, la vulnerabilidad se conecta con la madurez emocional y la exposición digital; en mujeres, con dinámicas estructurales de violencia y cosificación que encuentran nuevas formas de expresión en la tecnología; en personas mayores, con la brecha tecnológica y la explotación de la confianza; y en la esfera colectiva, con la facilidad para manipular la información pública. La clave común a todos estos supuestos es que la IA aumenta la capacidad de adaptación del ataque a las debilidades específicas del destinatario.[49]

Desde una perspectiva jurídico-penal, esta constatación resulta importante por dos razones.

En primer lugar, porque obliga a incorporar una lectura más sensible a la posición real de la víctima en el entorno digital.

En segundo lugar, porque pone de relieve que la respuesta frente a la ciberdelincuencia no puede descansar únicamente en la sanción ex post, sino que debe integrar estrategias preventivas, educativas y de reducción del riesgo orientadas a los colectivos más expuestos.[50]

5. Inteligencia artificial aplicada a la prevención, detección e investigación de ciberdelitos

La IA no solo actúa como factor de sofisticación de la ciberdelincuencia, sino también como instrumento de defensa frente a ella. Esta doble dimensión obliga a analizar su utilización en tareas de prevención, detección e investigación desde una perspectiva distinta a la empleada en capítulos anteriores. Ya no se trata aquí de estudiar cómo la IA fortalece la capacidad ofensiva del delincuente, sino de examinar en qué medida puede contribuir a la protección de sistemas, datos y bienes jurídicos sin comprometer las garantías propias del Estado de Derecho.

En este ámbito, la principal aportación de la IA radica en su capacidad para procesar grandes volúmenes de información en tiempo reducido, identificar patrones anómalos, correlacionar eventos dispersos y mejorar la capacidad de anticipación y respuesta frente a incidentes de seguridad. Esta funcionalidad resulta especialmente valiosa en entornos digitales donde el número de señales es inasumible mediante revisión humana exclusiva y donde la velocidad del ataque exige mecanismos de reacción mucho más ágiles que los disponibles en modelos tradicionales.[51]

La incorporación de IA a la ciberseguridad introduce, por tanto, un cambio de paradigma: de una lógica predominantemente reactiva se pasa a una

[48] LLEDÓ BENITO, I.: Op. Cit. (*La robótica y la inteligencia artificial en la nueva era de la revolución industrial 4.0. Los desafíos jurídicos, éticos y tecnológicos de los robots inteligentes*), pág. 81.

[49] SAVE THE CHILDREN: Op. Cit. (*Violencia viral. Análisis de la violencia contra la infancia y la adolescencia en el entorno digital*), págs. 17 y 32; LLEDÓ BENITO, I.: Op. Cit. (*La robótica y la inteligencia artificial en la nueva era de la revolución industrial 4.0. Los desafíos jurídicos, éticos y tecnológicos de los robots inteligentes*), pág. 81; AUFERIL, B.: Op. Cit. ("Auge de la ciberdelincuencia impulsada por inteligencia artificial"); SIXMA VAN HEEMSTRA FERNÁNDEZ, A.: Op. Cit. (*La Desinformación como Amenaza para la Democracia: El Caso del Brexit*), pág. 4.

[50] SAVE THE CHILDREN: Op. Cit. (*Violencia viral. Análisis de la violencia contra la infancia y la adolescencia en el entorno digital*), pág. 73.

[51] NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY: *Framework for Improving Critical Infrastructure Cybersecurity: version 1.1*, 2018, págs. 4 y 6; GALÁN CORDERO, C. y GALÁN PASCUAL, C.: *Aproximación a la Inteligencia Artificial y la ciberseguridad (BP/30)*, CCN-CERT, 2023, págs. 32-33.

lógica de prevención avanzada y detección temprana. Sin embargo, esa evolución no es jurídicamente neutra. La posibilidad de clasificar riesgos, priorizar amenazas o generar alertas automatizadas plantea interrogantes sobre el tratamiento de datos, la opacidad del sistema, la supervisión humana y la licitud de utilizar esos resultados en contextos con eventual trascendencia penal.

5.1. Usos defensivos de la IA en ciberseguridad

Los sistemas de IA aplicados a la ciberseguridad presentan una utilidad defensiva clara cuando se emplean para mejorar la visibilidad sobre el entorno digital y reforzar la capacidad de respuesta ante comportamientos anómalos. Frente a los sistemas tradicionales basados únicamente en reglas o firmas conocidas, los modelos de aprendizaje automático permiten detectar patrones de comportamiento que pueden revelar amenazas desconocidas o variantes nuevas de ataques ya conocidos.[52]

Entre las aplicaciones defensivas más relevantes destacan, en primer lugar, la detección de amenazas y el análisis de comportamiento, que permiten identificar accesos inusuales, desviaciones en tráfico de red, movimientos laterales o patrones de uso incompatibles con el funcionamiento ordinario del sistema. En segundo término, la respuesta automática y orquestación, mediante la cual determinadas acciones de contención o mitigación pueden activarse con una intervención humana reducida o posterior. En tercer lugar, la predicción de amenazas, entendida como la capacidad de reconocer señales previas o regularidades que permitan anticipar riesgos de seguridad antes de que el daño se materialice plenamente.[53]

Asimismo, la IA puede aplicarse al análisis de vulnerabilidades, al pentesting automatizado, a la detección de campañas de phishing, a la clasificación de software malicioso y a la simulación de escenarios de ataque destinados a reforzar la robustez de la infraestructura. Estas funciones muestran que la IA, bien utilizada, no constituye únicamente una herramienta de reacción, sino también de preparación, ensayo y mejora de la resiliencia frente a amenazas futuras.[54]

5.2. Aplicación práctica en el sector público y privado

La prevención y detección de riesgos en el entorno digital no se desarrolla exclusivamente dentro de la esfera estatal. Por el contrario, se articula en un modelo mixto, en el que intervienen fuerzas y cuerpos de seguridad, centros nacionales de ciberseguridad, operadores esenciales, entidades privadas y proveedores tecnológicos. Esta realidad hace que la utilización de IA en ciberseguridad se despliegue en espacios institucionales muy diversos y obliga a entender la defensa digital como una función compartida.[55]

Desde una perspectiva funcional, los sistemas de IA se integran hoy en arquitecturas orientadas a correlacionar eventos, priorizar alertas, detectar incidentes en tiempo real y

[52] GALÁN CORDERO, C. y GALÁN PASCUAL, C.: Op. Cit. (*Aproximación a la Inteligencia Artificial y la ciberseguridad (BP/30)*); RUSSELL, S. y NORVIG, P.: Op. Cit. (*Inteligencia Artificial: un enfoque moderno*), pág. 30.

[53] GALÁN CORDERO, C. y GALÁN PASCUAL, C.: Op. Cit. (*Aproximación a la Inteligencia Artificial y la ciberseguridad (BP/30)*), págs. 32-35; NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY: Op. Cit. (*Framework for Improving Critical Infrastructure Cybersecurity: version 1.1*), págs. 4 y 6.

[54] INSTITUTO NACIONAL DE CIBERSEGURIDAD (INCIBE): "Pentesting"; GALÁN CORDERO, C. y GALÁN PASCUAL, C.: Op. Cit. (*Aproximación a la Inteligencia Artificial y la ciberseguridad (BP/30)*), pág. 40.

[55] NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY: Op. Cit. (*Framework for Improving Critical Infrastructure Cybersecurity: version 1.1*), págs. 4 y 6.

reducir el tiempo de respuesta. El interés jurídico de esta aplicación práctica no reside tanto en el detalle técnico de la herramienta como en el hecho de que la IA pasa a ocupar una posición relevante en la gobernanza de la seguridad digital. Cuanto mayor sea la dependencia de estos sistemas, mayor será también la necesidad de asegurar que su funcionamiento sea verificable, auditable y compatible con exigencias normativas claras.[56]

Esta cuestión resulta especialmente importante en el ámbito privado. Las entidades financieras, los prestadores de servicios digitales o las empresas que gestionan infraestructuras críticas pueden detectar tempranamente incidentes o comportamientos anómalos y generar información de gran utilidad para la prevención del delito. Sin embargo, esa utilidad no puede traducirse en una delegación indiscriminada de funciones cuasi-investigadoras ni en una circulación descontrolada de datos o inferencias algorítmicas.

La experiencia comparada y española muestra, además, que estas utilidades no son meramente hipotéticas. En el espacio penal ya se han ensayado herramientas algorítmicas de apoyo como VERIPOL para el análisis lingüístico de denuncias potencialmente falsas, CATT para el análisis de chats en contextos de explotación de menores, o sistemas de valoración de riesgo como VioGén y RisCanvi. Aunque muchos de estos instrumentos no fueron diseñados específicamente para la ciberdelincuencia, ilustran bien una idea central: la IA puede servir para priorizar, clasificar y orientar la investigación, pero sus resultados deben concebirse como apoyo técnico y nunca como sustituto de la decisión policial, fiscal o jurisdiccional.

5.3. IA y prevención del delito: límites del paradigma predictivo

Uno de los puntos más delicados de este debate aparece cuando la prevención técnica se aproxima al terreno del análisis predictivo del delito. La IA puede identificar regularidades, patrones de riesgo y combinaciones de variables que, desde el punto de vista técnico, resultan útiles para anticipar posibles incidentes de seguridad. Sin embargo, el paso de esa lógica preventiva al ámbito penal exige una cautela reforzada.[57]

La razón es clara: una correlación estadística o una predicción automatizada no equivalen, por sí solas, a una base legítima de imputación ni a un fundamento suficiente para restringir derechos. La prevención no puede confundirse con sospecha penal bastante, ni la detección automatizada puede sustituir la valoración individualizada exigida por los principios del proceso penal. La presunción de inocencia, la exigencia de prueba suficiente y la proporcionalidad impiden que un sistema algorítmico se convierta por sí mismo en criterio autónomo de responsabilidad o en justificación bastante para decisiones limitativas de derechos.[58]

Además, los modelos predictivos no están exentos de problemas estructurales. Pueden reproducir sesgos preexistentes en los datos de entrenamiento, operar con criterios difícilmente explicables o generar resultados que, aun siendo funcionalmente útiles, no resulten fácilmente traducibles a categorías jurídicas comprensibles. De ahí que la utilización de IA con fines preventivos solo pueda considerarse legítima si se inserta en un marco de supervisión humana, trazabilidad

[56] GALÁN CORDERO, C. y GALÁN PASCUAL, C.: Op. Cit. (*Aproximación a la Inteligencia Artificial y la ciberseguridad (BP/30)*); ISO/IEC 42001:2023; ISO/IEC 27001:2022.

[57] LLEDÓ BENITO, I.: Op. Cit. (*La robótica y la inteligencia artificial en la nueva era de la revolución industrial 4.0. Los desafíos jurídicos, éticos y tecnológicos de los robots inteligentes*), pág. 81.

[58] Reglamento (UE) 2016/679, cit.; Reglamento (UE) 2024/1689, cit.

suficiente, documentación del sistema y control del riesgo, evitando cualquier deriva hacia un modelo de derecho penal predictivo incompatible con el Estado de Derecho.[59]

Precisamente por ello, los sistemas predictivos empleados en ámbitos próximos al proceso penal han sido recibidos con cautela por la jurisprudencia. Tanto en materia de valoración del riesgo de violencia de género como en el ámbito penitenciario, los resultados algorítmicos se han entendido como aproximativos u orientativos, insuficientes por sí solos para fundamentar decisiones restrictivas sin contraste adicional. La lección trasladable al ámbito de la ciberdelincuencia es clara: cualquier modelo predictivo debe quedar sometido a verificación humana, motivación reforzada y control judicial efectivo, especialmente cuando opera con cuestionarios, indicadores de riesgo o inferencias probabilísticas.

5.4. Ventajas operativas y riesgos de dependencia tecnológica

La utilización de IA en ciberseguridad ofrece ventajas operativas innegables. Permite actuar sobre volúmenes masivos de datos, detectar incidentes con mayor rapidez, reforzar la defensa frente a ataques desconocidos y mejorar la capacidad de adaptación frente a amenazas dinámicas. En entornos complejos, estas capacidades pueden resultar decisivas para reducir tiempos de exposición, contener daños y mejorar la resiliencia de las organizaciones.[60]

Sin embargo, esas ventajas no pueden ocultar los riesgos de una dependencia tecnológica excesiva. Cuanto más se confíe en sistemas automatizados para clasificar amenazas, priorizar alertas o sugerir respuestas, más importante será examinar la calidad de los datos utilizados, la robustez del modelo, la existencia de mecanismos de auditoría y la posibilidad real de comprender el sentido de sus resultados.

Un sistema técnicamente avanzado pero opaco o insuficientemente controlado puede generar una apariencia de seguridad superior a la seguridad efectivamente alcanzada.[61]

En definitiva, la IA puede reforzar de forma muy significativa la prevención, detección e investigación de ciberdelitos, pero solo será jurídicamente admisible si sus ventajas operativas quedan acompañadas de supervisión humana significativa, gobernanza robusta, trazabilidad, auditoría y respeto a los derechos fundamentales. Precisamente sobre estas exigencias gira el capítulo siguiente.[62]

6. Límites jurídicos y garantías en el uso de inteligencia artificial frente a la ciberdelincuencia

La utilización de IA en tareas de prevención, detección e investigación de ciberdelitos no puede valorarse exclusivamente desde la óptica de la eficacia técnica. En un Estado de Derecho, cualquier herramienta empleada con incidencia potencial sobre personas concretas, sobre sus datos o sobre sus posiciones procesales debe someterse a límites materiales y procedimentales estrictos.

En consecuencia, la cuestión central no consiste en determinar únicamente si la IA funciona, sino si su utilización resulta jurídicamente legítima, controlable y

[59] ISO/IEC 42001:2023, cit.; Reglamento (UE) 2024/1689, cit.; LLEDÓ BENITO, I.: Op. Cit. (*La robótica y la inteligencia artificial en la nueva era de la revolución industrial 4.0. Los desafíos jurídicos, éticos y tecnológicos de los robots inteligentes*), pág. 81.

[60] GALÁN CORDERO, C. y GALÁN PASCUAL, C.: Op. Cit. (*Aproximación a la Inteligencia Artificial y la ciberseguridad (BP/30)*).

[61] ISO/IEC 42001:2023, cit.; ISO/IEC 27001:2022, cit.; LLEDÓ BENITO, I.: Op. Cit. (*La robótica y la inteligencia artificial en la nueva era de la revolución industrial 4.0. Los desafíos jurídicos, éticos y tecnológicos de los robots inteligentes*), pág. 81.

[62] Reglamento (UE) 2016/679, cit.; Reglamento (UE) 2024/1689, cit.; ISO/IEC 42001:2023, cit.; ISO/IEC 27001:2022, cit.

compatible con los derechos fundamentales.[63]

Este problema es especialmente intenso en el ámbito de la ciberdelincuencia. La investigación digital implica con frecuencia acceso a dispositivos, tratamiento masivo de información, conservación de datos, análisis automatizado de patrones y utilización de sistemas opacos o complejos. Cuando a ello se añade la IA, el riesgo de afectación sobre la intimidad, el secreto de las comunicaciones, la protección de datos, la presunción de inocencia y el derecho de defensa se incrementa de forma notable. De ahí que la legitimidad de estas tecnologías dependa de su inserción en un marco garantista suficientemente robusto.

6.1. Protección de datos, intimidad y tratamiento masivo de información

Uno de los primeros límites jurídicos aparece en el terreno de la protección de datos personales. Buena parte de los sistemas de IA aplicados a la ciberseguridad trabajan sobre registros, metadatos, patrones de comportamiento, credenciales, comunicaciones o trazas de actividad que pueden contener información personal o hacer identificable a una persona física. En este contexto, el RGPD se convierte en una referencia imprescindible, en la medida en que exige licitud del tratamiento, limitación de la finalidad, minimización de datos, integridad, confidencialidad y responsabilidad proactiva.[64]

Estas exigencias no desaparecen por el hecho de que la finalidad perseguida sea la seguridad. Al contrario: cuando se pretende implantar sistemas automatizados de vigilancia, detección o clasificación de riesgos, la necesidad de justificar la base jurídica del tratamiento y de limitarlo a lo estrictamente necesario se vuelve aún más intensa. Especial importancia adquiere aquí el artículo 22 RGPD, en la medida en que protege frente a determinadas decisiones automatizadas con efectos significativos, y el artículo 35, relativo a la evaluación de impacto.

La afectación no se limita, además, al plano de los datos personales. En el ámbito procesal penal, el uso de medidas de investigación tecnológica puede incidir directamente sobre la intimidad, el secreto de las comunicaciones y la protección de datos personales, tal como reconoce expresamente la propia Ley Orgánica 13/2015, que reformó la LECrim para regular estas medidas y fortalecer las garantías procesales.[65]

6.2. Sesgos, opacidad y explicabilidad algorítmica

Un segundo límite jurídico deriva de la opacidad de determinados sistemas de IA y de la posibilidad de que reproduzcan o intensifiquen sesgos presentes en los datos de entrenamiento o en la propia lógica del modelo. Esta cuestión resulta especialmente sensible cuando el sistema no se limita a detectar una anomalía técnica, sino que produce inferencias sobre conductas, riesgos o perfiles con posible trascendencia jurídica.[66]

La opacidad algorítmica plantea un problema estructural para el Derecho. Si el funcionamiento del sistema no puede ser comprendido de manera suficiente, resulta muy difícil controlar la racionalidad de sus resultados, auditar sus errores o someter sus inferencias a contradicción. En el ámbito penal, donde la prueba debe ser inteligible y susceptible de discusión por las partes, esta dificultad no es menor: una herramienta incomprensible puede debilitar gravemente la

[63] LLEDÓ BENITO, I.: Op. Cit. (*La robótica y la inteligencia artificial en la nueva era de la revolución industrial 4.0. Los desafíos jurídicos, éticos y tecnológicos de los robots inteligentes*), pág. 81.

[64] Reglamento (UE) 2016/679, cit., arts. 5, 22, 32 y 35.

[65] ESPAÑA: Ley Orgánica 13/2015, de 5 de octubre; Preámbulo. BOE.

[66] LLEDÓ BENITO, I.: Op. Cit. (*La robótica y la inteligencia artificial en la nueva era de la revolución industrial 4.0. Los desafíos jurídicos, éticos y tecnológicos de los robots inteligentes*), pág. 81.

posibilidad de control judicial y el ejercicio efectivo del derecho de defensa.

De forma paralela, los sesgos constituyen una amenaza relevante para la igualdad y la no discriminación. Cuando un modelo aprende de datos históricos marcados por asimetrías previas, puede reproducirlas al clasificar riesgos o priorizar alertas. Por ello, tanto el AI Act como los marcos de gobernanza técnica insisten en la calidad de los datos, la gestión del riesgo, la monitorización del funcionamiento y la necesidad de mecanismos de revisión y corrección.[67]

6.3. Supervisión humana, proporcionalidad y control judicial

La idea de supervisión humana constituye uno de los ejes centrales de la regulación contemporánea de la IA. En realidad, esta exigencia responde a una intuición jurídica básica: cuando una herramienta tecnológica puede afectar a derechos fundamentales o influir en decisiones con relevancia penal, no puede quedar fuera del control humano significativo.[68]

En el plano procesal penal español, la lógica es análoga. La LECrim, tras la reforma introducida por la LO 13/2015, somete las medidas de investigación tecnológica a los principios de especialidad, idoneidad, excepcionalidad, necesidad y proporcionalidad. El artículo 588 bis a) exige, además, que no se utilicen para prevenir o descubrir delitos o despejar sospechas sin base objetiva, y vincula la proporcionalidad a la ponderación entre el sacrificio de derechos y el beneficio perseguido para el interés público.[69]

Esta previsión es de enorme importancia para el objeto de este trabajo, pues impide que la utilización de IA en la investigación se apoye en lógicas genéricas de vigilancia o en sospechas abiertas no individualizadas. Si una medida tecnológica limita derechos, debe estar conectada con un hecho concreto, ser útil para su esclarecimiento, no poder sustituirse por otra menos lesiva y mantenerse dentro de una ponderación razonable.

6.4. Prueba digital, derecho de defensa y validez procesal

El uso de IA en la investigación de ciberdelitos plantea también interrogantes muy relevantes en materia de prueba digital. No basta con que la tecnología permita detectar o sugerir una determinada hipótesis; es preciso que la obtención, conservación e incorporación de la información al proceso penal respeten las garantías legales y constitucionales exigibles.

Aquí adquiere especial relevancia la regulación de la LECrim sobre medidas de investigación tecnológica. El artículo 588 sexies a exige motivación individualizada cuando el acceso a ordenadores, dispositivos electrónicos o repositorios telemáticos de datos resulte previsible, y añade que la simple incautación del dispositivo no legitima por sí sola el acceso a su contenido.

Asimismo, el régimen de registros remotos del artículo 588 septies a delimita los supuestos en que la medida puede acordarse y exige concreción sobre el objeto, alcance y modo de ejecución. Por su parte, el artículo 588 octies regula la orden de conservación de datos para preservar información hasta obtener la autorización judicial correspondiente.[70]

Estas previsiones muestran que la tecnología no queda al margen del proceso, sino que se integra en él bajo condiciones estrictas. Desde el punto de vista del derecho de defensa, ello es esencial. Una persona investigada debe poder discutir no solo el resultado

[67] Reglamento (UE) 2024/1689, cit.; ISO/IEC 42001:2023, cit.

[68] Ibídem.

[69] ESPAÑA: Ley de Enjuiciamiento Criminal, art. 588 bis a. BOE.

[70] ESPAÑA: Ley de Enjuiciamiento Criminal, arts. 588 sexies a, 588 septies a y 588 octies. BOE.

incriminatorio, sino también el modo en que la información fue obtenida, la fiabilidad del sistema utilizado y la corrección del razonamiento que conecta ese dato con la imputación concreta.

No menos importante es evitar un desplazamiento retórico del debate probatorio. La mera invocación abstracta de que una captura de pantalla, un audio, un vídeo o una imagen "podrían haber sido generados o alterados por IA" no basta, por sí sola, para neutralizar su valor procesal. En un entorno en el que la manipulación digital es técnicamente posible, la impugnación debe concretar las razones de la sospecha y, cuando proceda, apoyarse en una pericial informática capaz de examinar autenticidad, integridad, metadatos, cadena de custodia y eventuales indicios de edición.

De otro modo, la referencia genérica al deepfake corre el riesgo de convertirse en una objeción especulativa incompatible con una valoración racional de la prueba.

Al mismo tiempo, tampoco puede presumirse una autenticidad inatacable de toda evidencia digital. Las deficiencias en la cadena de custodia, la falta de claridad sobre el origen del archivo o la inexistencia de verificación técnica no determinan automáticamente la nulidad, pero sí pueden debilitar su fuerza acreditativa cuando generan dudas fundadas sobre la integridad del soporte.

De ahí que la práctica de prueba pericial informática, la posibilidad de contradicción y la libre valoración judicial resulten especialmente decisivas en causas donde intervienen contenidos audiovisuales, documentos electrónicos o datos obtenidos mediante sistemas algorítmicos.

6.5. Hacia un modelo garantista de utilización de IA en seguridad y justicia penal

De todo lo anterior se desprende que la utilización de IA frente a la ciberdelincuencia solo puede considerarse jurídicamente legítima si se integra en un modelo garantista. Ese modelo exige, en primer lugar, una delimitación clara entre funciones auxiliares de apoyo analítico y decisiones con relevancia jurídica directa. La IA puede servir para detectar patrones, priorizar alertas o sugerir líneas de investigación; no puede, sin embargo, desplazar la valoración individualizada ni sustituir el razonamiento humano allí donde están en juego derechos fundamentales o consecuencias penales.

En segundo término, dicho modelo exige gobernanza robusta: documentación técnica, registros, trazabilidad, evaluación de riesgos, auditoría y supervisión humana. En esto convergen de manera muy clara el AI Act, el RGPD y la ISO/IEC 42001, al exigir transparencia, explicabilidad, control, gestión del riesgo y revisión continua.[71]

En tercer lugar, un modelo garantista requiere mantener el papel central del juez y del control judicial cuando la tecnología se proyecta sobre derechos fundamentales. La LO 13/2015 y la regulación de la LECrim evidencian precisamente esa lógica: las medidas tecnológicas más invasivas no se abandonan a la pura decisión policial o técnica, sino que se someten a autorización, motivación, duración limitada, control y cese.[72]

Por último, este modelo exige comprender que la lucha contra la ciberdelincuencia no puede descansar únicamente en la expansión punitiva ni en la sofisticación técnica de la vigilancia. Debe completarse con prevención, educación digital, cooperación institucional e instrumentos de control sobre la propia tecnología. Solo así la IA podrá contribuir a la seguridad sin erosionar las bases garantistas del sistema penal.

[71] Reglamento (UE) 2016/679, cit.; Reglamento (UE) 2024/1689, cit.; ISO/IEC 42001:2023, cit.

[72] ESPAÑA: Ley Orgánica 13/2015, de 5 de octubre; ESPAÑA: Ley de Enjuiciamiento Criminal, arts. 588 bis a y ss. BOE.

Desde esta óptica, cobra también relevancia la creación de mecanismos estables de supervisión, evaluación y transparencia sobre los sistemas de IA empleados en seguridad y justicia. La complejidad técnica del fenómeno aconseja no confiar únicamente en controles ad hoc dentro del proceso, sino complementar éstos con estructuras de gobernanza capaces de revisar sesgos, tasas de error, finalidades de uso y compatibilidad con los derechos fundamentales. En un terreno tan sensible como la investigación digital, la legitimidad tecnológica exige no solo eficacia, sino también rendición de cuentas institucional.

7. Conclusiones

La presente investigación ha permitido identificar la incidencia ambivalente de la inteligencia artificial en la ciberseguridad y en la ciberdelincuencia, así como los principales límites jurídicos que deben condicionar su utilización en tareas de prevención, detección e investigación. No obstante, conviene señalar de forma expresa algunas limitaciones del trabajo, tanto por exigencia metodológica como por honestidad académica.

En primer lugar, se trata de una investigación de carácter esencialmente jurídico-dogmático y documental. El estudio se ha construido a partir del análisis de normativa, doctrina e informes institucionales, lo que permite ofrecer una base sistemática y argumentativamente sólida, pero al mismo tiempo impide incorporar una dimensión empírica propia. No se han realizado entrevistas a operadores jurídicos, peritos, miembros de fuerzas y cuerpos de seguridad ni especialistas técnicos, ni tampoco un estudio cuantitativo autónomo sobre la incidencia práctica de las herramientas de inteligencia artificial en la investigación de ciberdelitos.

En segundo término, el objeto analizado se inserta en un entorno tecnológico especialmente dinámico. La evolución de la inteligencia artificial, de sus aplicaciones ofensivas y defensivas, y del ecosistema de amenazas digitales es tan rápida que algunas de las manifestaciones estudiadas pueden transformarse en plazos relativamente breves. Ello obliga a asumir que determinadas conclusiones prácticas del trabajo, aunque válidas en el momento de su formulación, deberán ser contrastadas con futuras actualizaciones normativas, técnicas y jurisprudenciales.

Asimismo, el trabajo se ha centrado específicamente en la aplicación de la inteligencia artificial a la prevención, detección e investigación de la ciberdelincuencia, dejando en un segundo plano otras cuestiones también relevantes, como la eventual responsabilidad derivada del diseño, comercialización o despliegue defectuoso de sistemas de inteligencia artificial, o el análisis exhaustivo de la jurisprudencia penal y procesal sobre evidencia digital. La delimitación del objeto de estudio ha resultado necesaria para garantizar la coherencia interna del TFM, pero al mismo tiempo ha supuesto la exclusión de líneas de análisis que podrían desarrollarse con mayor profundidad en investigaciones posteriores.

Por otra parte, el tratamiento de los colectivos especialmente vulnerables se ha realizado desde una perspectiva jurídico-criminológica, atendiendo a la especial exposición de menores, mujeres, personas mayores y de la esfera pública informativa frente a nuevas formas de victimización digital. Sin embargo, esta aproximación no ha pretendido constituir un estudio estadístico integral ni una investigación sociológica completa sobre el impacto diferencial del fenómeno, sino una identificación razonada de los principales espacios de vulnerabilidad reforzada.

Pese a estas limitaciones, la investigación sí permite sostener con suficiente solidez que la inteligencia artificial actúa actualmente como un factor de transformación estructural de la ciberdelincuencia y, al mismo tiempo, como una herramienta potencialmente útil para su prevención y detección, siempre que su utilización quede sometida a límites jurídicos efectivos, a supervisión humana significativa y a un marco robusto de garantías.

7.1. Recomendaciones

A la vista de lo expuesto, pueden formularse una serie de recomendaciones orientadas a reforzar la coherencia del marco jurídico y operativo aplicable a la inteligencia artificial en el ámbito de la ciberseguridad y de la ciberdelincuencia.

En primer lugar, resulta aconsejable avanzar hacia una mayor sistematización del tratamiento jurídico de la criminalidad digital y de los problemas específicos vinculados al uso de inteligencia artificial en su comisión, prevención e investigación. No se trata necesariamente de multiplicar nuevas figuras penales, sino de evitar dispersión interpretativa, mejorar la seguridad jurídica y ofrecer criterios más claros para la delimitación de conductas, riesgos y garantías aplicables.

En segundo término, convendría reforzar normativamente la exigencia de que la inteligencia artificial utilizada en contextos de seguridad o investigación penal opere siempre como herramienta auxiliar y no como sustituto de la valoración humana. Toda utilización de sistemas automatizados con potencial incidencia sobre derechos fundamentales debería quedar sometida a criterios expresos de necesidad, proporcionalidad, supervisión humana, trazabilidad y posibilidad real de revisión. En este punto, resulta especialmente importante impedir que las inferencias algorítmicas se conviertan en fundamento autónomo de imputación o en base suficiente para restringir derechos sin contraste individualizado.

En tercer lugar, sería recomendable fortalecer las obligaciones de transparencia, documentación, auditabilidad y gobernanza de los sistemas de inteligencia artificial empleados tanto por operadores públicos como por entidades privadas que participen en tareas de ciberseguridad. La eficacia técnica, por sí sola, no puede bastar como criterio de legitimidad. Es necesario garantizar que el funcionamiento del sistema pueda ser comprendido, revisado y sometido a control, especialmente cuando sus resultados influyan en decisiones con relevancia jurídica o procesal.

En cuarto lugar, debe impulsarse una formación especializada y continua de jueces, fiscales, abogados, fuerzas y cuerpos de seguridad, peritos y demás operadores jurídicos en materias como inteligencia artificial, prueba digital, análisis algorítmico y ciberseguridad. Sin una comprensión suficiente de estas herramientas, el control jurídico de su utilización corre el riesgo de convertirse en una supervisión meramente formal, incapaz de detectar sesgos, errores o afectaciones desproporcionadas a derechos fundamentales.

De igual modo, debería impulsarse la especialización en pericia informática aplicada a la impugnación de evidencias potencialmente manipuladas mediante IA, de forma que los órganos judiciales dispongan de criterios técnicos sólidos para valorar capturas de pantalla, audios, vídeos, imágenes sintéticas y documentos electrónicos. La generalización de los deepfakes aconseja evitar tanto la credulidad automática frente al soporte digital como el escepticismo genérico no sustentado en indicios verificables.

En quinto término, conviene reforzar la prevención no penal del fenómeno mediante políticas de alfabetización digital, concienciación y protección reforzada de colectivos especialmente vulnerables. La respuesta frente a la ciberdelincuencia impulsada por inteligencia artificial no puede descansar exclusivamente en la sanción ex post, sino que debe incorporar estrategias preventivas capaces de reducir la exposición al engaño, mejorar la identificación temprana del riesgo y fortalecer la resiliencia social frente a nuevas formas de fraude, manipulación y victimización digital.

Asimismo, en el terreno específico de la desinformación y de la afectación colectiva, la respuesta institucional debería priorizar la alfabetización digital, la verificación transparente y la prevención del daño antes que fórmulas expansivas de criminalización

que puedan tensionar indebidamente la libertad de expresión e información.

Finalmente, desde una perspectiva de lege ferenda, sería oportuno seguir profundizando en el desarrollo de marcos normativos que articulen de manera más precisa la relación entre inteligencia artificial, investigación tecnológica y garantías procesales. En particular, resultaría conveniente clarificar los límites del uso de herramientas algorítmicas en la fase de investigación, reforzar la exigencia de motivación cuando intervengan sistemas automatizados y asegurar que toda evidencia digital obtenida con apoyo tecnológico pueda ser sometida a contradicción efectiva, revisión judicial y control de fiabilidad.

En definitiva, la inteligencia artificial puede desempeñar un papel relevante en la protección frente a la ciberdelincuencia, pero su integración en el ámbito jurídico-penal solo será legítima si se produce dentro de un modelo garantista, técnicamente informado y compatible con los principios esenciales del Estado de Derecho.

Referencias

AUFERIL, B.: "Auge de la ciberdelincuencia impulsada por inteligencia artificial". En LinkedIn, 2024. Disponible en: https://www.linkedin.com/pulse/auge-de-la-ciberdelincuencia-impulsada-por-artificial-bruno-auferil--usnvf/

BOE: Instrumento de Ratificación del Convenio sobre la Ciberdelincuencia, hecho en Budapest el 23 de noviembre de 2001. Disponible en: https://www.boe.es/diario_boe/txt.php?id=BOE-A-2010-14221

BOE: Ley de Enjuiciamiento Criminal. Texto consolidado. Disponible en: https://www.boe.es/buscar/act.php?id=BOE-A-1882-6036

BOE: Ley Orgánica 13/2015, de 5 de octubre, de modificación de la Ley de Enjuiciamiento Criminal para el fortalecimiento de las garantías procesales y la regulación de las medidas de investigación tecnológica. Disponible en: https://www.boe.es/buscar/act.php?id=BOE-A-2015-10725

CENTRO CRIPTOLÓGICO NACIONAL (CCN-CERT): Aproximación a la Inteligencia Artificial y la ciberseguridad (BP/30). Ministerio de Defensa, octubre 2023.

DOMÍNGUEZ BARRAGÁN, María. Luisa.: "Inteligencia artificial en el proceso penal: ¿uso, recurso o abuso?", Inteligencia artificial y Derecho penal: desafíos éticos, aplicaciones prácticas y nuevos horizontes en la era digital. J.M. Bosch Editor, 2025, pp. 305-332. https://doi.org/10.2307/jj.39224064.13

ELDIARIO.COM: "FraudGPT, el método de ciberdelincuencia con inteligencia artificial". 24 de enero de 2024. Disponible en: https://eldiario.com/2024/01/24/fraudgpt-metodo-ciberdelincuencia-inteligencia-artificial/

ESPAÑA: Ley Orgánica 10/1995, de 23 de noviembre, del Código Penal.

ESPAÑA: Ley Orgánica 13/2015, de 5 de octubre, de modificación de la Ley de Enjuiciamiento Criminal para el fortalecimiento de las garantías procesales y la regulación de las medidas de investigación tecnológica.

ESPAÑA: Real Decreto de 14 de septiembre de 1882, por el que se aprueba la Ley de Enjuiciamiento Criminal.

EY: FERRÉ, X.: "Cómo la inteligencia artificial está cambiando la ciberdelincuencia". 25 de mayo de 2023. Disponible en: https://www.ey.com/es_es/cybersecurity/como-la-inteligencia-artificial-esta-cambiando-la-ciberdelincuencia

EWEEK: HITER, S.: "Generative AI and Cybersecurity". Junio 2023. Disponible en: https://www.eweek.com/artificial-intelligence/generative-ai-and-cybersecurity/

FLORES MARTÍN, Jackeline.: "La utilidad de los mecanismos de la inteligencia artificial

en la estrategia procesal penal", Inteligencia artificial y Derecho penal: desafíos éticos, aplicaciones prácticas y nuevos horizontes en la era digital. J.M. Bosch Editor, 2025, pp. 277-304. https://doi.org/10.2307/jj.39224064.12

GALÁN CORDERO, Carlos. y GALÁN PASCUAL, Carlos. *Aproximación a la Inteligencia Artificial y la ciberseguridad.* 2022.

ICSE - INTERPOL - ECPAT INTERNATIONAL: Towards a Global Indicator on Unidentified Victims in Child Sexual Exploitation Material. 2018.

INSTITUTO NACIONAL DE CIBERSEGURIDAD (INCIBE): "Pentesting". Disponible en: https://www.incibe.es/aprendeciberseguridad/pentesting

ISO/IEC 27001:2022: Information security, cybersecurity and privacy protection - Information security management systems- Requirements.

ISO/IEC 42001:2023: Information technology -Artificial intelligence- Management system.

LIZ RIVAS, Lenny. Violencia y agresión entre iguales a través de las TICS: Cyberbulling. AlmaMater. Cuadernos de Psicosociobiología de la Violencia: Educación y Prevención, nº 5, 2024, Dykinson, 2024 pp. 89-105. https://doi.org/10.14679/3314

LLEDÓ BENITO, I.: La robótica y la inteligencia artificial en la nueva era de la revolución industrial 4.0. Los desafíos jurídicos, éticos y tecnológicos de los robots inteligentes. Edit. Dykinson, Madrid, 2021.

LÓPEZ DE MÁNTARAS, Ramón. y MESEGUER GONZÁLEZ, Pedro.: Inteligencia artificial. Los Libros de la Catarata, Madrid, 2017.

MAZURIER, Pablo, Andrés., DELGADO MORÁN, Juan, José & PAYA SANTOS, Claudio, Augusto. Gobernanza constructivista de la internet. *Teoría y Praxis*, 2019 17(34), 107-130. https://doi.org/10.5377/typ.v1i34.14823

MINISTERIO DEL INTERIOR: Informe sobre la cibercriminalidad en España 2022. Gobierno de España, 2023.

NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (NIST): Framework for Improving Critical Infrastructure Cybersecurity: version 1.1. 2018.

PARLAMENTO EUROPEO Y CONSEJO DE LA UNIÓN EUROPEA: Reglamento (UE) 2016/679, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos.

PARLAMENTO EUROPEO Y CONSEJO DE LA UNIÓN EUROPEA: Reglamento (UE) 2024/1689, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial.

PAYÁ SANTOS, Claudio Augusto, DELGADO MORÁN, Juan José. Violencia de género en los jóvenes.: factores de protección frente a la violencia de género. en; "Vulnerabilidad de las víctimas desde la perspectiva de género. Una visión criminológica". 1st ed., 69–84. Dykinson. 2021.https://doi.org/10.2307/j.ctv282jjsk.6.

RUSSELL, Stuart. y NORVIG, Peter.: Inteligencia Artificial: un enfoque moderno. 2.ª ed., Pearson Educación, S.A., 2013.

SAVE THE CHILDREN: Violencia viral. Análisis de la violencia contra la infancia y la adolescencia en el entorno digital. 2019.

SIXMA VAN HEEMSTRA FERNÁNDEZ, Anna.: La Desinformación como Amenaza para la Democracia: El Caso del Brexit. Documento de Opinión IEEE 42/2023.

SUBIJANA ZUNZUNEGUI, Ignacio. José.: "El ciberterrorismo: Una perspectiva legal y judicial". En Eguzkilore, núm. 22, San Sebastián, 2008.

WESTERLUND, Mika.: "The Emergence of Deepfake Technology: A Review". Vol. 9, noviembre 2019. https://doi.org/10.22215/timreview/1282