Revista Peruana de Ciencias Penales | Número 39 • 2025 • pp. 40-68
DOI: https://doi.org/10.56176/rpcp.39.2025.150

# LA INTELIGENCIA ARTIFICIAL: SU USO ILÍCITO Y EL IMPACTO EN EL DERECHO PENAL

## *Artificial intelligence: illegal use and the impact on criminal law*

ÁNGELA CASALS FERNÁNDEZ* **

Recibido: 03.NOV.2025

Aprobado: 18.DIC.2025



**RESUMEN:**

La inteligencia artificial es uno de los fenómenos tecnológicos más desafiantes del siglo XXI. En el presente artículo se analiza cómo la autonomía técnica, la opacidad algorítmica y la imprevisibilidad de los resultados desafían los principios clásicos de acción, culpabilidad y responsabilidad. A partir del Reglamento (UE) 2024/1689 de Inteligencia Artificial, se aborda la clasificación de los sistemas de riesgo y las prácticas prohibidas, así como los nuevos escenarios delictivos derivados de la IA, desde los deepfakes hasta el fraude automatizado. La dogmática penal debe adaptarse al riesgo tecnológico, preservando la proporcionalidad, la legalidad y la tutela de los bienes jurídicos.

**PALABRAS CLAVE:** Inteligencia artificial, uso ilícito, delitos tecnológicos, derecho penal

**ABSTRACT:**

Artificial intelligence is one of the most challenging technological phenomena of the 21st century. This article analyses how technical autonomy, algorithmic opacity and the unpredictability of results challenge the classic principles of action, guilt and responsibility. Based on Regulation (EU) 2024/1689, on Artifical

* Profesora Ayudante. Doctora de Derecho Penal (Acreditada Contratada Doctora) en la Universidad Nacional de Educación a Distancia. Email: angelacasals@der.uned.es. Código ORCID: https://orcid.org/0000-0001-6602-3713.

** Artículo derivado del proyecto de investigación «Retos jurídicos en la aplicación de la IA predictiva en el sector sanitario» (IA JurisSalud). Referencia PID2024-160176NB-I00, Ministerio de Ciencia e Innovación, de la Universidad Autónoma de Barcelona. Investigadora principal: Dra. Dña. Sancha Camacho Clavijo.



ISSN 2306-7608 (Impresa) • ISSN 2810-8396 (En línea)

Intelligence it addresses the classification of risk systems and prohibited practices, as well as new criminal scenarios arising from AI, from deepfakes to automated fraud. Criminal dogma must adapt to technological risk, preserving proportionality, legality and the protection of legal rights.

**KEYWORDS:** artificial intelligence, illegal use, technology crimes, criminal law

## 1. INTRODUCCIÓN

El desarrollo de la inteligencia artificial (IA) constituye uno de los fenómenos tecnológicos más disruptivos de la era contemporánea. Para muchos es considerada la "Cuarta Revolución Industrial", toda vez que se ha producido un profundo cambio digital, promovido por redes inteligentes interconectadas que unen procesos industriales, tecnológicos y sociales, transformando la producción, el empleo y las interacciones humanas mediante la automatización sofisticada y el análisis masivo de datos. Podemos considerar, por lo tanto, que ha surgido una industria basada en fábricas inteligentes (García Mendiola, 2025, p. 300).

Esta creciente capacidad de los sistemas automatizados para ejecutar tareas complejas, aprender de la experiencia y tomar decisiones con un grado variable de autonomía ha transformado de manera irreversible los ámbitos económico, social y jurídico. En este contexto, el Derecho penal, como *ultima ratio* del ordenamiento jurídico y garante de los bienes jurídicos más esenciales, se encuentra frente a una serie de desafíos inéditos que ponen a prueba sus categorías tradicionales de imputación, culpabilidad y punibilidad

Actualmente, la IA no solo actúa como herramienta de apoyo en la investigación y persecución del delito, sino también como posible medio comisivo e, incluso, como factor que reconfigura las fronteras entre la acción humana y la intervención tecnológica. La posibilidad de que algoritmos o sistemas automatizados generen resultados lesivos con una dubitada intervención directa del ser humano plantea interrogantes sobre la atribución de responsabilidad penal, la validez de las pruebas obtenidas mediante sistemas automatizados o la compatibilidad de determinadas técnicas predictivas con los derechos fundamentales reconocidos en la Constitución Española y en el Derecho de la Unión Europea.

Es indudable que el progreso de la IA en los últimos años ha superado cualquier previsión tecnológica y jurídica previa. El salto cualitativo que han experimentado los sistemas de aprendizaje automático, el procesamiento masivo de datos y, más recientemente, los modelos generativos de lenguaje e imagen, ha transformado no solo la economía digital, sino también las estructuras sociales y los mecanismos

de decisión. Lo que en la década anterior se entendía como una herramienta auxiliar, hoy actúa como motor autónomo de producción, análisis y creación. Esta evolución, acelerada por la disponibilidad de datos y la potencia de cálculo, plantea al Derecho, y especialmente al Derecho penal, un desafío estructural: el de mantener la capacidad de tutela de bienes jurídicos en un entorno donde la acción humana se diluye tras la capa algorítmica (Varona Gómez, 2024, p. 12).

El Derecho penal se fundamenta en la imputación objetiva, la culpabilidad y la proporcionalidad de la sanción. Sin embargo, la IA introduce variables que erosionan estos pilares: autonomía técnica, opacidad decisional, imprevisibilidad de los resultados y, en algunos casos, ausencia de dolo o culpa humana directa (Silva Sánchez, 2025, p. 87). En consecuencia, el modelo clásico de responsabilidad penal se ve tensionado por un contexto en el que las conductas ilícitas pueden producirse mediante la intervención o el fallo de sistemas que «aprenden» por sí mismos. Ello obliga a reconsiderar la atribución de responsabilidad: si recae sobre el programador, el usuario, la empresa o, incluso, si se requiere un nuevo tipo de imputación objetiva por riesgo tecnológico (García Mendiola, 2025, p. 302).

Durante la última década, el uso criminal de la IA ha crecido en variedad y sofisticación. Los sistemas de generación de contenido, *deepfakes*, se emplean en delitos contra la intimidad, la imagen o el honor, mientras que los algoritmos de suplantación de voz y texto han incrementado la eficacia de estafas y fraudes telemáticos. La Policía Nacional y la Guardia Civil han advertido en múltiples ocasiones que el incremento de las estafas digitales en España está directamente vinculado al uso de inteligencia artificial para clonar identidades y manipular comunicaciones. El caso reciente de un joven detenido en Toledo por crear y distribuir imágenes pornográficas falsas de 26 mujeres reales, algunas menores de edad, generadas mediante IA, muestra el impacto tangible de estas tecnologías en la criminalidad contemporánea (SER Toledo, 2025). Asimismo, en Barcelona, la investigación penal a cuatro menores por manipular imágenes de sus compañeras de instituto para producir material sexual falso evidencia la capacidad de la IA para convertir un simple dispositivo doméstico en instrumento delictivo (Llanas, 2025). Estos hechos ponen de relieve que el riesgo penal de la IA no es una abstracción, sino una realidad creciente.

En la actualidad, el Código Penal español no tipifica de forma expresa los delitos cometidos con apoyo de IA, por lo que las conductas deben encajarse en figuras tradicionales como la revelación de secretos, el descubrimiento ilícito de imágenes o la suplantación de identidad (Silva Sánchez, 2025, p. 93). Sin embargo, la autonomía y el alcance de los sistemas actuales hacen que las tipificaciones resulten insuficientes; la frontera entre autor, partícipe y herramienta se difumina, dificultando la imputación. De hecho, las investigaciones judiciales recientes han mostrado las dificultades periciales para acreditar el origen humano de un contenido manipulado


ISSN 2306-7608 (Impresa) • ISSN 2810-8396 (En línea)

por IA o para determinar la intencionalidad de quien lo difundió (Serrano Ferrer, 2021, p. 54).

Desde la perspectiva del bien jurídico protegido, el uso ilícito de la IA afecta simultáneamente a la intimidad, la seguridad, la propiedad, así como a la confianza social en la información. La manipulación de contenidos audiovisuales puede generar daños reputacionales irreversibles y facilitar extorsiones, campañas de desinformación o chantajes. Más aún, la capacidad de los modelos generativos para producir material de apariencia verosímil plantea un problema probatorio de primer orden: ¿cómo determinar la autenticidad de una prueba digital en un proceso penal cuando incluso un perito especializado puede ser engañado por una falsificación algorítmica? (Colina Ramírez, 2023, p. 125). Esta cuestión obliga a reforzar los mecanismos de trazabilidad y a debatir sobre la presunción de autenticidad de la prueba digital, así como sobre la eventual introducción de certificados de procedencia basados en *blockchain* o metadatos forenses.

A ello se suma un segundo plano de riesgo, el uso institucional de la IA dentro del propio sistema de justicia penal. Los sistemas de predicción de reincidencia o de evaluación del riesgo de fuga, inspirados en modelos estadounidenses, se han planteado en Europa como posibles herramientas de apoyo a jueces y fiscales. Sin embargo, su adopción suscita preocupaciones en torno a la transparencia del algoritmo, los sesgos de entrenamiento y el derecho a un juicio justo. Si una persona es valorada por un sistema opaco cuya lógica ni siquiera los magistrados pueden explicar, se vulnera el principio de culpabilidad personal y la exigencia de motivación de las resoluciones judiciales (Lang Irrazábal, 2022, p. 37). En este punto, el desafío no es solo técnico, sino ético y constitucional, cómo compatibilizar la eficiencia del dato con el respeto a la dignidad humana.

El marco regulatorio español y europeo intenta dar respuesta a estos desafíos, aunque con un retraso estructural frente al ritmo del desarrollo tecnológico. La Unión Europea ha avanzado con el Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial que clasifica los sistemas según su nivel de riesgo y prevé sanciones para usos prohibidos o de alto riesgo. No obstante, este instrumento pertenece al ámbito administrativo y de mercado, no al penal. España, por su parte, ha anunciado iniciativas legislativas para obligar a identificar los contenidos generados por IA y sancionar su uso indebido. Pese a ello, persiste un vacío en cuanto a la responsabilidad penal por daños o por la generación de material ilícito. El legislador penal aún no ha adaptado la dogmática de autoría, causalidad y culpabilidad a los entornos algorítmicos, lo que deja espacio para la impunidad en conductas que, aunque gravemente lesivas, no encajan en tipos preexistentes (Valls Prieto, 2022, p.4).

En este contexto, diversos autores proponen la incorporación de nuevas figuras delictivas específicas para el uso abusivo de IA, así como la revisión de los criterios de imputación en delitos tecnológicos. Una línea doctrinal aboga por establecer una responsabilidad penal por riesgo tecnológico, similar a la aplicada en materia medioambiental o de seguridad industrial, donde la imputación se base en la creación o mantenimiento de un sistema autónomo potencialmente lesivo (Silva Sánchez, 2011). Otra corriente, más restrictiva, defiende mantener la estructura clásica de culpabilidad, entendiendo la IA como mera extensión instrumental de la conducta humana (Bustos Ramírez, 2024). Ambas visiones coinciden, sin embargo, en que la falta de previsión normativa genera inseguridad jurídica y dificulta la persecución eficaz de nuevas formas de criminalidad digital. Como señala Lloria García (2013) la tipificación de nuevas conductas al hilo del cambio tecnológico puede ser necesaria pues «no siempre los instrumentos tradicionales del derecho penal son válidos para resolver las cuestiones que surgen a propósito de las lesiones a bienes jurídicos nuevos o la afectación de los tradicionalmente tutelados con una mayor intensidad». Esta operación no es sencilla, existiendo en muchos casos dudas sobre los bienes jurídicos afectados por los comportamientos cometidos a través de las nuevas tecnologías, y diferentes instrumentos supranacionales que empujan a la criminalización que pueden crear disrupciones en el ordenamiento nacional.

La problemática se agrava en el terreno internacional. La IA permite la comisión de delitos transnacionales, la ocultación del rastro digital y la manipulación de jurisdicciones. Así, una imagen falsa generada en un servidor fuera de la Unión Europea puede difundirse en segundos en España, afectando a víctimas locales sin que exista un autor nacional identificable. Este desajuste entre territorialidad del delito y globalidad del medio exige repensar los mecanismos de cooperación judicial y las normas de competencia internacional. Del mismo modo, la cibercriminalidad basada en IA desafía los principios de proporcionalidad de las penas, pues los daños pueden multiplicarse exponencialmente con un coste mínimo para el infractor (Mínguez Rosique; Gallego Arribas, 2025, p. 17).

Frente a este panorama, la respuesta penal debe ser prudente pero decidida. Una sobrerreacción legislativa podría vulnerar derechos fundamentales o criminalizar conductas sin dolo ni riesgo real; una respuesta tardía, en cambio, favorecería la impunidad y la pérdida de confianza social en la justicia. Por ello, se requiere un enfoque multidisciplinar que combine la actualización del Código Penal con políticas preventivas, educación digital, mecanismos de trazabilidad tecnológica y cooperación internacional. El Derecho penal, en definitiva, debe conservar su función de *ultima ratio*, pero adaptarse a un escenario en el que la acción humana y la inteligencia artificial se entrelazan.


ISSN 2306-7608 (Impresa) • ISSN 2810-8396 (En línea)

La reflexión científica en este ámbito no puede limitarse a denunciar los riesgos, sino que ha de contribuir a redefinir los principios de imputación, prueba y culpabilidad a la luz de los sistemas autónomos. La IA no es un agente moral, pero sí una realidad capaz de generar consecuencias jurídicas y daños sociales. En esa frontera, entre la autonomía técnica y la responsabilidad humana, se decidirá el futuro del Derecho penal en la era de la IA y será precisamente en la capacidad del sistema para responder con garantías donde se medirá no solo su eficacia, sino su legitimidad.

## 2. MARCO CONCEPTUAL Y NORMATIVO DE LA INTELIGENCIA ARTIFICIAL: REGLAMENTO (UE) 2024/1689 DEL PARLAMENTO EUROPEO Y DEL CONSEJO, DE 13 DE JUNIO DE 2024, POR EL QUE SE ESTABLECEN NORMAS ARMONIZADAS EN MATERIA DE INTELIGENCIA ARTIFICIAL

El concepto de inteligencia artificial fue utilizado, por primera vez, por John McCarthy, profesor de matemáticas del Darmont College (New Hampshire), en el año 1955, definiéndola como «la ciencia y la ingeniería de crear máquinas inteligentes, especialmente aquellas programadas con computación inteligente». McCarthy participaba en un proyecto de investigación con otros científicos de diversas disciplinas, con el fin de crear una máquina que utilizara lenguaje, formara abstracciones, resolviera problemas, entre otras cosas, mediante la imitación de los procesos del pensamiento humano. Si bien en un principio estos trabajos tuvieron bastante aceptación en el mundo científico, a mediados de los años setenta fueron dejados de lado, debido a las dificultades que presentaban en aquel momento. Durante un largo período esta disciplina tuvo un escaso desarrollo, resurgiendo a finales de la década de los noventa (Lang Irrazábal, 2022, p.32).

Actualmente, y desde un punto de vista técnico, la inteligencia artificial puede definirse como el conjunto de sistemas capaces de realizar tareas que, si fueran ejecutadas por un ser humano, requerirían inteligencia, es decir, aprendizaje, razonamiento, percepción o toma de decisiones. Según el Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial (en adelante, Reglamento de Inteligencia Artificial), en su artículo 3 define sistema de IA como «un sistema basado en una máquina que está diseñado para funcionar con distintos niveles de autonomía y que puede mostrar capacidad de adaptación tras el despliegue, y que, para objetivos explícitos o implícitos, infiere de la información de entrada que recibe la manera de generar resultados de salida, como predicciones, contenidos, recomendaciones o decisiones, que pueden influir en entornos físicos o virtuales».

En el plano jurídico, la IA plantea una dificultad adicional, su conceptualización debe atender no solo a sus capacidades funcionales, sino también a las implicaciones normativas derivadas de su autonomía. El Derecho no puede limitarse a reproducir definiciones técnicas, sino que debe valorar en qué medida el comportamiento de un sistema automatizado puede ser jurídicamente relevante, es decir, susceptible de imputación o de control normativo.

Debemos tener en cuenta que el término inteligencia «artificial» implica una distinción respecto a la inteligencia «natural», la propia de los seres humanos, y se refiere a que el origen de la inteligencia es el resultado de un esfuerzo informático intencionado en lugar de la inteligencia de una persona. Por eso se habla, principalmente, de la existencia de dos enfoques básicos para la IA. En primer lugar, la IA basada en el conocimiento o *knowledge-based AI*, trabaja con una representación simbólica explícita del conocimiento, es decir, a partir de expertos humanos o documentos se crea un algoritmo para razonar e inferir soluciones a problemas o consultas en un ámbito particular. Ejemplo de ello son los sistemas expertos (Martino, 1992, p. 57). Y, en segundo lugar, la IA basada en datos o *data-driven AI*, más conocida como *machine learning*, la cual se centra en el aprendizaje a partir de ejemplos o de la experiencia en el uso del sistema. Los datos observados representan información incompleta sobre los acontecimientos y los algoritmos de aprendizaje tratan de generalizar esa información para hacer predicciones sobre los sucesos conocidos, es por ello por lo que, debido a la ingente cantidad de datos y recursos informáticos necesarios para ser eficaz, ha sido posible su avance, gracias a Internet y a la computación en la nube. Por eso, actualmente, se ha producido el auge del *big data*, al estar ligado al desarrollo actual del aprendizaje automático. Indudablemente, cada enfoque tiene sus propias ventajas y limitaciones. El aprendizaje automático es más útil en operaciones que requieren la identificación de patrones y, en cambio, la IA basada en el conocimiento sigue siendo útil para tareas en campos específicos donde conocimiento y razonamiento están bien sistematizados. No obstante, se sabe que en la práctica muchos sistemas inteligentes utilizan diversos componentes provenientes de ambos enfoques, esto se conoce como «sistema multiagente», los *multi-agent systems* o MAS (Barrio Andrés, 2024, p. 24).

La doctrina distingue entre IA débil, orientada a la ejecución de tareas específicas mediante aprendizaje estadístico o algoritmos deterministas, e IA fuerte, caracterizada por su capacidad de aprendizaje autónomo y razonamiento adaptativo. Desde la perspectiva penal, esta clasificación adquiere relevancia porque cuanto mayor es la autonomía del sistema, menor es la capacidad de control directo por parte del ser humano, lo que complica la atribución de responsabilidad.

En el ámbito práctico, los sistemas utilizados por las fuerzas y cuerpos de seguridad del Estado o en el análisis de datos judiciales suelen encuadrarse dentro de


ISSN 2306-7608 (Impresa) • ISSN 2810-8396 (En línea)

la IA débil: herramientas predictivas, reconocimiento facial o clasificación automatizada de riesgos. Sin embargo, los desarrollos hacia sistemas con mayor autonomía decisional, por ejemplo, vehículos autónomos o sistemas de vigilancia inteligente, introducen nuevas zonas grises entre acción humana y decisión algorítmica (Lang Irrazábal, 2022, p. 35).

En cuanto a su regulación, el primer marco normativo mundial sobre ética de la IA ha sido la «Recomendación sobre la Ética de la Inteligencia Artificial» de 23 noviembre de 2021, adoptado por los 193 Estados miembros de la UNESCO. Los objetivos de la Recomendación se sitúan en el punto 8 donde expresa que son los siguientes «a) proporcionar un marco universal de valores, principios y acciones para orientar a los Estados en la formulación de sus leyes, políticas u otros instrumentos relativos a la IA, de conformidad con el Derecho internacional; b) orientar las acciones de las personas, los grupos, las comunidades, las instituciones y las empresas del sector privado a fin de asegurar la incorporación de la ética en todas las etapas del ciclo de vida de los sistemas de IA; c) proteger, promover y respetar los derechos humanos y las libertades fundamentales, la dignidad humana y la igualdad, incluida la igualdad de género; salvaguardar los intereses de las generaciones presentes y futuras; preservar el medio ambiente, la biodiversidad y los ecosistemas; y respetar la diversidad cultural en todas las etapas del ciclo de vida de los sistemas de IA; d) fomentar el diálogo multidisciplinario y pluralista entre múltiples partes interesadas y la concertación sobre cuestiones éticas relacionadas con los sistemas de IA; e) promover el acceso equitativo a los avances y los conocimientos en el ámbito de la IA y el aprovechamiento compartido de los beneficios, prestando especial atención a las necesidades y contribuciones de los países de ingreso mediano bajo, incluidos los PMA, los PDSL y los PEID». Además, recomienda a los Estados miembro invertir en competencias digitales y de alfabetización mediática e información para fomentar el pensamiento crítico sobre el uso de la IA.

La Recomendación hace mención al conjunto de principios éticos y jurídicos que operan como límites materiales, siendo los siguientes: en primer lugar, la proporcionalidad y la inocuidad (25 y 26); debemos tener en cuenta que el método de IA elegido debería ser adecuado y proporcional para lograr un objetivo legítimo determinado. Su utilización no debe constituir una violación o un abuso de los derechos humanos, no debiendo utilizarse con fines de calificación social o vigilancia masiva. En segundo lugar, la seguridad y la protección (27), propiciándose mediante el desarrollo de marcos de acceso a los datos sostenibles, respetando la privacidad y fomentando la utilización de datos de calidad. En tercer lugar, la equidad y la no discriminación (28, 29 y 30), debiendo adoptar un enfoque inclusivo para garantizar que los beneficios de las tecnologías de la IA estén disponibles y sean accesibles para todos, teniendo en cuenta las necesidades específicas de los diferentes grupos de edad, los sistemas culturales, los diferentes grupos lingüísticos, las personas con

discapacidad, las niñas y las mujeres y las personas desfavorecidas, marginadas y vulnerables o en situación de vulnerabilidad. En estos puntos, además, se hace mención expresa a que los Estados Miembros deberían esforzarse por reducir las brechas digitales y garantizar el acceso inclusivo al desarrollo de la IA y la participación en él. En cuarto lugar, la sostenibilidad (31); la evaluación continua de los efectos humanos, sociales, culturales, económicos y ambientales de las tecnologías de la IA debería llevarse a cabo con pleno conocimiento de las repercusiones de dichas tecnologías en la sostenibilidad como un conjunto de metas en constante evolución en toda una serie de dimensiones, como las que se definen actualmente en los Objetivos de Desarrollo Sostenible de las Naciones Unidas. En quinto lugar, el derecho a la intimidad y la protección de datos (32, 33 y 34); la privacidad debe ser respetada, protegida y promovida, para ello los marcos de protección de datos y todo mecanismo conexo deberían tomar como referencia los principios y normas internacionales de protección de datos relativos a la recopilación, la utilización y la divulgación de datos personales y al ejercicio de sus derechos por parte de los interesados, garantizando al mismo tiempo un objetivo legítimo y una base jurídica válida para el tratamiento de los datos personales, incluido el consentimiento informado. En sexto lugar, la supervisión y decisión humanas (35 y 36); se debe velar por que siempre sea posible atribuir la responsabilidad ética y jurídica, en cualquier etapa del ciclo de vida de los sistemas de IA, a personas físicas o a entidades jurídicas existentes. Y como punto clave, puede ocurrir que, en algunas ocasiones, los seres humanos decidan depender de los sistemas de IA por razones de eficacia, pero la decisión de ceder el control en contextos limitados seguirá recayendo en los seres humanos, ya que estos pueden recurrir a los sistemas de IA en la adopción de decisiones y en la ejecución de tareas, pero un sistema de IA nunca podrá reemplazar la responsabilidad final de los seres humanos y su obligación de rendir cuentas. En séptimo lugar, la transparencia y la explicabilidad (37 a 41); ambas son condiciones previas fundamentales para garantizar el respeto, la protección y la promoción de los derechos humanos, las libertades fundamentales y los principios éticos. Desde un punto de vista sociotécnico, una mayor transparencia contribuye a crear sociedades más pacíficas, justas, democráticas e inclusivas. Posibilita un escrutinio público que puede reducir la corrupción y la discriminación, y también puede ayudar a detectar y prevenir los efectos negativos sobre los derechos humanos. En octavo lugar, la responsabilidad y rendición de cuentas (42 y 43); los actores de la IA y los Estados Miembros deberían respetar, proteger y promover los derechos humanos y las libertades fundamentales, y deberían también fomentar la protección del medio ambiente y los ecosistemas, asumiendo su responsabilidad ética y jurídica respectiva, de conformidad con el derecho nacional e internacional, en particular las obligaciones de los Estados Miembros en materia de derechos humanos. Además, deberían elaborarse mecanismos adecuados de supervisión, evaluación del impacto, auditoría y diligencia debida, incluso en lo


ISSN 2306-7608 (Impresa) • ISSN 2810-8396 (En línea)

que se refiere a la protección de los denunciantes de irregularidades, para garantizar la rendición de cuentas respecto de los sistemas de IA y de su impacto a lo largo de su ciclo de vida. Y, en noveno lugar, la gobernanza y colaboración adaptativas y de múltiples partes interesadas (46 y 47); se debe tener en cuenta que en la utilización de datos deben respetarse el derecho internacional y la soberanía nacional, esto quiere decir que los Estados pueden regular los datos generados dentro de sus territorios o que pasan por ellos y adoptar medidas para la regulación efectiva de los datos, en particular su protección, sobre la base del respeto del derecho a la privacidad.

Por otro lado, y siendo de plena actualidad, nos encontramos el Reglamento de Inteligencia Artificial, también conocida por la Ley Europea de Inteligencia Artificial, que constituye la primera ley integral que se aplicará a los agentes tanto públicos como privados, de dentro y fuera de la Unión Europea, en la medida en que el sistema de IA se introduzca en el mercado de la Unión Europea o su uso afecte a personas establecidas en ella. La Unión Europea se sitúa así a la vanguardia de la regulación de la IA, la otra regulación general importante, en Estados Unidos, constituida por la Orden Ejecutiva del presidente Joe Biden, norma con fuerza de ley, de 30 de octubre de 2023, es mucho menos intensa en su carga obligacional y no proviene de una fuente parlamentaria. No obstante, la regulación europea no sólo contempla mecanismos de control y regulación, sino también, y es importante destacarlo, de fomento, como la regulación de los espacios controlados de pruebas, *sandboxes,* y las medidas a favor del desarrollo por las pymes de sistemas de IA, destinadas a incentivar un desarrollo de estas tecnologías sostenible en términos sociales.

El Reglamento de Inteligencia Artificial, aprobado en 2024, constituye, como hemos mencionado anteriormente, el primer marco normativo integral sobre la materia. Entró en vigor el 1 de agosto de 2024, pero su aplicación se realiza por fases. Las prohibiciones y obligaciones de transparencia se aplicarán desde el 2 de febrero de 2025. Las normas sobre gobernanza y modelos de propósito general entrarán en vigor el 2 de agosto de 2025. La plena aplicación del reglamento, incluyendo las normas sobre sistemas de alto riesgo, se completará el 2 de agosto de 2026. Su objetivo es garantizar que los sistemas de IA introducidos en el mercado europeo sean seguros y respeten los derechos fundamentales y los valores de la Unión. La norma establece un sistema de clasificación por niveles de riesgo, que impone mayores obligaciones a los sistemas de alto riesgo, entre los que se incluyen los utilizados en la administración de justicia y la aplicación del Derecho penal.

Aunque el Reglamento no crea tipos penales ni modifica directamente el Código Penal español, su incidencia indirecta es profunda. La utilización de sistemas de IA en la persecución penal o en la evaluación del riesgo de reincidencia deberá someterse a criterios de transparencia, supervisión humana y trazabilidad. De igual modo, la infracción de estos estándares podría derivar en responsabilidades penales

si el uso de la tecnología genera daños graves o vulnera derechos fundamentales, en especial los reconocidos en los artículos 18 y 24 de la Constitución Española.

Es de interés revisar lo mencionado en el artículo 5 del Reglamento de Inteligencia Artificial donde se indican las prácticas IA prohibidas:

«a) La introducción en el mercado, la puesta en servicio o la utilización de un sistema de IA que se sirva de técnicas subliminales que trasciendan la conciencia de una persona o de técnicas deliberadamente manipuladoras o engañosas con el objetivo o el efecto de alterar de manera sustancial el comportamiento de una persona o un colectivo de personas, mermando de manera apreciable su capacidad para tomar una decisión informada y haciendo que tomen una decisión que de otro modo no habrían tomado, de un modo que provoque, o sea razonablemente probable que provoque, perjuicios considerables a esa persona, a otra persona o a un colectivo de personas.

b) La introducción en el mercado, la puesta en servicio o la utilización de un sistema de IA que explote alguna de las vulnerabilidades de una persona física o un determinado colectivo de personas derivadas de su edad o discapacidad, o de una situación social o económica específica, con la finalidad o el efecto de alterar de manera sustancial el comportamiento de dicha persona o de una persona que pertenezca a dicho colectivo de un modo que provoque, o sea razonablemente probable que provoque, perjuicios considerables a esa persona o a otra.

c) La introducción en el mercado, la puesta en servicio o la utilización de sistemas de IA para evaluar o clasificar a personas físicas o a colectivos de personas durante un período determinado de tiempo atendiendo a su comportamiento social o a características personales o de su personalidad conocidas, inferidas o predichas, de forma que la puntuación ciudadana resultante provoque una o varias de las situaciones siguientes: i) un trato perjudicial o desfavorable hacia determinadas personas físicas o colectivos de personas en contextos sociales que no guarden relación con los contextos donde se generaron o recabaron los datos originalmente, ii) un trato perjudicial o desfavorable hacia determinadas personas físicas o colectivos de personas que sea injustificado o desproporcionado con respecto a su comportamiento social o la gravedad de este.

d) La introducción en el mercado, la puesta en servicio para este fin específico o el uso de un sistema de IA para realizar evaluaciones de riesgos de personas físicas con el fin de valorar o predecir el riesgo de que una persona física cometa un delito basándose únicamente en la elaboración


ISSN 2306-7608 (Impresa) • ISSN 2810-8396 (En línea)

del perfil de una persona física o en la evaluación de los rasgos y características de su personalidad; esta prohibición no se aplicará a los sistemas de IA utilizados para apoyar la valoración humana de la implicación de una persona en una actividad delictiva que ya se base en hechos objetivos y verificables directamente relacionados con una actividad delictiva.

e) La introducción en el mercado, la puesta en servicio para este fin específico o el uso de sistemas de IA que creen o amplíen bases de datos de reconocimiento facial mediante la extracción no selectiva de imágenes faciales de internet o de circuitos cerrados de televisión.

f) La introducción en el mercado, la puesta en servicio para este fin específico o el uso de sistemas de IA para inferir las emociones de una persona física en los lugares de trabajo y en los centros educativos, excepto cuando el sistema de IA esté destinado a ser instalado o introducido en el mercado por motivos médicos o de seguridad.

g) La introducción en el mercado, la puesta en servicio para este fin específico o el uso de sistemas de categorización biométrica que clasifiquen individualmente a las personas físicas sobre la base de sus datos biométricos para deducir o inferir su raza, opiniones políticas, afiliación sindical, convicciones religiosas o filosóficas, vida sexual u orientación sexual; esta prohibición no incluye el etiquetado o filtrado de conjuntos de datos biométricos adquiridos lícitamente, como imágenes, basado en datos biométricos ni la categorización de datos biométricos en el ámbito de la garantía del cumplimiento del Derecho.

h) El uso de sistemas de identificación biométrica remota «en tiempo real» en espacios de acceso público con fines de garantía del cumplimiento del Derecho, salvo y en la medida en que dicho uso sea estrictamente necesario para alcanzar uno o varios de los objetivos siguientes: i) 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, ii) 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, iii) 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».

Este artículo constituye el núcleo ético-jurídico del Reglamento de Inteligencia Artificial, al fijar los límites absolutos al desarrollo y utilización de sistemas de IA en la Unión Europea. Frente a los modelos basados en el riesgo, alto, limitado o mínimo, este artículo establece una categoría de prácticas directamente prohibidas, cuya ilicitud no depende de evaluación previa ni de medidas de mitigación, su sola existencia contraviene los valores fundamentales de la Unión y los derechos reconocidos en la Carta. La lógica del precepto es preventiva y protectora, orientada a salvaguardar la autonomía individual, la dignidad humana y la no discriminación. De ahí que los supuestos enumerados se centren en la manipulación del comportamiento humano, la explotación de vulnerabilidades, la vigilancia masiva y la discriminación biométrica o social. Se busca evitar la creación de entornos donde la IA se convierta en instrumento de control o exclusión social.

El apartado (a) prohíbe las técnicas subliminales o manipuladoras que menoscaben la capacidad de decisión libre e informada. La noción de «técnicas que trascienden la conciencia» implica cualquier forma de influencia conductual automatizada que actúe sobre sesgos cognitivos sin conocimiento del individuo, una práctica especialmente relevante en la publicidad digital o la microsegmentación política.

El apartado (b) extiende la prohibición a los sistemas que se aprovechan de vulnerabilidades ligadas a la edad, la discapacidad o la precariedad socioeconómica. Este inciso refuerza la idea de protección de colectivos en situación de desigualdad estructural, introduciendo una dimensión social que trasciende la mera protección de datos.

Los apartados (c) y (d) se dirigen a impedir el uso de la IA como instrumento de *profiling* social y penal. En particular, el (c) veta los denominados «sistemas de puntuación social», inspirados en el modelo chino, que suponen una forma de sanción extrajurídica y perpetúan desigualdades. El (d) prohíbe la predicción delictiva basada en rasgos personales, por considerar que convierte correlaciones estadísticas en juicios morales o penales, rompiendo el principio de responsabilidad individual.

Los apartados (e) al (h) abordan la dimensión biométrica y de vigilancia. Se prohíbe la creación de bases de datos faciales masivas sin consentimiento, la inferencia emocional en contextos laborales o educativos, por su potencial disciplinario, la categorización biométrica con fines discriminatorios, y el uso de identificación remota en tiempo real salvo excepciones muy tasadas de seguridad pública. Este último punto representa uno de los debates más tensos del Reglamento, al intentar equilibrar seguridad y derechos fundamentales.

Hay que tener en cuenta que el Reglamento dispone de las oportunas sanciones en su Capítulo XII, en concreto en el artículo 99.3 nos dice que «el no respeto de la prohibición de las prácticas de IA a que se refiere el artículo 5 estará sujeto a


ISSN 2306-7608 (Impresa) • ISSN 2810-8396 (En línea)

multas administrativas de hasta 35.000.000 € o, si el infractor es una empresa, de hasta el 7 % de su volumen de negocios mundial total correspondiente al ejercicio financiero anterior, si esta cuantía fuese superior». Si fueran instituciones, órganos u organismos de la UE, según el artículo 100.2 nos encontraríamos con que «el no respeto de la prohibición de las prácticas de IA a que se refiere el artículo 5 estará sujeto a multas administrativas de hasta 1.500.000 €».

En último lugar, queremos mencionar la Agencia Española de Supervisión de Inteligencia Artificial (AESIA) creada a través de la disposición adicional centésimo trigésima de la Ley 22/2021, de 28 de diciembre, de Presupuestos Generales del Estado para el año 2022. Es una agencia del sector público estatal adscrita al Ministerio para la Transformación Digital y de la Función Pública a través de la Secretaría de Estado de Digitalización e Inteligencia Artificial, que tiene por objeto garantizar un uso seguro y ético de la IA en España. La labor de la AESIA, junto con la de otros organismos públicos, es eliminar o reducir los riesgos que la IA puede generar en áreas como la seguridad, la integridad, la intimidad, la salud, la igualdad de trato y los derechos fundamentales de las personas. Corresponde a la AESIA la supervisión, asesoramiento, concienciación y formación para la adecuada implementación de toda la normativa nacional y europea en torno al adecuado uso y desarrollo de los sistemas de inteligencia artificial. En sus funciones de supervisión le corresponde la función de inspección, comprobación y sanción previstas en la normativa europea de Inteligencia Artificial, en concreto, el Reglamento de Inteligencia Artificial.

## 3. ALGUNAS CUESTIONES SOBRE LA INTELIGENCIA ARTIFICIAL COMO HERRAMIENTA AL SERVICIO DE LA JUSTICIA PENAL

Es indudable que la IA actúa con una velocidad mucho mayor que un ser humano, procesando en apenas segundos datos que a una persona le pueden llevar horas o incluso días. Y si, además, añadimos que no se cansa, no se puede distraer o, incluso, no enferma, tenemos una gran ventaja en cuanto a evitar errores en el traspaso de información. Esto, trasladado a la justicia, puede producir que el juez cuente con una mejor asistencia en su labor y que tanto él, como los empleados del Juzgado, puedan dedicar su tiempo a tareas de mayor importancia y de exclusividad humana (Lang Irrazábal, 2022, p. 34). Por eso, el uso de la inteligencia artificial en la administración de justicia y en las labores de investigación criminal ha experimentado un notable crecimiento en los últimos años, pero también es cierto que en el ámbito penal el uso de la IA es muy cuestionado. Entre las principales aplicaciones destacan el reconocimiento facial, la predicción de conductas delictivas, la identificación automatizada de sospechosos o la clasificación de riesgos de reincidencia.

No obstante, el empleo de estas herramientas plantea cuestiones jurídicas de primer orden. En particular, la legitimidad de la intervención penal basada en modelos predictivos o *predictive policing* (Cinelli; Manrique Gan, 2019, p. 5) exige compatibilizar la eficacia tecnológica con los derechos fundamentales a la intimidad, la protección de datos personales y la presunción de inocencia. El sesgo algorítmico puede generar discriminación indirecta por motivos de raza, género, origen o condición socioeconómica, vulnerando el artículo 14 de la Constitución Española y el artículo 21 de la Carta de Derechos Fundamentales de la Unión Europea. Desde la óptica del Derecho penal, un sistema que clasifica a individuos o territorios como «de riesgo» en función de datos sesgados puede conducir a una persecución selectiva, erosionando el principio de igualdad ante la ley y el principio de culpabilidad. Además, el empleo de técnicas predictivas puede tensionar la frontera entre prevención y punición. Si el Derecho penal se fundamenta en la reacción frente a hechos cometidos y no en la mera peligrosidad, la utilización de algoritmos que anticipan conductas puede derivar en una deriva de derecho penal del enemigo, incompatible con el Estado de Derecho (Cinelli; Manrique Gan, 2019, p. 12).

La Ley Orgánica 7/2021, de 26 de mayo, de protección de datos personales tratados para fines de prevención, detección, investigación y enjuiciamiento de infracciones penales, establece límites estrictos a la automatización de decisiones, exigiendo la intervención humana significativa y la posibilidad de revisión individualizada. Asimismo, el Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, 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, o también conocido como Reglamento General de Protección de Datos, en su artículo 22.1 expresa que «todo interesado tendrá derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado, incluida la elaboración de perfiles, que produzca efectos jurídicos en él o le afecte significativamente de modo similar». En el ámbito penal, esto se traduce en la necesidad de garantizar que las decisiones que puedan derivar en imputación, condena o medidas restrictivas de derechos no dependan únicamente de algoritmos.

Otro ámbito de especial relevancia es la utilización de sistemas de IA en la obtención y valoración de prueba. Los algoritmos empleados para analizar grabaciones, identificar patrones de voz o autenticar documentos digitales deben someterse a los principios de cadena de custodia, integridad probatoria y contradicción. El Tribunal Supremo ha señalado en diversas resoluciones como, por ejemplo, STS 259/2020, de 5 de junio (ECLI:ES:TS:2020:1326) que las pruebas obtenidas mediante medios tecnológicos deben garantizar la posibilidad de verificación y contradicción por las partes, de lo contrario, se vulneraría el derecho de defensa reconocido en el artículo 24 de la Constitución Española, en relación con el principio de contradicción (Cuadrado Salinas, 2025, p. 125). En consecuencia, un resultado derivado de un sistema


ISSN 2306-7608 (Impresa) • ISSN 2810-8396 (En línea)

algorítmico no transparente o no auditable podría ser impugnado por vulneración del derecho de defensa (artículo 24 Constitución Española). En este sentido, el principio de explicabilidad, que anteriormente vimos en la doctrina europea sobre IA fiable, se erige como requisito indispensable para la validez de las pruebas automatizadas, toda vez que el juez, el fiscal y la defensa deben poder conocer cómo el sistema llegó a una determinada conclusión.

Por lo tanto, el uso de la IA en el proceso penal, aunque útil, implica una potencial afectación a derechos como la presunción de inocencia, la igualdad de armas, el derecho a la privacidad y la tutela judicial efectiva.

Asimismo, como gran parte de la doctrina considera, la confianza excesiva en la objetividad de los algoritmos, fenómeno conocido como *automation bias*, puede conducir a que operadores jurídicos asuman los resultados tecnológicos como neutrales, sin someterlos a suficiente escrutinio (Freile Mansilla, 2025, p. 195). Desde una perspectiva constitucional, tal delegación acrítica resulta incompatible con el principio de culpabilidad personal y con el mandato de motivación judicial de las resoluciones (artículo 120.3 Constitución Española). En conclusión, la IA puede aportar eficiencia al sistema penal, pero solo si su utilización se somete a estrictos controles de transparencia, trazabilidad y revisión humana.

## 4. LA INTELIGENCIA ARTIFICIAL COMO INSTRUMENTO DEL DELITO

La IA no solo se emplea como instrumento de persecución del delito, sino también como medio comisivo. En la actualidad, numerosos delitos tradicionales encuentran en la tecnología un nuevo cauce de ejecución: fraudes automatizados, falsificación de contenidos mediante *deepfakes*, ataques informáticos autónomos o manipulación de mercados a través de algoritmos de alta frecuencia. Estos fenómenos cuestionan la suficiencia de los tipos penales existentes, que fueron diseñados bajo la premisa de la acción humana directa. La dificultad radica en determinar el grado de dominio del hecho que conserva el sujeto cuando el resultado es generado por un sistema de IA que actúa de manera parcialmente autónoma.

La IA no es un sujeto de derecho, ni se le reconoce, de momento personalidad jurídica, por lo que la responsabilidad penal por las infracciones cometidas utilizándola se imputaría a las personas físicas que utilizaran esta tecnología, o que la programaran para delinquir. El Código Penal español ofrece mecanismos para abordar estas situaciones, principalmente a través de la autoría mediata y la comisión por omisión, según los artículos 28 y 11 del Código Penal. Sin embargo, la casuística tecnológica exige reinterpretar estas categorías a la luz de la imputación objetiva: ¿hasta qué punto el resultado puede atribuirse al sujeto cuando media un sistema

que toma decisiones propias? El uso de programas capaces de aprender y adaptarse multiplica el alcance de los delitos informáticos. La ciberdelincuencia autónoma es una amenaza emergente, sistemas que, una vez lanzados, pueden identificar vulnerabilidades, modificarse a sí mismos y actuar sin intervención humana continua (Miró Llinares, 2018, p. 175).

El Convenio sobre la Ciberdelincuencia, hecho en Budapest el 23 de noviembre de 2001 y la Directiva (UE) 2013/40 del Parlamento Europeo y del Consejo, de 12 de agosto de 2013, relativa a los ataques contra los sistemas de información, establecen la responsabilidad penal por ataques informáticos, pero parten del presupuesto de una acción humana directa. Cuando el agente actúa solo en la fase inicial, por ejemplo, programando un *malware* que luego actúa por sí mismo, surge la cuestión de si su responsabilidad alcanza todos los resultados posteriores o solo los previsibles al momento de la programación. En este punto cobra relevancia la noción de riesgo permitido, el sujeto puede quedar exento de responsabilidad si adoptó todas las medidas de control razonables y el resultado deriva de un comportamiento del sistema objetivamente imprevisible. Sin embargo, cuando la creación o el despliegue del sistema implica una temeraria indiferencia ante los posibles daños, podría configurarse una modalidad de dolo eventual o, al menos, una conducta imprudente grave.

La autoría mediata tradicional requiere el control de la voluntad de otro, un instrumento humano, en cambio, en el uso de IA, el control se ejerce sobre un programa. La cuestión es si puede hablarse de «instrumentalización tecnológica» análoga a la humana. Parte de la doctrina propone extender el concepto de autoría mediata al programador o usuario que diseña o utiliza la IA con fines delictivos, manteniendo el dominio funcional del proceso mediante la configuración del algoritmo y el conocimiento de sus efectos previsibles (Salvadori, 2025, p. 70). Otros autores sostienen que la imputación debe limitarse a los resultados previsibles conforme al principio de culpabilidad.

Y es aquí donde debemos tener en cuenta que el operador que utiliza un sistema de IA en la ejecución de una actividad debe mantener un deber de vigilancia y control sobre su funcionamiento. Este deber deriva tanto de la posición de garante (artículo 11 del Código Penal) como de los principios de diligencia y previsibilidad tecnológica. Así, si un profesional emplea una herramienta de IA con potencial para causar daños, por ejemplo, en el ámbito financiero, médico o de seguridad, está obligado a verificar su funcionamiento, comprender sus limitaciones y prevenir sus posibles efectos lesivos. En otras palabras, quien despliega un sistema autónomo en el mercado asume una posición de garante frente a los daños que dicho sistema pueda causar. El incumplimiento de ese deber puede dar lugar a responsabilidad penal por imprudencia o, en casos graves, por dolo eventual. En la práctica, la jurisprudencia española ha admitido imputaciones por omisión en contextos de control tecnológi-


ISSN 2306-7608 (Impresa) • ISSN 2810-8396 (En línea)

co, especialmente en materia de seguridad industrial y transporte. La extensión de estos criterios al ámbito algorítmico parece coherente con la función preventiva del Derecho penal, aunque siempre bajo el principio de proporcionalidad y con cautela frente a una expansión excesiva de la punición.

Pero sin lugar a duda, una de las cuestiones más debatidas en la doctrina contemporánea es si la inteligencia artificial podría, en un futuro, llegar a ser considerada sujeto de imputación penal. La idea de otorgar una «personalidad electrónica» a determinados sistemas altamente autónomos, sugerida en algunos documentos del Parlamento Europeo desde 2017, pretende responder al vacío existente entre la creciente autonomía de la IA y las estructuras jurídicas tradicionales basadas exclusivamente en la responsabilidad humana. No obstante, la doctrina penal española y europea se ha mostrado mayoritariamente contraria a esta posibilidad. El Derecho penal parte de la base de que solo el ser humano, dotado de conciencia, libertad y capacidad de autodeterminación, puede ser sujeto de reproche jurídico. La imputabilidad penal presupone una capacidad de comprender la ilicitud del hecho y de actuar conforme a esa comprensión. Las máquinas, por muy avanzadas que sean, no poseen intencionalidad, voluntad moral ni conciencia de sí mismas; son, en última instancia, instrumentos programados por personas. La IA no tiene libertad, ni deberes jurídicos, ni la posibilidad de ser disuadida mediante sanción. Según Silva Sánchez (2011, p. 45), la culpabilidad no puede trasladarse a entes sin conciencia moral sin vaciar de contenido el principio de humanidad del Derecho penal. Aunque, esto nos suena, en cierta manera, a un pasado no tan lejano sobre la responsabilidad penal de las personas jurídicas y la famosa frase acuñada en 1881 por Franz von Liszt de *societas delinquere non potest*. Aunque sigue siendo un tema de constante debate doctrinal, desde la reforma del Código Penal de 2010, las personas jurídicas tienen responsabilidad penal directa e independiente de la de sus administradores, directivos o miembros, como así expresa el artículo 31 bis del Código Penal. Así que, tomándonos una licencia y haciendo un cierto paralelismo simbólico podríamos preguntarnos *¿IA delinquere non potest?*

Ellamey y Elwakad (2023) proponen una solución alternativa, conceder a la IA una personalidad jurídica electrónica, de manera que sea posible que la IA asuma responsabilidad penal y civil por sus acciones, independiente de los seres humanos que la diseñaron o la utilizan. Este enfoque reconoce la autonomía de los sistemas de IA en la toma de decisiones, lo que permitiría que sean sancionados directamente por acciones ilícitas, sin necesidad de imputar la culpa exclusivamente a los operadores humanos (Loaiza Moreno; Soto Soto; Hoyos Escaleras, 2024, p. 2163).

Por lo tanto, como vemos se han explorado posibles modelos intermedios de atribución penal, especialmente, para casos de alto grado de autonomía tecnológica. En primer lugar, un modelo de responsabilidad vicaria o indirecta, por el

cual los resultados producidos por la IA se imputarían al programador o al usuario responsable del sistema. En segundo lugar, un modelo de responsabilidad funcional o estructural, que consideraría a la IA como un sistema operativo dentro de una organización, atribuyendo el resultado al ente colectivo que se beneficia de su uso. Y, en tercer lugar, un modelo de responsabilidad objetiva limitada, propuesto en algunos foros europeos, que establecería un régimen sancionador no penal, de carácter cuasi-administrativo, basado en la idea de riesgo tecnológico (Loaiza Moreno; Soto Soto; Hoyos Escaleras, 2024, p. 2167). Pero, si bien es cierto, todos estos modelos colisionan, en mayor o menor medida, con el principio de culpabilidad subjetiva y con la necesidad de motivación personal de la conducta delictiva. El Derecho penal español, en su configuración actual, no admite sanciones sin culpabilidad, ni siquiera bajo la forma de una responsabilidad objetiva encubierta. La respuesta frente a los daños producidos por IA autónomas, por el momento, deberían buscarse en otros ámbitos, como la responsabilidad civil o administrativa.

Por otro lado, el empleo de IA en la ejecución de actividades económicas, industriales o de servicios plantea la cuestión de cómo valorar el dolo o la imprudencia cuando el resultado lesivo se produce por la actuación de un sistema automatizado. En materia de dolo, debe analizarse si el sujeto conocía el riesgo inherente al funcionamiento del sistema y, aun así, decidió actuar. En cuanto a la imprudencia, la clave es la previsibilidad tecnológica, es decir, el resultado será imputable penalmente si era razonablemente previsible en atención al estado de la técnica y el sujeto no adoptó las medidas necesarias para evitarlo. Así, un programador que introduce deliberadamente un código destinado a evadir controles de seguridad podría responder por dolo eventual. En cambio, un usuario que emplea una IA comercial sin conocer defectos imprevisibles podría quedar exento de responsabilidad penal, aunque no de responsabilidad civil.

Y, en último lugar, debemos tener en cuenta que el operador humano que controla o supervisa un sistema de IA ostenta una posición de garante respecto de los riesgos que genera, conforme al artículo 11 del Código Penal. Esto significa que debe impedir la producción de resultados lesivos derivados del funcionamiento del sistema cuando tiene capacidad de control sobre él. Por tanto, si un sistema automatizado comete un daño previsible y el responsable no interviene a tiempo, puede incurrir en un delito por omisión impropia. Además, la doctrina entiende que la posición de garante surge no solo del contrato o de la ley, sino también de la creación previa de una fuente de riesgo (Lacruz López, 2015, p. 296). Quien introduce en el tráfico jurídico un sistema de IA debe asumir los deberes de supervisión y corrección derivados de ese riesgo.

No podemos olvidarnos de lo mencionado en el punto anterior, toda vez que en el Reglamento de Inteligencia Artificial impone que todo sistema de alto riesgo


ISSN 2306-7608 (Impresa) • ISSN 2810-8396 (En línea)

esté sometido a una supervisión humana efectiva. Desde el punto de vista penal, esta obligación puede considerarse un deber jurídico cuyo incumplimiento constituya la base de la imputación. El control humano no se limita a la mera posibilidad de intervenir, sino que exige comprensión suficiente del funcionamiento del sistema y capacidad real para corregir o detener su actuación. En el ámbito empresarial, la delegación ciega en decisiones automatizadas sin control experto podría ser calificada de negligencia grave, especialmente si se vulneran derechos fundamentales o se generan daños relevantes (Barrio Andrés, 2024, p. 68). Como es entendible, el Derecho penal no castiga la innovación tecnológica, pero sí la irresponsabilidad en su uso.

Debemos tener en cuenta más situaciones que son de preocupación a nivel penal. Como hemos podido ver, la expansión de la inteligencia artificial ha generado un desfase evidente entre el avance tecnológico y la capacidad del Derecho penal para responder de manera coherente y proporcionada. Las categorías clásicas de acción, culpabilidad e imputación se fundamentan sobre la noción de conducta humana, libre y consciente. Frente a sistemas que aprenden y actúan de forma autónoma, el principio de legalidad penal se enfrenta al desafío de tipificar conductas cuya materialización resulta imprevisible o cuya autoría es difusa. La respuesta jurídica debe mantener la racionalidad del sistema penal y reservar la sanción criminal a los casos de verdadera necesidad en los que el daño tecnológico no pueda ser prevenido eficazmente mediante otras ramas del Derecho, manteniendo, de esta manera, el principio de intervención mínima.

Por ello, ante esta nueva realidad, la doctrina penal se encuentra ante la necesidad de poder considerar el reinterpretar algunos conceptos sin desnaturalizarlos. En primer lugar, la noción de acción podría ser ampliada para incluir las conductas que consisten en diseñar, programar o poner en marcha sistemas capaces de generar resultados lesivos. En segundo lugar, el concepto de imputación objetiva debería incorporar la idea de riesgo tecnológico permitido, que delimite cuándo un resultado producido por IA es atribuible a su operador y cuándo pertenece al ámbito del azar o la imprevisibilidad técnica. Del mismo modo, la posición de garante adquiere una nueva relevancia para quienes introducen en el mercado o utilizan sistemas de IA de alto riesgo, teniendo que asumir un deber especial de vigilancia y corrección. Esta extensión del deber jurídico no supone una expansión punitiva injustificada, sino una adaptación de la dogmática al contexto tecnológico actual. Y, por último, la culpabilidad deberá valorarse atendiendo al nivel de conocimiento técnico exigible al sujeto conforme a su posición y a la complejidad del sistema. La falta de formación o de control sobre la tecnología no puede servir de excusa general, pero sí puede modular el grado de reprochabilidad.

## 5. DELITOS COMETIDOS CON INTELIGENCIA ARTIFICIAL: NUEVAS MODALIDADES Y REINTERPRETACIONES TÍPICAS

El despliegue de sistemas de IA está incidiendo directamente en la tipificación y ejecución de numerosos delitos previstos en el Código Penal español, tanto en su modalidad clásica como en la aparición de formas novedosas de criminalidad tecnológica. No se trata solo de nuevos tipos penales, sino también de la reinterpretación funcional de figuras tradicionales a la luz de medios comisivos automatizados. Podemos hablar de dos vertientes, por un lado, el uso malintencionado de la tecnología avanzada para manipular, suplantar, falsificar y, por otro lado, la capacidad de la inteligencia artificial para perfeccionar acciones delictivas, aumentado el espectro del cibercrimen y creando escenarios inéditos para la comisión de ilícitos (Romera Bravo, 2025, p. 46).

El siguiente catálogo no pretende ser exhaustivo, sino ilustrativo de la incidencia transversal de la IA en la criminalidad contemporánea. En primer lugar, y atendiendo a los delitos contra la intimidad, la propia imagen y el honor, vemos que la IA ha multiplicado la capacidad de vulnerar derechos personalísimos. Constituyen un claro ejemplo, los denominados *deepfakes*, contenidos falsificados mediante aprendizaje profundo. Estas tecnologías permiten generar imágenes o vídeos falsos de personas reales, frecuentemente con contenido sexual o difamatorio, lo que encaja en los tipos del artículo 197 del Código Penal: delitos de descubrimiento y revelación de secretos.

También, por otro lado, la generación automatizada de perfiles personales o la vigilancia algorítmica intensiva puede suponer una vulneración de la intimidad, especialmente cuando se realiza sin base legal o consentimiento, pudiendo dar lugar a delitos del artículo 197 ter del Código Penal, es decir, el acceso ilícito a datos personales.

En segundo lugar, respecto a los delitos contra el patrimonio y el orden socioeconómico, la IA se ha convertido en herramienta privilegiada tanto para la comisión como para la detección de fraudes. Destacamos las siguientes tipologías: las estafas automatizadas (artículo 248 Código Penal), donde la IA generativa permite crear mensajes, voces o imágenes que suplantan identidades o reproducen de forma verosímil interacciones humanas (*voice cloning, phishing* inteligente), induciendo al error a víctimas sin contacto directo con el autor. Un caso reciente fue el acontecido a la empresa inglesa de diseño e ingeniería Arup en mayo de 2024, a través del cual los criminales utilizaron videoconferencias 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 solicitaban hacer distintos depósitos y transferencias bancarias a uno de sus empleados de sus oficinas en Hong Kong. El empleado


ISSN 2306-7608 (Impresa) • ISSN 2810-8396 (En línea)

accedió a hacer las transferencias y se estima que la empresa sufrió unas pérdidas de unos 25.6 millones de dólares (Invezz, 2024).

El delito de manipulación de mercados (artículo 284 Código Penal), donde los algoritmos de alta frecuencia o *trading bots* pueden alterar de forma artificial los precios o crear falsas señales de oferta y demanda. El delito de blanqueo de capitales (artículo 301 Código Penal), donde los sistemas de IA pueden ser usados para ocultar el origen de fondos ilícitos mediante transacciones automatizadas o estructuras opacas basadas en *criptoactivos*, dificultando la trazabilidad de las operaciones (Casals Fernández, 2022, p. 430).

En cuarto lugar, revisando los delitos contra la seguridad informática la IA ha revolucionado la ciberdelincuencia, permitiendo ataques más sofisticados, adaptativos y autónomos. Podemos incluir aquí el acceso ilícito a sistemas (artículo 197 bis Código Penal) mediante algoritmos capaces de aprender vulnerabilidades. La interferencia en sistemas o datos (artículo 264 Código Penal) ejecutada por *malware* inteligente que se replica o modifica sin control humano. La particularidad penal radica en determinar si el programador conserva dominio del hecho o si la propagación del ataque escapa a su previsión, planteando cuestiones de dolo eventual e imprudencia grave.

En quinto lugar, respecto a los delitos contra la seguridad vial y la integridad física, vemos que el avance de los vehículos autónomos y de los sistemas de asistencia basados en IA plantea nuevos supuestos de responsabilidad penal en materia de homicidios y lesiones imprudentes (artículos 142 y 152 Código Penal). Cuando un vehículo opera parcialmente de forma automatizada, la atribución de responsabilidad dependerá del grado de control humano exigido. Si el sistema comete un error previsible y el conductor o fabricante omitió las medidas de control o actualización, podría apreciarse imprudencia profesional. La falta de protocolos normativos claros para delimitar la responsabilidad entre conductor, fabricante y software intensifica la necesidad de un marco de responsabilidad compartida.

En sexto lugar, atendiendo a los delitos contra la Administración Pública y la Justicia, donde ya hemos visto que el uso de IA por parte de las Administraciones plantea riesgos específicos. La manipulación o sesgo de algoritmos empleados en procesos selectivos, concesiones o sanciones podría configurar delitos de prevaricación (artículo 404 Código Penal) o falsedad documental (artículo 390 Código Penal) si la decisión automatizada vulnera el ordenamiento jurídico. Asimismo, la incorporación de IA en tareas judiciales o policiales sin garantías suficientes podría dar lugar a delitos de revelación de secretos o de abuso de autoridad si se accede indebidamente a datos sensibles o se utilizan sin base legal.

En séptimo lugar, en cuanto a los delitos contra la libertad sexual, la generación y difusión de material pornográfico sintético sin consentimiento, el llamado *deepfake porn*, encaja en los tipos del artículo 183 ter CP y 197.7 CP, ampliados por la Ley Orgánica 10/2022, de 6 de septiembre, de garantía integral de la libertad sexual. La IA agrava el daño reputacional y psicológico, al permitir replicar y distribuir falsificaciones hiperrealistas, lo que justifica su consideración como modalidad cualificada o, al menos, circunstancia agravante derivada del medio comisivo tecnológico.

Y, en octavo lugar, en cuanto a los delitos contra la seguridad nacional y el orden público, somos conocedores de que los sistemas de IA pueden emplearse en campañas de desinformación, manipulación electoral o sabotaje cibernético, susceptibles de encuadrarse en los artículos 544 y siguiente del Código Penal (atentados contra el orden público) o 580 del Código Penal (colaboración con organizaciones terroristas).

La Estrategia de Seguridad Nacional 2023 advirtió sobre el riesgo de la utilización de IA, mencionando que «[…] se ha extendido el uso de la inteligencia artificial gracias a la irrupción de herramientas online que ponen esta tecnología a disposición de los ciudadanos de manera gratuita. La IA puede ser utilizada para generar contenido falso, fraudes o desinformación de manera rápida y convincente. Se pueden crear noticias falsas, artículos o videos que aparenten ser reales, lo que dificulta que las personas distingan entre información verdadera y falsa. Esto ha conllevado, además de las ventajas de su uso responsable, un aumento de denuncias relacionadas con su uso malicioso incluso derivando en casos de *sextorsión* mediante la generación de vídeos o fotografías falsos».

En conclusión, la inteligencia artificial no solo crea nuevos espacios de criminalidad, sino que reconfigura los límites típicos de delitos ya existentes. La clave reside en interpretar el Código Penal con criterio funcional y flexible, preservando los principios de legalidad y culpabilidad, pero evitando zonas de impunidad derivadas de la complejidad tecnológica.

## 6. CONCLUSIONES

La IA ha irrumpido en la sociedad y, especialmente en el ámbito penal, como un fenómeno estructural que trasciende el mero cambio tecnológico para interpelar directamente los fundamentos dogmáticos del Derecho penal clásico. Se ha puesto de manifiesto que la autonomía técnica, la opacidad algorítmica y la imprevisibilidad de los sistemas inteligentes desestabilizan las categorías tradicionales de acción, culpabilidad e imputación. Esta constatación no es únicamente teórica: los casos recientes de *deepfakes,* fraudes automatizados o manipulación de datos judiciales


ISSN 2306-7608 (Impresa) • ISSN 2810-8396 (En línea)

ilustran una realidad en la que el elemento humano se diluye progresivamente tras la capa algorítmica. El resultado es una tensión estructural entre la expansión de los riesgos tecnológicos y la capacidad del sistema penal para responder con coherencia normativa y garantías.

Desde una perspectiva dogmática, el primer gran desafío radica en la redefinición del concepto de acción penalmente relevante. La conducta tradicionalmente entendida como manifestación de voluntad humana encuentra en la IA un intermediario que puede alterar el curso causal sin intervención consciente del sujeto. En este contexto, la noción de acción debería ampliarse para abarcar las conductas consistentes en diseñar, programar, desplegar o supervisar sistemas capaces de generar resultados típicos. No se trata de atribuir personalidad penal a la máquina, sino de reconocer que el comportamiento penalmente relevante puede situarse en fases previas, como la configuración, el entrenamiento o la puesta en marcha, que generan un riesgo jurídicamente desaprobado.

En segundo término, la imputación objetiva demanda una reformulación en clave tecnológica. El operador que introduce un sistema de IA en el tráfico jurídico crea un riesgo que puede materializarse sin intervención ulterior. De ahí que la categoría de riesgo permitido deba redefinirse para distinguir entre el desarrollo responsable y el uso temerario de la tecnología. En la práctica, el Derecho penal deberá valorar si el sujeto adoptó las medidas de control y supervisión exigibles según el estado de la técnica, integrando un criterio de diligencia tecnológica análogo al de la diligencia profesional en sectores regulados. La infracción de deberes de vigilancia sobre sistemas autónomos podría fundamentar la imputación por imprudencia grave o dolo eventual, siempre bajo un juicio de previsibilidad razonable. Esta evolución no implica una expansión punitiva, sino una adecuación del tipo de injusto a los riesgos de la sociedad algorítmica.

Un tercer eje de análisis se centra en la culpabilidad. El principio de responsabilidad personal, inseparable de la capacidad de autodeterminación consciente, se ve amenazado por la complejidad técnica y la asimetría de conocimiento entre desarrolladores, usuarios y víctimas. La culpabilidad en el entorno de la IA debe valorarse conforme al nivel de conocimiento técnico exigible al sujeto según su rol, ya sea programador, operador o entidad beneficiaria, y no según la mera previsión abstracta del daño. Ello conduce a una graduación funcional de la culpabilidad, donde la posición de garante y el grado de control efectivo sobre el sistema determinan la intensidad del reproche. La ignorancia técnica no exime, pero modula la imputación, por lo tanto, la ausencia de formación o de protocolos de supervisión puede revelar una imprudencia profesional más que un dolo directo.

Desde una dimensión institucional, evidenciamos la insuficiencia del marco normativo penal vigente. El Código Penal español carece de tipos específicos para

conductas cometidas mediante IA, obligando a subsumirlas en figuras tradicionales como revelación de secretos, falsificación documental, estafa, suplantación de identidad, que no contemplan la autonomía operativa ni la escala de daño generada por algoritmos. Este vacío no se resuelve con una mera tipificación simbólica, sino con una revisión integral de los elementos del tipo penal, la autoría y la participación. Se impone la necesidad de integrar criterios de imputación por riesgo tecnológico en supuestos en los que la conducta humana inicial de diseño, despliegue o falta de control, sea la fuente de un resultado lesivo automatizado.

En cuanto al Reglamento de Inteligencia Artificial, aunque eminentemente administrativo, proyecta efectos indirectos sobre el ámbito penal. Su clasificación de sistemas por niveles de riesgo y su enumeración de prácticas prohibidas constituyen parámetros normativos para delimitar el riesgo permitido y, por extensión, la imputación penal. Las infracciones graves de los principios de transparencia, trazabilidad o supervisión humana, cuando generan daños a bienes jurídicos protegidos, pueden considerarse indicios de negligencia o dolo eventual. De este modo, el Derecho penal puede apoyarse en estándares europeos de gobernanza algorítmica sin necesidad de crear tipos autónomos desconectados del sistema general. No obstante, esta integración requiere coherencia interpretativa y coordinación entre las jurisdicciones penal y administrativa, evitando duplicidades o vulneraciones del principio *ne bis in idem.*

En el terreno de la prueba penal, la IA introduce una problemática específica: la autenticidad y fiabilidad de los contenidos digitales. La proliferación de *deepfakes* y la manipulación algorítmica de evidencias comprometen el principio de contradicción y la presunción de inocencia. La doctrina y la jurisprudencia deberán reforzar los estándares de trazabilidad probatoria, exigiendo la verificabilidad de los procesos automatizados y la posibilidad de auditoría independiente de los algoritmos empleados. El principio de explicabilidad, derivado del Derecho europeo, adquiere aquí valor constitucional como garantía de defensa. Ningún resultado derivado de un sistema opaco puede servir de base exclusiva para una condena penal sin vulnerar el artículo 24 de la Constitución Española.

Asimismo, la incorporación de sistemas predictivos o de análisis de riesgo en la administración de justicia plantea una tensión entre eficiencia y garantías. La tentación de sustituir el juicio humano por la valoración algorítmica amenaza los pilares de la culpabilidad personal y de la motivación judicial. El uso de la IA en la fase de investigación debe quedar estrictamente sometido a control jurisdiccional y a criterios de transparencia verificable. Debemos subrayar que la confianza ciega en la neutralidad tecnológica, el *automation bias*, es incompatible con la legitimidad del poder punitivo.


ISSN 2306-7608 (Impresa) • ISSN 2810-8396 (En línea)

Desde un enfoque crítico-propositivo, el artículo permite extraer varias líneas de reforma y/o reflexión. En primer lugar, se considera una revisión del Código Penal orientada a incorporar una cláusula general sobre responsabilidad penal por riesgo tecnológico, siguiendo modelos ya presentes en el ámbito medioambiental o de seguridad industrial. Dicha cláusula no supondría penalizar la innovación, sino establecer deberes jurídicos claros de supervisión, auditoría y prevención para quienes desarrollan o utilizan IA de alto riesgo. En segundo lugar, resulta urgente armonizar el derecho penal con el derecho administrativo tecnológico, de modo que los incumplimientos graves de obligaciones de transparencia o control previstos en el Reglamento de Inteligencia Artificial puedan constituir indicios relevantes de imprudencia penal.

Consideramos que debe promoverse la creación de programas de *compliance* algorítmico en el ámbito empresarial y público. La cultura de cumplimiento ya consolidada en materia de corrupción o medioambiente debe extenderse al uso ético y seguro de la IA. Estos programas no solo cumplen una función preventiva, sino también probatoria, al acreditar la diligencia debida en caso de incidente tecnológico. La futura acción penal frente a delitos cometidos con IA debería considerar la existencia o no de estos mecanismos como criterio de graduación de la culpabilidad corporativa.

Una línea de actuación que consideramos también relevante es la formación especializada de los operadores jurídicos. Los jueces y fiscales deberían adquirir competencias técnicas mínimas para evaluar el funcionamiento y los riesgos de los sistemas automatizados, sin olvidarnos de la importancia de los peritos, como ayuda imprescindible para la intelección técnica de la IA. Sin comprensión tecnológica, la tutela judicial efectiva se vacía de contenido. De igual modo, las periciales informáticas deben someterse a estándares unificados de auditoría y certificación, garantizando independencia y reproducibilidad de los resultados.

Por último, el debate sobre la personalidad jurídica electrónica de la IA, aunque conceptualmente estimulante, debe abordarse con cautela. La atribución de responsabilidad penal a sistemas sin conciencia moral colisiona con los principios de humanidad y culpabilidad. En cambio, podrían explorarse mecanismos cuasi-penales de responsabilidad objetiva limitada, de naturaleza sancionadora o administrativa, orientados a reparar los daños y prevenir la reiteración, sin desnaturalizar el núcleo ético del Derecho penal. La expansión del castigo sin culpa sería, en definitiva, más peligrosa para el Estado de Derecho que la propia impunidad tecnológica que pretende corregir.

En el plano internacional, el uso transfronterizo de la IA impone la necesidad de reforzar la cooperación judicial digital y los mecanismos de intercambio de información forense. Los delitos cometidos mediante algoritmos operativos en

múltiples jurisdicciones cuestionan la territorialidad clásica de la competencia penal. La Unión Europea debería avanzar hacia un marco común de persecución de la criminalidad algorítmica, apoyado en normas uniformes de conservación de pruebas digitales y en la armonización de criterios de imputación. España, como Estado miembro, tendría así un punto de partida sólido para adaptar su legislación sin fragmentar el espacio penal europeo.

En síntesis, la IA actúa como un catalizador de transformación del Derecho penal. No exige una ruptura con sus principios, sino una reinterpretación evolutiva. La acción, la imputación y la culpabilidad deben reconfigurarse a la luz del riesgo tecnológico, manteniendo siempre la proporcionalidad y la tutela de los bienes jurídicos. El Derecho penal seguirá siendo la *ultima ratio*, pero su eficacia dependerá de su capacidad para entender el nuevo contexto digital sin renunciar a su función garantista.

La conclusión más profunda que deja el trabajo es que el verdadero desafío no reside en castigar a la máquina, sino en preservar la humanidad del castigo. La IA no tiene voluntad ni conciencia, pero quienes la crean y la utilizan sí. En ese tránsito entre autonomía técnica y responsabilidad humana se juega la legitimidad del sistema penal del siglo XXI. El Derecho penal del futuro deberá ser, más que nunca, racional y tecnológicamente ilustrado, capaz de sancionar la imprudencia digital sin criminalizar la innovación, es decir, proteger la dignidad humana frente a la despersonalización algorítmica. Solo así la justicia podrá seguir siendo humana en una era cada vez, aparentemente, menos humana.

## BIBLIOGRAFÍA

Alkorta Idiakez, I. (2025). *La regulación de los productos sanitarios con Inteligencia Artificial*. Tirant lo Blanch.

Barrio Andrés, M. (2024). Objeto, ámbito de aplicación y sentido del Reglamento Europeo de Inteligencia Artificial. En: Barrio Andrés, M. (Dir.), *El Reglamento Europeo de Inteligencia Artificial.* Tirant lo Blanch.

Casals Fernández, A. (2022). Las criptomonedas frente al delito de blanqueo de capitales y la complejidad de la prueba pericial en el ámbito ciberdelincuente. *Anuario de Derecho Penal y Ciencias Penales*, *LXXV*(1), 421-446. https://doi.org/10.53054/adpcp.v75i1.9697.

Colina Ramírez, E. I. (2023). Inteligencia Artificial ¿otro cambio de paradigma en el derecho penal? *Cuadernos de política criminal,* (141), 123-150. https://doi.org/10.14679/2718.


ISSN 2306-7608 (Impresa) • ISSN 2810-8396 (En línea)

Cerdio Herrán, J. A. (2025). *IA para el derecho o del poder del modelado explícito en la era de la IA opaca*. Tirant lo Blanch.

Cinelli, V. / Manrique Gan, A. (2019). El uso de programas de análisis predictivo en la inteligencia policial: una comparativa europea. *Revista de Estudios en Seguridad Internacional*, *5*(2), 1-19. https://doi.org/10.18847/1.10.1.

Cuadrado Salinas, C. (2025). *Inteligencia artificial en la justicia penal. Desafíos, oportunidades y límites. Entre la eficacia y la garantía de los derechos fundamentales*. Tirant lo Blanch.

Freile Mansilla, R. (2025). Prueba electrónica, informe pericial informático e inteligencia artificial. En: González-Torre, A.P. / Tarodo Soria, S. (Dirs.), *Desafíos regulatorios de la inteligencia artificial*. Tirant lo Blanch.

Ellamey, Y. / Elwakad, A. (2023). The criminal responsibility of artificial intelligence systems: A prospective analytical study. *Corporate Law* & *Governance Review, 5*(1), 92-100. https://doi.org/10.22495/clgrv5i1p8.

García Mendiola, M. G. (2025). Aspectos jurídicos de la inteligencia artificial generativa: un nuevo paradigma. En: González-Torre, A. P. / Tarodo Soria, S. (Dirs.), *Desafíos regulatorios de la inteligencia artificial*. Tirant lo Blanch.

Invezz (17 de mayo de 2024). *Deepfake: la empresa británica Arup es víctima de una estafa de 25 millones de dólares, ¿cómo puede protegerse?* https://es.tradingview.com/news/invezz:d3919989709cd:0/.

Lacruz López, J. M. (2015). El delito como conducta típica y IV: Los tipos de los injusto de los delitos de omisión. En: Gil Gil, A. / Lacruz López, J. M. / Melendo Pardos, M.; Núñez Fernández, J., *Curso de Derecho Penal Parte General*. Dykinson.

Lang Irrazábal, M.C. (2022). La inteligencia artificial en la Administración de Justicia. en *Ars Iuris Salmanticensis, Tribuna de Actualidad*, *10*(2), diciembre, 31-39. https://doi.org/10.14201/AIS20221023139.

Llanas, S. (27 de febrero de 2025). *Investigados penalmente cuatro menores por compartir imágenes de contenido sexual producidas con IA en un instituto de Barcelona*. Diario El País https://elpais.com/espana/catalunya/2025-02-27/investigados-penalmente-cuatro-menores-por-manipular-y-compartir-imagenes-de-contenido-sexual-producidas-con-ia-en-un-instituto-de-barcelona.html?utm_source=chatgpt.com

Loaiza Moreno, J. D. / Soto Soto, F. F. / Hoyos Escaleras, A. M. (2024). Revolucionando la justicia: el impacto de la inteligencia artificial en el derecho

penal. *Estudios y Perspectiva Revista Científica y Académica*, *4*(3). https://doi.org/10.61384/r.c.a..v4i3.537.

Martino, A. (1992): *Expert system in law.* North-Holland.

Mínguez Rosique, M. / Gallego Arribas, D. (Coords.) (2025). *Cibercrimen: Tendencias y desafíos actuales*. Boletín Oficial del Estado.

Miró Llinares, F. (2018). Apuntes sobre Derecho penal e Inteligencia Artificial. En: Morales Prats, F. / Tamarit Sumalla, J. M. / García Albero, R., *Represión Penal y Estado de Derecho. Homenaje al Profesor Gonzalo Quintero Olivares*. Thomson Reuters

Romera Bravo, F.J. (2025). El impacto de la inteligencia artificial en el derecho penal: delitos emergentes y marco jurídico europeo. *Revista Jurídica Conede*, (3), 41-50. https://dialnet.unirioja.es/servlet/articulo?codigo=10064915.

Salvadori, I. (2025). Cibercrimen y uso ilícito de la inteligencia artificial: retos y desafíos del Derecho penal. En: Mínguez Rosique, M. / Gallego Arribas, D. (Coords.): *Cibercrimen: Tendencias y desafíos actuales*. Boletín Oficial del Estado.

Ser Toledo (20 de octubre de 2025). *Detenido un joven de 28 años por crear y distribuir imágenes pornográficas de 26 mujeres reales con Inteligencia Artificial*. https://cadenaser.com/castillalamancha/2025/10/20/detenido-un-joven-de-28-anos- por-crear-imagenes-pornograficas-de-26-mujeres-reales-con-inteligencia-artificial-ser-toledo/?utm_source=chatgpt.com

Serrano Ferrer, M. P. (2021). *Derecho Penal y nuevas tecnologías*. Thomson Reuters-Aranzadi.

Silva Sánchez, J. M. (2011). *La expansión del Derecho penal: aspectos de la política criminal en las sociedades postindustriales*. (3º ed.). Edisofer.

Valls Prieto, J. (2022). Sobre la responsabilidad penal por la utilización de sistemas inteligente. *Revista Electrónica de Ciencia Penal y Criminología*, (24-27). https://revistacriminologia.com/24/recpc24-27.pdf.

Varona Gómez, D. (2024): Algoritmos e inteligencia artificial en el sistema de justicia penal. *InDret. Revista para el análisis del derecho*, (4). https://indret.com/algoritmos-e-inteligencia-artificial-en-el-sistema-de-justicia-penal/.


ISSN 2306-7608 (Impresa) • ISSN 2810-8396 (En línea)