

DETERMINACIÓN DE LA RESPONSABILIDAD PENAL POR HECHOS REALIZADOS POR LA INTELIGENCIA ARTIFICIAL¹

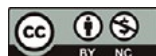
ESTABLISHMENT OF CRIMINAL LIABILITY FOR FACTS COMMITTED BY ARTIFICIAL INTELLIGENCE

■ **DR. IGNACIO FRANCISCO BENÍTEZ ORTÚZAR**

Catedrático de Derecho penal, Universidad de Jaén, España²

<https://orcid.org/0000-0002-4955-9409>

ibenitez@ujaen.es



Resumen

Los desafíos que plantea la determinación de la responsabilidad penal por hechos lesivos cometidos por sistemas de inteligencia artificial son objeto de análisis en el presente artículo. A partir de la distinción entre una dimensión tecnológica débil: previsible y controlable por humanos, y una fuerte, con aprendizaje autónomo, se examina la insuficiencia de los modelos tradicionales de imputación penal, como la autoría mediata o la responsabilidad por el producto defectuoso, para resolver los problemas que

- 1 Investigación realizada en el Proyecto PID2022-138770OB-11 «Derecho penal y distribución de la riqueza en la sociedad tecnológica (II)», financiado por MICIU/AEI/10.13039/501100011033 y por FEDER, UE.
- 2 El autor es doctor en Derecho (Universidad de Granada), licenciado en Derecho (Universidad de Granada), licenciado en Criminología (Universidad de Murcia) y experto en Criminología (Universidad de Granada); presidente del Tribunal Administrativo del Deporte de Andalucía (2019-actualidad); investigador responsable del grupo de Investigación SEJ 428, «Derecho penal, criminología, democracia y derechos fundamentales»; coordinador del Máster Oficial en Justicia Penal y Sistema Penitenciario, al igual que del Programa de Doctorado en Derecho de la Universidad de Jaén; presidente y socio-fundador de UNIXA CRIDES S. L., Spin off UJA; miembro de honor de la Sociedad Cubana de Ciencias Penales y Profesor Distinguido de la Facultad de Derecho de la Universidad de la Habana (Cuba).

plantea la inteligencia artificial. El texto explora las propuestas de la Unión Europea, centradas en el control humano y la gestión de riesgos; debate acerca de la posible creación de una *personalidad electrónica* para sistemas autónomos y concluye que, si bien el marco actual prioriza la supervisión humana, la evolución hacia IA más autónomas podría exigir una reconceptualización de la teoría del delito y la responsabilidad penal.

Palabras clave: Responsabilidad penal; inteligencia artificial; imputación objetiva.

Abstract

This article analyzes the challenges inherent in determining criminal liability for harmful acts committed by artificial intelligence systems. Proceeding from the distinction between weak artificial intelligence: predictable and controlled by human, and strong artificial intelligence, characterized by autonomous learning, it examines the inadequacy of traditional models of criminal imputation, such as mediate perpetration or product liability, to resolve the issues posed by artificial intelligence. The text explores European Union proposals, centered on human control and risk management; debates the potential creation of an electronic personhood for autonomous systems; and concludes that, while the current framework prioritizes human supervision, the evolution towards more autonomous AI may necessitate a reconceptualization of the theory of crime and criminal liability.

Keywords: Criminal Liability; Artificial Intelligence; Objective Imputation.

Sumario

I. Introducción; II. Encaje del concepto de delito en el mundo digital. El ciberdelito; III. La responsabilidad por la lesión de bienes jurídicos a consecuencia de actuaciones de la inteligencia artificial; 3.1. Inteligencia artificial débil: Previsible y controlable; 3.2. Inteligencia artificial fuerte: Aprendizaje autónomo; IV. Conclusiones; V. Referencias.

I. INTRODUCCIÓN

El término inteligencia artificial (IA) fue acuñado en 1955 por John McCarthy, un profesor de matemáticas de Dartmouth que propuso la realización de un proyecto de investigación en el que participarían 10 destacados científicos para intentar averiguar cómo podían conseguir que las máquinas utilizaran el lenguaje, formasen abstracciones y conceptos, y resolvieran problemas que hasta ese momento solo estaban al alcance del ser humano, hasta superar a este (McCarthy *et al.*, s.f., s.p.).

El 25 de abril de 2018 se publicó la Comunicación de la Comisión Europea sobre IA para Europa, que estima que dicha expresión se aplica a los sistemas que manifiestan un comportamiento inteligente, pues son capaces de analizar su entorno y pasar a la acción —con cierto grado de autonomía— con el fin de alcanzar objetivos específicos (p. 1).

Recientemente, el Reglamento (UE) 2024/1689, del Parlamento europeo y del Consejo, de 13 de junio de 2024, por el que se establecen las normas en materia de IA y se modifican los reglamentos (CE) n.º 300/2008, (UE) n.º 167/2013, (UE) n.º 2018/858, (UE) 2018/1139, y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 Y (UE) 2020/1828 (Reglamento de IA), en adelante *Ley de IA de la Unión Europea*, dispone en su artículo 3.1, que se entenderá por

sistema de IA un sistema basado en una máquina que está diseñado [sic] 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. (2024, p. 46)

La creación de artificios que llegan a actuar con cierta independencia del ser humano también va a influir con fuerza en el ámbito propio de la criminalidad. La era de la informática y las tecnologías de la información y la comunicación (TIC), además de facilitar el desarrollo de la vida social e individual del ser humano, permitieron la aparición de nuevas formas de realización de la actividad delictiva. Los aspectos relacionados con la telemática, la irrupción de internet como

una red de redes de alcance auténticamente global, hicieron posible, también, una inmediatez en el resultado, unido a la posibilidad de la distancia en la acción; circunstancias que, en el marco de la actividad criminal y de la eficacia del Derecho punitivo de los Estados, ya advirtieron la aparición de verdaderos problemas en la persecución del delito, si se partía de las previsiones jurídicas tradicionales. Esta realidad se hace más palpable con la evolución de la tecnología y la implementación del *big data* y el mundo de los algoritmos en *aparatos* a los que se les dota de la llamada IA, con una capacidad de aprendizaje cada vez más autónomo con respecto a la actuación de sus programadores, creadores y/o de sus usuarios. La gran capacidad de almacenamiento de datos unida a la velocidad de su procesamiento la hace, incluso, una tarea imposible de realizar por la mente humana (Herrera, 2022, p. 21).

En cualquier caso, el debate acerca de la irrupción de la IA en la sociedad actual no difiere mucho de otras «revoluciones» tecnológicas o científicas a las que se ha enfrentado la historia de la humanidad, con incidencia directa en los ordenamientos jurídicos, tanto internos como internacionales. Al respecto, estos siempre «se han decantado por enfoques preventivos mediante diversas técnicas de control y supervisión que no entorpezca [sic] el avance del conocimiento científico y tecnológico cuando no sea realmente necesario», para dejar, así, al Derecho penal «en el segundo plano que siempre se le ha conferido, esto es, el de última *ratio* y de subsidiariedad» (Romeo, 2022a, p. 78).

De inicio, surgen cuestiones que el legislador debe resolver con cierta urgencia; por un lado, ¿deben responder penalmente los seres humanos que, de alguna forma, están a cargo de la supervisión y el control de los sistemas inteligentes?, ¿los sistemas inteligentes dotados de autonomía cumplen con los elementos de la teoría jurídica del delito para ser sujetos de imputación penal?, o, por el contrario, ¿basta con la determinación de la responsabilidad civil por los daños y perjuicios de los sistemas dotados de IA? (Romeo, 2022b, p. 8).

Dos son las cuestiones a plantear: 1. El encaje en el concepto del delito de las conductas llevadas a cabo con las TIC, incluida la IA; y 2. La responsabilidad por la lesión de bienes jurídicos a consecuencia de la IA.

II. ENCAJE DEL CONCEPTO DE DELITO EN EL MUNDO DIGITAL: EL CIBERDELITO

La actividad fraudulenta o criminógena relacionada con la telemática, por lo general, no da lugar a la creación de bienes jurídicos de nuevo cuño. Al contrario, en la mayoría de los casos, por no decir en todos, el bien jurídico tutelado que se ve lesionado o puesto en peligro como consecuencia de la aplicación de los desarrollos tecnológicos en un bien jurídico tradicional, cuya tutela ya es objeto de protección por el Derecho penal (privacidad, igualdad/no discriminación, dignidad, libertad, libertad de expresión, vida, integridad...) (Valls, 2021, p. 85). Las nuevas formas de lesión o puesta en peligro de esos valores fundamentales para la sociedad y el individuo se ven facilitados por las innovaciones tecnológicas y la IA.

Asimismo, no debe obviarse otra particularidad a la que da lugar la inmediatez del resultado y la distancia en la actividad, con relación a la investigación penal de este tipo de delitos, que permiten actuar en cualquier lugar del mundo desde distancias remotas, lo cual obliga a un esfuerzo de los Estados en la cooperación policial y judicial para perseguirlos, pues sin ello, el mejor texto punitivo material sería completamente inaplicable.

Se han acuñado —ya de forma consolidada a lo largo de las últimas décadas— los términos *ciberdelito*, *ciberdelincuencia* y *ciberseguridad* (Benítez, 2008, pp. 112-113; 2021, pp. 79-128), los cuales comprenden aquellos delitos que se cometen, en particular, por medio de las TIC, incluida la IA. La actividad realizada con las TIC facilita la realización de tareas que, hasta hace poco tiempo, eran exclusivas de los seres humanos; ahora, «los sistemas electrónicos tienen la capacidad de reproducir o simular las formas de trabajo propias de la mente humana» (Castillo, 2001, p. 35), con eficacia en distintos sectores de la actividad cotidiana y, también, de la criminal.

El «Convenio sobre ciberdelincuencia», del Consejo de Europa —firmado en Budapest, el 23 de noviembre de 2001—, entre los ciberdelitos, separa los *delitos contra la confidencialidad, la integridad y la disponibilidad de los datos y sistemas informáticos* (como parte de los cuales considera solo el acceso ilícito, la interceptación ilícita, la interferencia en el sistema y el abuso de dispositivos), de los *delitos informáticos* (que abarcan, entre otros, la falsificación informática y

el fraude informático). Así, mantiene una relación de género a especie, en la que los delitos informáticos son apenas una parte de los ciberdelitos.

Desde semejante planteamiento, en España, la «Estrategia nacional de ciberseguridad» (2019), ya incluía una definición de *cibercriminalidad*, en el sentido de que:

El término cibercriminalidad hace referencia al conjunto de actividades ilícitas cometidas en el ciberespacio que tienen por objeto los elementos, sistemas informáticos o cualesquiera [sic] otros bienes jurídicos, siempre que en su planificación, desarrollo y ejecución resulte determinante la utilización de herramientas tecnológicas; en función de la naturaleza del hecho punible en sí, de la autoría, de su motivación, o de los daños infligidos, se podrá hablar así de ciberterrorismo, de ciberdelito, o en su caso de *hacktivismo*. (pp. 43437-43455)

Desde un punto de vista sistemático podrían distinguirse, incluso, dos modalidades de ciberdelitos, agrupados en los de resultado y los de medio; el primer caso se refiere a las conductas que vulneran los sistemas que utilizan tecnologías de la información; y los segundos, a las conductas que utilizan las novedades tecnológicas cual vehículo para atentar contra la propiedad, la intimidad, la privacidad o la indemnidad sexual, entre otros bienes jurídicos. En esta línea se pronunciaba, también, la «Estrategia nacional de ciberseguridad» (2019), cuando afirmaba:

son tres los ámbitos en los que se desenvuelve la lucha contra la cibercriminalidad: I) el ciberespacio como objetivo directo de los hechos delictivos, o de la amenaza; II) el ciberespacio como medio clave para su comisión; y III) el ciberespacio como medio u objeto directo de investigación en cualquier tipo de hecho delictivo. (pp. 43437-43455)

Este tercer ámbito es instrumental en la investigación criminal. Sin embargo, en el ámbito jurídico-penal tal clasificación no es operativa, en tanto que el ataque a los sistemas informáticos, en sí mismos, alcanza relevancia penal porque estos contienen una información relevante de personas, empresas, etc., y los tipos penales se incluyen en los respectivos títulos y capítulos del Libro segundo del Código penal. Tratar de incluir todas las posibilidades delictivas imaginables que puedan cometerse por medio de las TIC, incluida la IA, en un

único Título del Código penal, dedicado a «los ciberdelitos», en aras de una mayor organización y sistematización de los tipos penales, conduciría, sencillamente, al fracaso legislativo.

No obstante, a lo largo del articulado del Código penal se encuentran diseminados múltiples tipos penales cuya comisión puede ser incluida en el término coloquial de *ciberdelincuencia* o *cibercriminalidad*, por lo cual, en la descripción de la conducta típica se hace expresa referencia a esta modalidad delictiva, cuando con la redacción tradicional no estaría contemplada (como ocurre, por ejemplo, con la dificultad de incluir las *Deepfakes* o imágenes creadas con IA entre los tradicionales delitos contra la intimidad o en los llamados daños informáticos, o en determinados delitos contra la indemnidad sexual de los menores de edad, perpetrados con IA).

En el ámbito operativo, con el objetivo de mejorar la lucha contra este tipo de criminalidad y lograr una mejor coordinación, la Fiscalía General del Estado en España, en la primera década del siglo xxi, creó un servicio específico de Criminalidad Informática —Instrucción 2, de 11 de octubre de 2011—, que distinguió tres grupos de delitos en la materia: a) Aquellos en los que el objeto de la actividad delictiva son los sistemas informáticos o las TIC; b) La actividad criminal que se sirve, para su ejecución, de las ventajas que ofrecen las TIC; y, c) Los que entrañan especial complejidad en la investigación, que la hace demandar conocimientos específicos en la materia.

Ahora bien, sobre todo para articular la investigación de estas conductas delictivas y perseguirlas con mínimas garantías de éxito, ante la globalización que implica el uso de internet y las tecnologías de la era informática, con la irrupción del *big data* y la IA, también, en el mundo de la delincuencia es necesario un consenso internacional que, de un lado, permita armonizar el Derecho penal sustantivo en la materia y, del otro, aporte instrumentos útiles y, a la vez, garantistas para agilizar la cooperación policial y judicial entre los Estados. En esta línea, el mayor consenso existente sobre este tipo de delincuencia lo conforman el «Convenio sobre ciberdelincuencia» del Consejo de Europa, de 23 de noviembre de 2001, y un protocolo adicional a este, de 2003 —«Protocolo adicional al Convenio sobre la ciberdelincuencia relativo a la penalización de actos de índole racista y xenófoba cometidos por medio de sistemas informáticos» (Consejo de Europa, 2003, s.p.).

En España, la «Estrategia de seguridad nacional» (2021) afirma:

Tecnologías como la Inteligencia Artificial y el *big data* subyacen cada vez más en ámbitos como el sanitario, el de transportes, el energético, el empresarial, el financiero, el educativo y el militar. La capacidad de procesamiento de grandes cantidades de datos se presenta como una característica avanzada para la consecución de los objetivos deseados. Su potencial de transformación y aplicación en procesos de análisis de riesgos y de alerta temprana es cada vez mayor. Pero el desarrollo de estas tecnologías también genera interrogantes relacionados con la seguridad. La aplicación de algoritmos para la toma automática de decisiones requiere un marco de protección de la privacidad y la no-discriminación. El empleo de sistemas autónomos también tiene implicaciones éticas que requieren mecanismos de control y parámetros que garanticen el respeto a los derechos humanos. En el medio-largo plazo, el salto tecnológico que supone la computación cuántica permitirá usos difíciles de prever hoy en día en materia de comunicaciones seguras, cifrado y descifrado y sistemas de vigilancia avanzados, entre otros. (pp. 167795-167830)

La importancia que se otorga a la ciberseguridad en la «Estrategia de seguridad nacional» (2021) es patente, al afirmarse:

En términos de ciberseguridad, se requiere garantizar el uso seguro y fiable del ciberespacio, para proteger los derechos y las libertades de los ciudadanos y promover el progreso socioeconómico. Para ello es importante incrementar las capacidades (tecnológicas, humanas y económicas) de la ciberseguridad nacional dirigidas a la prevención, detección, respuesta, recuperación, investigación y defensa activa. (p. 167821)

Para ello y en la Línea de Acción 17, considerando el ciberespacio como uno de los espacios comunes globales, junto al marítimo, aéreo y ultraterrestre, se promueve el avance en la integración del modelo de gobernanza de la ciberseguridad en el marco del Sistema de Seguridad Nacional (p. 167822). La Ley de IA de la Unión Europea (2024), que propone el establecimiento de una IA centrada en el ser humano y fiable (Artículo 1) ni recoge nuevos delitos en la materia, ni reconoce a la IA una capacidad jurídica autónoma. El texto legis-

lativo europeo establece la autorización de la IA en los Estados de la Unión en atención al nivel de riesgo, entendido, expresamente, como «la combinación de la probabilidad de que se produzca un perjuicio y la gravedad de dicho perjuicio» (Artículo 3.2). La disposición exige un sistema de control sobre los sistemas de IA de alto riesgo³ y establece las obligaciones de los proveedores, importadores y distribuidores de estos sistemas, y sus responsables, a lo largo de la cadena de valor de la IA, a la vez que obliga a los Estados a imponer un sistema de sanciones (Artículo 99 y siguientes).

En todo caso, el Informe sobre la cibercriminalidad en España, de 2023 (Gobierno de España), tomando datos del Sistema Estadístico de Criminalidad, sitúa la cibercriminalidad⁴ en un 19,2 %, respecto del conjunto de la criminalidad, lo que supone una línea en constante crecimiento desde el año 2016, cuando se ubicaba en el 4,6 %. De esta cifra, el 90,5 % corresponde a fraudes informáticos (estafas) y el 3,7 % a amenazas y coacciones.

La tabla siguiente muestra el comportamiento de las variables anteriores, según los datos de los Estudios sobre cibercriminalidad en España, de 2016 a 2023, publicados por el Ministerio del Interior:

- 3 Artículo 9.2. El sistema de gestión de riesgos se entenderá como un proceso iterativo continuo planificado y ejecutado durante todo el ciclo de vida de un sistema de IA de alto riesgo, que requerirá revisiones y actualizaciones sistemáticas periódicas. Constará de las siguientes etapas:
 - a) la determinación y el análisis de los riesgos conocidos y previsibles que el sistema de IA de alto riesgo pueda plantear para la salud, la seguridad o los derechos fundamentales cuando el sistema de IA de alto riesgo se utilice de conformidad con su finalidad prevista;
 - b) la estimación y la evaluación de los riesgos que podrían surgir cuando el sistema de IA de alto riesgo se utilice de conformidad con su finalidad prevista y cuando se le dé un uso indebido razonablemente previsible;
 - c) la evaluación de otros riesgos que podrían surgir, a partir del análisis de los datos recogidos con el sistema de vigilancia poscomercialización a que se refiere el artículo 72;
 - d) la adopción de medidas adecuadas y específicas de gestión de riesgos diseñadas para hacer frente a los riesgos detectados con arreglo a la letra a).
- 4 Tomando como referencia los hechos conocidos por las fuerzas y los cuerpos de seguridad, incluidos todos los delitos que hayan empleado las TIC en su comisión.

Año	Cibercriminalidad sobre criminalidad total	Fraudes informáticos	Acceso e interceptación ilícita	Contra el honor	Propiedad industrial e intelectual	Delitos sexuales	Falsificación informática	Interferencia de datos y sistema	Amenazas y coacciones
2013	(42.403)	62,9 %	4,3 %	4,6 %	0,4 %	1,8 %	3,8 %	0,8 %	21,4 %
2014	(49.935)	65,8 %	3,7 %	4,4 %	0,4 %	1,9 %	3,7 %	0,9 %	19 %
2015	(60.154)	67,9 %	4 %	3,5 %	0,3 %	2 %	3,9 %	1,5 %	16,8 %
2016	4,6 % (66.586)	68,9 %	3,9 %	2,3 %	0,2 %	1,8 %	4,1 %	1,7 %	17,2 %
2017	5,7 % (81.307)	74,4 %	3,1 %	1,9 %	0,1 %	1,6 %	3,6 %	1,4 %	13,9 %
2018	7,5 % (110.613)	80,2 %	2,5 %	1,3 %	0,2 %	1,3 %	2,8 %	0,9 %	10,8 %
2019	9,9 % (218.302)	88,1 %	1,8 %	0,7 %	0,1 %	0,8 %	2 %	0,7 %	2 %
2020	16,3 % (287.963)	89,6 %	1,6 %	0,5 %	0,0 %	0,6 %	2,2 %	0,6 %	4,9 %
2021	15,6 % (305.477)	87,4 %	1,7 %	0,5 %	0,0 %	0,5 %	3,4 %	0,7 %	5,7 %
2022	16,1 % (374.737)	89,7 %	1,5 %	0,3 %	0,0 %	0,4 %	3,3 %	0,4 %	4,3 %
2023	19,2 % (472.125)	90,5 %	1,6 %	0,2 %	0,0 %	0,4 %	3,2 %	0,4 %	3,7 %

FUENTE: Elaboración del autor, sobre los datos del Ministerio del Interior.

El Informe anual sobre seguridad de 2023 hacía especial hincapié en el peligro de la utilización de la IA para elaborar campañas de desinformación, con las consecuencias que ello tiene, sobre todo en procesos electorales en los Estados democráticos, así como la especial vulnerabilidad del ciberespacio, advertencia que se hace más relevante en el Informe de situación de cibercriminalidad en España de 2024, cuando se afirma, expresamente:

La adopción de la IA se ha acelerado en el ámbito de la ciberseguridad, proporcionando herramientas de análisis predictivo y detección automatizada de amenazas. Sin embargo, al mismo tiempo, los ciberdelincuentes están aprovechando la IA para desarrollar ataques más precisos y difíciles de detectar. En 2024, solo el 2% de las organizaciones españolas ha implementado estrategias adecuadas para defenderse contra ciberataques habilitados por IA, lo que demuestra la necesidad urgente de adaptar las defensas frente a estas amenazas emergente [sic]. (p. 5)

En España, en fecha relativamente reciente, se constituyó la prevista Agencia Española de Supervisión de la Inteligencia Artificial (2023, pp. 122289-122316) como organismo público responsable de la supervisión, el asesoramiento, la concientización y la formación dirigidas a entidades públicas y privadas para la adecuada implementación de toda la normativa nacional y europea en torno al adecuado uso de los sistemas de IA y su desarrollo, con funciones inspectoras, y de comprobación y sanción.

III. LA RESPONSABILIDAD POR LA LESIÓN DE BIENES JURÍDICOS A CONSECUENCIA DE ACTUACIONES DE LA IA

Un robot o una máquina dotado(a) de IA (o incluso, un *software* programado con IA) puede causar, perfectamente, un daño que afecte bienes jurídicos protegidos por el Derecho penal. Aceptando esta realidad como cierta, debe plantearse si se cuenta en la actualidad con mecanismos jurídicos (civiles y penales) suficientes o si debe procederse a una revisión de los existentes, para adaptarlos o crear los demás que se requieran, en aras de poder aportar una respuesta satisfactoria en este tipo de situaciones, cuya incidencia se incre-

mentará poco a poco, a la vista de los grandes avances que se están produciendo en materia tecnológica.

Este reto alcanza mayores cotas de desafío cuanto más autónoma se configura la máquina y debe arbitrarse una respuesta garantista, también, en el supuesto en que el daño sea producido por un robot capaz de *decidir* por sí mismo. En este sentido, se expresa la Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión en cuanto a normas de Derecho civil sobre robótica, en la cual se reflexiona acerca de que, a mayor autonomía de los robots, mayor dificultad para que estos sean considerados instrumentos en manos de otros agentes. Ello obliga a replantearse si la normativa general sobre responsabilidad es suficiente o si, por el contrario, deben crearse nuevos principios sobre la responsabilidad jurídica de aquellos cuando se produzca un daño no atribuible a un agente humano, dada la insuficiencia de la normativa actual ante supuestos en que los robots puedan adoptar decisiones autónomas.

Como reza la referida Resolución, «la humanidad se encuentra a las puertas de una era en la que robots, bots, androides y otras formas de IA, cada vez más sofisticadas, parecen dispuestas a desencadenar una nueva revolución industrial» (Parlamento de Europa, 2017, pp. 1-2), por lo que resulta esencial una ponderación de las consecuencias jurídicas y éticas de dicha materia. Dicho instrumento alude, también, a la posibilidad de que la IA llegue a superar la inteligencia humana y admite que el desarrollo de máquinas inteligentes y autónomas no solo generan ventajas, sino, además, preocupaciones; asimismo, la tendencia hacia la automatización requiere que los implicados en el desarrollo y la comercialización de aplicaciones de IA incorporen, desde el principio, características de seguridad y ética, y estén preparados para aceptar la responsabilidad jurídica con respecto a la calidad de la tecnología que producen.

Con matices en cuanto a la ubicación de los elementos en las distintas concepciones dogmáticas del delito, este —en tanto categoría jurídica— gira, siempre, en torno al ser humano como sujeto responsable de una conducta típica, antijurídica, culpable y punible. Varía la ubicación de los elementos, pero en todo concepto de delito este aparece unido, necesariamente, a la persona, con su vertiente objetiva y, lo más complejo y apasionante en el Derecho penal, su

faz subjetiva, inseparable del ser humano. Sin embargo, ahora se enfrenta el dilema de saber «cuándo o con qué grado de inteligencia se supera el *umbral de responsabilidad*, entendiendo por tal aquel en el que el agente artificial puede ser hecho responsable por sus decisiones» (Cuesta, 2019, p. 52); o, en otras palabras, cuándo el ente dotado de IA deja de ser considerado *objeto* del ordenamiento jurídico y adopta las características de la categoría del *sujeto* en las relaciones jurídicas.

Con muchas reservas, el Derecho penal liberal se fue adaptando a las complejas estructuras sociales e incluyó la previsibilidad de la responsabilidad penal de las personas jurídicas. El Derecho y, en concreto, el penal, no puede ser ajeno a la sociedad en la que desempeña sus efectos. No cabe duda de que, transcurridas las dos primeras décadas del siglo xxi, la realidad social no puede disociarse de los efectos que el *big data* y la IA producen en todas las facetas y, también, como no podría ser de otra manera, en el mundo de la delincuencia. Al respecto, no ha de ignorarse que el Derecho, y también el penal, es una creación humana, un instrumento normativo del que se valen los hombres para asegurar la convivencia social (Romeo, 2022b, p. 8). La cuestión es la toma de decisiones acerca de si el sistema dotado de IA puede ser sujeto de responsabilidad y —de no serlo— cuáles serían los presupuestos y las condiciones para atribuir la responsabilidad a determinadas personas responsables de la supervisión y el control de aquel (Valls, 2022).

Las similitudes con la articulación de una respuesta jurídica, también penal, con respecto a las lesiones a bienes jurídicos realizadas en el ámbito de la persona jurídica, que ha dado lugar a admitir una responsabilidad penal de estos entes colectivos, sin duda, existe; ahora bien, cuanto más va avanzando el aprendizaje autónomo de la IA, más se aleja la posibilidad de plantear modelos de responsabilidad similares a los de entes colectivos dotados de personalidad jurídica propia.

Los distintos modelos de responsabilidad penal de la persona jurídica, al margen de que se establezca una responsabilidad directa o no, siempre parten de la comisión de un hecho típico, antijurídico y culpable ejecutado por una persona física, bien por aquella que tiene el poder de dirección de esta, bien por una falta de vigilancia y control de quien actúa en beneficio o por cuenta de ella. Ahora bien,

en el ámbito de la responsabilidad derivada del acto de la máquina dotada de IA, el problema se plantea cuando no es posible vincular el hecho lesivo al bien jurídico protegido con un actuar humano doloso o imprudente (Quintero, 2017, p. 22). En casos de IA generativa —debido al autoaprendizaje autónomo de la IA— se puede alcanzar un resultado que no era previsible en absoluto, sin que ni tan siquiera pueda afirmarse con seguridad una relación de causalidad con un acontecer humano, a cuyo efecto, incluso el caso fortuito es insuficiente (Rosal, 2023, p. 14; Lledó, 2022, p. 89).

En esta línea, la llamada *Ley de IA de la Unión Europea* incide especialmente en el análisis previo de los riesgos derivados de la IA, para que la creación de IA generativa, su comercialización y utilización siempre sea controlable por el ser humano. Desde esta perspectiva, señala Romeo (2022a), con pleno acierto, dos premisas que ayudarán a situar de forma correcta la cuestión: «el principio irrenunciable del control o dominio humano de los sistemas de IA», y «la oportunidad de establecer procedimientos de *compliance* o de atribución de obligaciones o deberes preventivos a los seres humanos que guardan determinada relación con tales sistemas, de forma similar a como se ha establecido para las personas jurídicas» (p. 103).

En cualquier caso, dependiendo del grado de autonomía de la decisión adoptada por la IA, se podrán plantear distintas propuestas de respuesta jurídico-penal (Rosal, 2023, pp. 21-44).

3.1. INTELIGENCIA ARTIFICIAL DÉBIL: PREVISIBLE Y CONTROLABLE

En estos casos se trata de entes sin capacidad de decisión propia, en los que existen personas físicas o jurídicas que están detrás del control real o potencial de la máquina autónoma.

No es posible plantear una *culpabilidad* de estos entes ni tampoco, en el ámbito penal, la *responsabilidad objetiva* por el resultado producido. Existen personas (físicas o jurídicas) responsables del control de los riesgos (previsibilidad y control) y ellas serán, también, las responsables de la actuación del ente dotado de IA ante un hipotético resultado lesivo.

- Se trata de sistemas de IA con una respuesta programada ante determinados estímulos, sin autonomía propia y, por tanto, sin capaci-

dad de actuación autónoma; tampoco tienen personalidad jurídica. Para estos casos, se ha planteado la posibilidad de aplicar distintas construcciones dogmáticas existentes: La teoría de la autoría mediata, para la cual el autor mediato sería —dependiendo del caso— el propietario o el usuario de la máquina (si no actuara de acuerdo con las instrucciones de funcionamiento), el programador, el fabricante o el sujeto que la comercializa. Para ello, se hace un paralelismo con la estructura teórica del *hombre de atrás* que realiza el tipo penal utilizando, en la ejecución del hecho, a un inimputable. La diferencia es evidente: el inimputable tiene personalidad jurídica (aunque no tenga reconocida la capacidad para ponerla en funcionamiento), mientras que el robot en la actualidad tiene el estatus jurídico de *cosa*. Comoquiera, teniendo en cuenta que el inimputable, en el caso del autor mediato, es considerado un *instrumento* de este, no existe problema en plantear que el robot es un *instrumento* del auténtico autor del delito, en el supuesto de la IA débil, controlable. En el tráfico jurídico y, también, en el Derecho penal, el robot no deja de tener el estatus de objeto, y no de sujeto del delito.

- Además, se plantea la posibilidad de atribuir la responsabilidad penal por el resultado lesivo producido al programador y al fabricante o, en su caso, al dueño o al usuario, por la violación del deber de garante que tiene sobre el robot, en dependencia de quién tenga el control sobre el riesgo en el caso concreto, lo cual permite plantear tanto la responsabilidad dolosa como la imprudente.
- Otra posibilidad, en caso de que la IA débil actúe al margen de lo previsible, es la de hacer responsable al programador o al fabricante atendiendo a la teoría de la responsabilidad por el producto defectuoso.
- Incluso, tomando como referencia estructuras evolucionadas, puede plantearse un modelo de prevención mediante programas de vigilancia y control similares a los modelos de cumplimiento normativo que se han implementado en los ordenamientos jurídicos para las personas jurídicas, con efectos de exención o atenuación de la responsabilidad penal.

En la línea de garantizar la supervisión y vigilancia humanas se han pronunciado las distintas resoluciones internacionales. Así, por ejemplo, en el ámbito europeo, el 8 de abril de 2019, la Comisión Europea publicó la Comunicación al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones:

«Generar confianza en la IA centrada en el ser humano», que recuerda el hecho de que la IA sirve para mejorar el bienestar de los humanos y considera que uno de los requisitos para obtener una IA fiable pasa por garantizar la intervención y supervisión humanas.

En este sentido, la Ley de IA de la Unión Europea de 2024 garantiza la previsión de que los sistemas que incorporen IA se encuentren siempre supervisados por una persona física, de forma que se prevengan o minimicen riesgos. En dichos casos, adquiere especial relevancia el límite máximo de riesgo permitido, tal como recoge la Ley de la IA de la Unión Europea de 2024.

3.2. INTELIGENCIA ARTIFICIAL FUERTE: APRENDIZAJE AUTÓNOMO

En el marco de la IA fuerte no son aplicables fórmulas propias de la comisión por omisión, la autoría mediata o la responsabilidad por el producto defectuoso, que son posibles con respecto a la IA débil. Tampoco son aplicables los argumentos dogmáticos que permiten atribuir responsabilidad penal a la persona jurídica. En estos casos comienza a hablarse de la necesidad de construir un tercer tipo de personalidad con capacidad para operar en el ámbito jurídico: *la personalidad electrónica* (Lledó, 2022, p. 143; Herrera, 2022, p. 36). La Resolución del Parlamento Europeo, de 16 de febrero de 2017, solicita a la Comisión, entre otras cosas, crear —a largo plazo— una personalidad jurídica específica para los robots, de forma que, como mínimo, los que sean autónomos y más complejos puedan ser considerados personas electrónicas responsables de reparar los daños que causen y, posiblemente, se aplique la personalidad electrónica a aquellos supuestos en los que los robots tomen decisiones autónomas inteligentes o interactúen con terceros de forma independiente.

Esta IA tiene autonomía propia, capacidad de actuar independientemente conforme con su autoaprendizaje, aunque en la actualidad no tienen personalidad jurídica ni, por tanto, capacidad jurídica individual. Cuando el robot está programado para desarrollar un proceso de aprendizaje, según su propia experiencia y eso dificulta el grado de predictibilidad de la respuesta, esto es, cuando la máquina está configurada para tomar sus propias decisiones, la respuesta jurídica tradicional presenta grandes fisuras, por cuanto es muy posible que,

en semejantes situaciones, se produzca la ruptura del nexo causal existente entre la actividad del productor, el distribuidor o el usuario y el daño causado por la máquina. La realidad muestra que la toma de decisión que hace actuar a la máquina inteligente se encuentra cada vez más alejada de la capacidad de decisión del ser humano que la programó, la produjo, la comercializó o la adquirió.

La responsabilidad en estos casos se ha afrontado en el seno de la Unión Europea con la limitación de la IA controlada por el ser humano, un sistema de control durante todo el proceso generativo, el control de riesgos, que incluye una vigilancia especial con respecto a los sistemas de alto riesgo (Artículo 6) y el establecimiento de una serie de prácticas de IA prohibidas (Artículo 5), si bien prevé un sistema de sanciones que deberá desarrollar cada Estado (Artículo 99) (Parlamento Europeo y Consejo de Europa, 2024, s.p.).

En esta línea, de *lege ferenda*, podría plantearse incluso una responsabilidad jurídica *ad hoc* para el robot inteligente, si bien, para ello, deberá crearse una nueva dogmática penal del delito cometido por la IA; una dogmática propia de la personalidad electrónica. El modelo europeo no prevé esa posibilidad, sino que limita el control siempre al ser humano y, por tanto, también la responsabilidad por la falta de control en alguna de las etapas de creación, producción, comercialización o utilización de la IA.

Todo lo expuesto obliga a plantear pautas para resolver los dilemas que se plantean. Los avances tecnológicos obligan a reevaluar conceptos tradicionales como:

- *La culpabilidad*, pues debe determinarse si siempre y en todo caso será el ser humano responsable por los actos de la máquina, lo que puede conllevar a supuestos de *responsabilidad objetiva*, denostada en el ámbito jurídico-penal. Tal situación puede plantearse en los casos de hechos lesivos causados por una IA fuerte, en cuyo control el ser humano ha fracasado. Difícilmente pueda afirmarse una relación de causalidad entre el fallo de vigilancia y control del ser humano, y el resultado causado por una decisión autónoma de la máquina, sin caer en la responsabilidad objetiva.
- Si resulta de aplicación en todo supuesto la ficción jurídica de la comisión por omisión al existir *posición de garante* y hasta dónde llegan sus límites.

- Hasta dónde se puede llegar con los criterios de imputación objetiva, para atribuir el resultado a una conducta humana precedente.
- Plantear la autoría mediata, para atribuir la responsabilidad a los programadores, fabricantes o usuarios, lo cual es difícil desde el momento en el que el autoaprendizaje es autónomo y se aleja cada vez más tanto del programador como del usuario.
- El papel que desempeña la responsabilidad penal de las personas jurídicas, que fuerza cierto paralelismo entre la ficción del ente colectivo y la máquina autónoma. Ello es difícil de asumir, pues, en última instancia, cuando se plantea la responsabilidad penal de la persona jurídica (directa o vicarial) se parte de que el delito lo ha realizado siempre una persona física, aun cuando esta no tenga responsabilidad criminal o, incluso, no haya llegado a identificarse.
- Aplicar a los programadores, fabricantes, comercializadores o usuarios (en dependencia del caso) de estos robots autónomos, las normas tradicionales de las *actuaciones en nombre de otro*, y trasladar la responsabilidad a la persona física o jurídica, lo cual parece descartado.

Comoquiera que este impacto puede representar una amenaza real, se precisa del estudio de los parámetros legales de dicho fenómeno, pues, más pronto que tarde, tanto las administraciones como los tribunales de justicia tendrán que enfrentarse en este ámbito a numerosos avatares de aún difícil respuesta (véase en relación con el vehículo autónomo o los drones de uso militar, por ejemplo).

Llegado este punto, es necesario plantearse si sería necesario que la máquina sienta, padezca, tenga conciencia de sí, de sus decisiones y del resultado de estas, o si bastaría con un modelo de toma de decisiones sin *sentimientos*. Obviamente, si se parte de los principios propios del Derecho penal liberal, sustentados en la culpabilidad, el hecho de que el robot dotado de IA fuerte pueda tomar decisiones basadas en un autoaprendizaje autónomo no implicaría, necesariamente, que esas decisiones sean racionales, sentimentales, emotivas, en definitiva... equiparables a las del ser humano, por lo que no cabe aplicar los postulados de la culpabilidad a esos entes autónomos. Otra opción, en su caso, sería conceptuarlos como *psicópatas* en el sentido más extremo, seres que actúan sin empatía, remordimientos ni sentimientos.

Entre las opciones que se plantean, se hallan:

- Responsabilidad propia de la máquina, a partir de la creación de una personalidad electrónica, esto es, la consideración de «personas artificiales», dotadas de personalidad jurídica propia, con capacidad jurídica y de obrar (Navarro, 2025, p. 201). En este caso habría que construir un nuevo concepto de delito que permita la atribución del hecho lesivo al ente dotado de IA con una construcción diferenciada de la *culpabilidad*. Al mismo tiempo, debería rediseñarse la respuesta penal, pues la pena y la medida de seguridad actuales están concebidas para el ser humano.
- Crear nuevos tipos penales con el Derecho penal tradicional, basados en el principio de precaución, y castigar la programación de robots dotados de una capacidad de autoaprendizaje que haga imprevisibles las respuestas futuras, su creación, posesión y utilización, en contra de los postulados de la Ley de IA de la Unión Europea. Ello sería factible con la introducción de modelos de control y vigilancia similares a los modelos de cumplimiento normativo de la persona jurídica, con respecto a los programadores, fabricantes, comercializadores, propietarios y usuarios. En estos casos, no se imputaría el resultado hipotéticamente producido, sino que serían delitos de mera actividad y peligro abstracto. El resultado, de darse, podría aparecer como un tipo agravado o una condición objetiva de punibilidad, según el caso.

Para concluir, han de dejarse abiertas las grandes cuestiones a las que tendrá que enfrentarse la dogmática penal en los próximos años: ¿es necesario un cambio en la estructura de la teoría jurídica del delito o construir otro concepto de *delito* para la sociedad del futuro, ajustado a un nuevo sujeto, distinto a la persona física y la jurídica: la *persona electrónica*? Todo dependerá del momento en el que la evolución de la IA permita afirmar una *acción* que produzca un cambio en el mundo exterior que, de ninguna forma remota, podría ser imputado a un ser humano, por su imprevisibilidad e imposibilidad de control. En ese momento, la máquina dejará de ser concebida como un objeto en el ordenamiento jurídico y habrá que pensar en respuestas no solo jurídico-civiles, sino también penales. Ante este panorama, parece sensato, de inmediato, aplicar el principio de precaución y evitar el desarrollo de la IA no controlable por el ser humano, en línea con la Ley de IA de la Unión Europea.

IV. CONCLUSIONES

La creación de artificios que llegan a actuar con cierta independencia del ser humano también va a influir con fuerza en el ámbito propio de la criminalidad. La era de la informática y las TIC, además de facilitar el desarrollo de la vida social e individual del ser humano, permitió la aparición de nuevas formas de realización de la actividad delictiva. Esta realidad se hace más palpable con la evolución de la tecnología y la implementación del *big data* y el mundo de los algoritmos en aparatos a los que se dota de la llamada *IA*, con una capacidad de aprendizaje cada vez más autónomo con respecto a la actuación de sus programadores, creadores y/o usuarios.

La llamada *Ley de IA de la Unión Europea* (reglamento de 2024) propone el establecimiento de una *IA* centrada en el ser humano y fiable, pero no recoge nuevos delitos en materia de *IA* ni reconoce a esta una capacidad jurídica autónoma.

En cualquier caso, dependiendo del grado de autonomía de la decisión adoptada por la *IA*, se podrán plantear distintas propuestas de respuesta jurídico-penal:

Para los casos de *IA* débil pudieran aplicarse distintas construcciones dogmáticas existentes:

- La teoría de la autoría mediata.
- La atribución de la responsabilidad penal por el resultado lesivo producido al programador y al fabricante o, en su caso, al dueño o usuario, con base en la violación del deber de garante que tiene sobre el robot.
- En caso de que la *IA* débil actúe al margen de lo previsible, podría hacerse responsable al programador o al fabricante atendiendo a la teoría de la responsabilidad por el producto defectuoso.
- Incluso, atendiendo a estructuras evolucionadas, es posible plantear un modelo de prevención mediante programas de vigilancia y control similares a los modelos de cumplimiento normativo que se han implementado en los ordenamientos jurídicos para las personas jurídicas, con efectos de exención o atenuación de la responsabilidad penal.

En el marco de la *IA* fuerte comienza a hablarse de la necesidad de ir construyendo un tercer tipo de personalidad con capacidad de operar

en el ámbito jurídico: la personalidad electrónica. Para tal supuesto, constituyen opciones:

- Responsabilidad propia de la máquina, mediante la creación de una personalidad electrónica, esto es, la consideración de «personas artificiales», dotadas de personalidad jurídica propia, con capacidad jurídica y de obrar.
- Reconocimiento de nuevos tipos penales con el Derecho penal tradicional, basados en el principio de precaución.

Sobre esta línea se debe continuar reflexionando.

V. REFERENCIAS

- Agustina, J. R. (2021). Nuevos retos dogmáticos ante la cibercriminalidad. ¿Es necesaria una dogmática del ciberdelito ante un nuevo paradigma? *Estudios Penales y Criminológicos*, 41, 705-777.
- Benítez Ortúzar, I. F. (2008). Informática y delito. Aspectos penales relacionados con las nuevas tecnologías. En Benítez Ortúzar, I. F. (Coord.). *Reforma del Código penal. Respuestas para una sociedad del siglo xxi*, 111-136. Dykinson.
- Benítez Ortúzar, I. F. (2021). Ciberdelitos. la implementación en el ordenamiento interno de los acuerdos internacionales en materia de cibercriminalidad. En Lledó Yagüe, F., Benítez Ortúzar, I. y Monje Balmaseda, O. (Dirs.). *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*, 79-128. Dykinson.
- Castillo Jiménez, C. (2001). Protección del derecho a la intimidad y uso de nuevas tecnologías de la información. *Derecho y Conocimiento*, 1, 35-48.
- Comisión Europea. (2018). Comunicación al Parlamento Europeo, al Consejo Europeo, al Consejo, al Comité Económico y Social Europeo, y al Comité de las Regiones. Inteligencia artificial para Europa. <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52018DC0237>
- Comisión Europea (2019). Comunicación al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo, y al Comité

- de las Regiones. Generar confianza en la inteligencia artificial centrada en el ser humano. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52019DC0168&from=LT>
- Consejo de Europa. (Septiembre 17, 2010). Convenio sobre ciberdelincuencia, de 23 de noviembre de 2001. *Boletín Oficial del Estado*, (226), 78847-78896. [https://www.boe.es/eli/es/ai/2001/11/23/\(1\)](https://www.boe.es/eli/es/ai/2001/11/23/(1))
- Consejo de Europa (2003). Protocolo adicional al Convenio sobre la ciberdelincuencia, relativo a la penalización de actos de índole racista y xenófoba cometidos por medio de sistemas informáticos. <https://www.coe.int/en/web/cybercrime/first-additional-protocol>
- Cuesta Aguado, P. M. (2019). Inteligencia artificial y responsabilidad penal. *Revista Penal México*, 9(16-17), 51-62. <https://revistacienciasinacipe.fgr.org.mx/index.php/01/article/view/336>
- Demetrio Crespo, E., Caro Coria, D. C. y Escobar Bravo, M. E. (Eds.). (2023). *Estudios de Derecho penal, neurociencias e inteligencia artificial*. Ediciones UCLM.
- Fiscalía General del Estado. (Octubre 11, 2011). Instrucción 2, sobre el Fiscal de Sala de Criminalidad Informática y las secciones de criminalidad informática de las Fiscalías. https://www.fiscal.es/memorias/estudios2016/INS/INS_02_2011.html
- Gobierno de España, Ministerio del Interior. (2023). Informe sobre la cibercriminalidad. https://www.interior.gob.es/opencms/export/sites/default/galleries/galeria-de-prensa/documentos-y-multimedia/balances-e-informes/2023/Informe-Cibercriminalidad_2023.pdf
- Herrera de las Heras, R. (2022). *Aspectos legales de la inteligencia artificial. Personalidad jurídica de los robots y responsabilidad civil*. Dykinson.
- Informe de situación de cibercriminalidad en España. (2024). https://ptedisruptive.es/wp-content/uploads/2024/12/INFORME-DE-SITUACION_CIBERSEGURIDAD_2024.pdf
- Lledó Benito, I. (2022). *El Derecho penal, robots, IA y cibercriminalidad: desafíos éticos y jurídicos. ¿Hacia una distopía?* Dykinson.

- Lledó Yagüe, F. (2022). *Los nuevos esclavos digitales del siglo XXI y la superación del hombre óptimo. ¿Hacia un nuevo derecho robótico?* Dykinson.
- McCarthy, J., Minsky, M. L., Rochester, N. y Shannon, C. E. (S.f.). A proposal for the Dartmouth Summer Research Project on Artificial Intelligence. <https://home.dartmouth.edu/about/artificial-intelligence-ai-coined-dartmouth>
- Miró Llinares, F. (2018). Inteligencia artificial y justicia penal: Más allá de los resultados lesivos causados por robots. *Revista de Derecho Penal y Criminología*, (20), 87-130.
- Morán Espinosa, A. (2021). Responsabilidad penal de la inteligencia artificial (IA). ¿La próxima frontera? *IUS*, 15(48), 289-323.
- Muñoz Vela, J. M. (Enero-marzo, 2022). Inteligencia artificial y responsabilidad penal. *Derecho Digital e Innovación*, (11), s.p. Wolters Kluwer.
- Navarro Cardoso, F. (2025). Anotaciones sobre la responsabilidad penal de las «personas artificiales». *Revista Penal*, (56), 198-207. <https://doi.org/10.36151/RP.56.10>
- Orden PCI/487/2019, de 26 de abril, «Estrategia nacional de ciberseguridad» (Abril 30, 2019). *Boletín Oficial del Estado*, (103), 43437-43455. <https://www.boe.es/buscar/doc.php?id=BOE-A-2019-6347>
- Parlamento Europeo. (Febrero 16, 2017). Resolución con recomendaciones destinadas a la Comisión, sobre normas de Derecho civil sobre robótica. https://www.europarl.europa.eu/doceo/documento/TA-8-2017-0051_ES.docx
- Parlamento Europeo y Consejo de Europa. (Julio 12, 2024). Reglamento (UE) 2024/1689. *Boletín Oficial del Estado*. <https://www.boe.es/buscar/doc.php?id=DOUE-L-2024-81079>
- Quintero Olivares, G. (2017). La robótica ante el Derecho penal: el vacío de respuesta jurídica a las desviaciones incontroladas. *Revista Electrónica de Estudios Penales y de la Seguridad*, (1), 1-22. <https://www.ejc.reeps.com/numeros-antiores/numero-1-segundo-semester-2017>

- Real Decreto 1150/2021, de 28 de diciembre, «Estrategia de seguridad nacional». (Diciembre 31, 2021). *Boletín Oficial del Estado*, (314), 167795-167830. https://www.boe.es/diario_boe/txt.php?id=BOE-A-2021-21884
- Real Decreto 729/2023, de 22 de agosto, «Estatuto de la Agencia Española de Inteligencia Artificial». (Septiembre 2, 2023). *Boletín Oficial del Estado*, (210), 122289-122316. <https://www.boe.es/eli/es/rd/2023//08/22/729>
- Romeo Casabona, C. (2022a). Avances científicos y Derecho penal. En Demetrio Crespo, E., Cuerda Martín, M. y García de la Torre García, F. (Coords.). *Derecho penal y comportamiento humano. Avances desde la neurociencia y la inteligencia artificial*, 77-107. Tirant Lo Blanch.
- Romeo Casabona, C. (2022b). La atribución de responsabilidad penal por los hechos cometidos por sistemas autónomos inteligentes, robótica y tecnologías conexas. *ULP Law Review i Revista de Direito da ULP*, 16, 7-16.
- Rosal Blasco, B. (Junio, 2023). ¿El modelo de la responsabilidad penal de las personas jurídicas para los daños punibles derivados del uso de la inteligencia artificial? *Revista de Responsabilidad Penal de Personas Jurídicas y Compliance*, 2, 1-49. <https://www.redepec.com/wp-content/uploads/pdf/-el-modelo-de-la-responsabilidad-penal-de-las-personas-jur-dicas-para-los-da-os-punibles-derivados-del-uso-de-la-inteligencia-artificial--3098.pdf>
- Valls Prieto, J. (2021). Inteligencia artificial, derechos humanos y bienes jurídicos. Aranzadi.
- Valls Prieto, J. (2022). Sobre la responsabilidad penal por la utilización de sistemas inteligentes. *Revista Electrónica de Ciencia Penal y Criminología*, (24), 24-27. <https://revistacriminologia.com/24recpc24-27.pdf>