IUS ET SCIENTIA
2024 • Vol. 10 • Nº 2 • ISSN 2444-8478
https://editorial.us.es/es/revistas/ius-et-scientia
https://dx.doi.org/10.12795/IETSCIENTIA • © Editorial Universidad de Sevilla 2024


# La inteligencia artificial en el proceso penal: eficiencia versus garantías

## ARTIFICIAL INTELLIGENCE IN THE CRIMINAL PROCESS: EFFICIENCY VERSUS GUARANTEES

**Irene Yáñez García-Bernalt**
Personal Investigador Postdoctoral en Derecho Procesal – Universidad de Salamanca
ireneygb@usal.es 0000-0001-8859-6029

Recibido: 03 de septiembre de 2024 | Aceptado: 28 de noviembre de 2024

RESUMEN

La incorporación de la Inteligencia Artificial (IA) al proceso penal genera tensiones entre la eficiencia judicial y el respeto de las garantías y derechos procesales fundamentales. La IA puede acelerar, sin duda, la tramitación del processo, mejorar la gestión y ofrecer herramientas predictivas destinadas a apoyar las decisions. No obstante, plantea serios riesgos como la Vulneración del derecho de defensa, la igualdad, la presunción de Inocencia y la protección de los datos personales, debido a los posibles sesgos y errores algorítmicos. Para garantizar un processo con todas las garantías será crucial implementar mecanismos de supervisión humana, transparencia y explicabilidad. El uso ético, legal y garantista se erige como esencial para mantener el equilibrio entre un proceso penal garantista y las tecnologías

PALABRAS CLAVE

Garantías
Investigación
Proceso penal
Derechos fundamentales
Inteligencia Artificial

ABSTRACT

The incorporation of Artificial Intelligence (AI) into criminal proceedings creates tensions between judicial efficiency and the protection of fundamental procedural guarantees and rights. AI can, undoubtedly, expedite case processing, improve management, and provide predictive tools to support decision-making. However, it also poses significant risks, such as the violation of the right to defense, equality, the presumption of innocence, and personal data protection, due to potential algorithmic biases and errors. To ensure a process that fully upholds procedural guarantees, it will be essential to implement mechanisms for human oversight, transparency, and explainability. An ethical, lawful, and rights-based use of AI is crucial to maintaining the balance between a guarantee-based criminal process and the integration of these technologies.

KEYWORDS

Guarantees
Investigation
Criminal process
Fundamental rights Artificial Intelligence

## I. Introducción: la irrupción de la Inteligencia Artificial en la sociedad

La Inteligencia Artificial (IA, en adelante) es una ya no tan nueva tecnología que ha irrumpido en nuestros quehaceres diarios de manera profunda y variada, transformando desde el modo en qué trabajamos, hasta nuestras formas de comunicación social. Cuando hablamos de IA nos referimos a una noción centrada en la capacidad de elegir la mejor acción para lograr un determinado objetivo teniendo en consideración ciertos criterios que se deben optimizar en función de los recursos disponibles, siendo así la racionalidad una parte significativa de esta herramienta[1]. Esta tecnología, cada vez más asentada y que quizá antes asociábamos a un mundo de ficción, hoy está presente en la mayoría de las aplicaciones, dispositivos y servicios que utilizamos diariamente. Nuestro contacto con esta tecnología es, cuanto menos, constante: desde el uso de relojes inteligentes que miden nuestras constantes vitales y niveles de estrés, hasta los sistemas virtuales de geolocalización, sistemas de luz inteligentes en nuestros hogares o mecanismos de alarmas. Podríamos afirmar, sin miedo alguno, que nos encontramos no ya una «revolución 4.0» sino en una «revolución 5.0» en la que, más pronto que tarde, irrumpirá su sexta edición.

El mundo del Derecho no es impermeable, de modo que la IA también ha empezado a jugar un papel, cada vez más protagonista, en los sistemas de administración de justicia de los diferentes Estados con el objetivo de aportar herramientas de mejora en la precisión, eficiencia y transparencia de los procedimientos dentro de cada orden jurisdiccional. Sin embargo, su implementación comporta también considerables preocupaciones a nivel ético y legal, puesto que sus decisiones pueden incidir en derechos fundamentales de los ciudadanos y afectar, gravemente, las garantías procesales que rigen en nuestro ordenamiento jurídico. Tales consecuencias nos llevan a defender la idea de que esta tecnología no debería sustituir al factor humano, sino que debe ser un complemento que nos permita avanzar hacia el futuro, a una combinación algoritmo-humano que pueda ofrecernos beneficios desde la óptica de la eficacia y la eficiencia (Martín Diz, 2024).



Poniendo el foco de atención en el proceso penal, estos nuevos sistemas basados en algoritmos entrenados a través de la técnica del *machine learning*, permiten obtener, gestionar y procesar una ingente cantidad de datos con la finalidad de elaborar informes que favorezcan la toma decisiones. En la fase de instrucción del proceso penal, estos reportes podrían ayudar a decidir sobre la apertura o no del juicio oral, la concurrencia de alguna de las causas para decretar el sobreseimiento o el archivo de la causa (Barona Vilar, 2021). Estos sistemas de IA pueden suponer –y suponen– una auténtica revolución en el modo de proceder en la fase de instrucción al ofrecer avanzadas herramientas para mejorar la eficiencia y precisión tales como el reconocimiento facial

* Este trabajo ha sido realizado dentro de la convocatoria de contratos predoctorales en período de orientación postdoctoral (POP) (Programa Propio III) USAL 2020, cofinanciada por el Banco Santander.

1. Véase en este sentido el documento *A definition of AI: main capabilities and disciplines* elaborado por el Grupo de Expertos Independientes en Inteligencia Artificial de la Comisión Europea, de 8 de abril de 2019.

IUS ET SCIENTIA • 2024
Vol. 10 • Nº 2 • pp. 80-100
ISSN 2444-8478 • https://dx.doi.org/10.12795/IESTSCIENTIA.2024.i02.04

y análisis de imágenes, análisis de datos y redes sociales, técnicas OSINT o chatbots y asistentes de IA. Sin embargo, los compromisos legales y éticos antes mencionados nos llevan a analizar y estudiar la posible afectación o compromiso de determinados derechos procesales fundamentales y otras garantías procesales inherentes al investigado.

## II. Marco legal del uso de las tecnologías disruptivas

Para situar al lector en contexto, es necesario explicar brevemente el marco legal de referencia sobre el uso de las tecnologías disruptivas en el proceso penal. Los avances en esta materia no dejan indiferentes a las políticas públicas y al poder legislativo de los países. Los constantes beneficios, pero también las preocupantes amenazas de esta inteligencia conducen a la necesidad de legislar sobre esta materia, en nuestro caso tanto a nivel de la Unión Europea (UE, en adelante), como a nivel interno a través del Derecho nacional.

### 2.1. La IA en el marco regulatorio de la Unión Europea

En primer lugar, aunque no se trata precisamente la incorporación de la IA al proceso penal, es importante tener presente la Directiva (UE) 2016/680 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativa a la protección de las personas físicas en el tratamiento de datos personales por parte de las autoridades competentes para fines de prevención, investigación, detección o enjuiciamiento de infracciones penales o de ejecución de sanciones penales y a la libre circulación de dichos datos[2]. Este instrumento pone de relieve la importancia de circulación de los datos personales entre las autoridades judiciales y policiales para mejorar la eficacia de la cooperación en materia penal haciendo hincapié en el fortalecimiento de los derechos de los interesados y de las obligaciones de los encargados del tratamiento de sus datos. Es una cuestión fundamental y a tener en cuenta en los posteriores textos normativos que regularán la incorporación de la IA a los sistemas de administración de justicia, por cuanto la protección de los derechos y garantías de los sujetos debe inclinar la balanza frente a la eficiencia que puede generar la utilización de las herramientas tecnológicas. Esta Directiva fortalece, pues, la protección de los datos personales en el contexto del proceso penal asegurando una mayor coherencia en la regulación de la protección de los datos en el seno de la UE, ajustándola a los avances tecnológicos en congruencia con las necesidades de protección de los derechos fundamentales de los ciudadanos comunitarios.

Cuatro años más tarde, en febrero de 2020 en los albores de la nueva pandemia mundial que acechaba a la sociedad globalizada, se publicaba el Libro Blanco sobre la IA por parte de la Comisión Europea. Se trata del resultado de la estrategia que la UE había adoptado en 2018 en materia de IA, representa un marco de referencia para el



2. Ténganse en cuenta otros instrumentos como la Directiva (UE) 2018/1972 del Parlamento Europeo y del Consejo, de 11 de diciembre de 2018, por la que se establece el Código Europeo de las Comunicaciones Electrónicas y el Reglamento 2021/694 del Parlamento Europeo y del Consejo de 29 de abril de 2021 por el que se establece el Programa Europa Digital y por el que se deroga la Decisión (UE) 2015/2240.

IUS ET SCIENTIA • 2024
Vol. 10 • Nº 2 • pp. 80-100
ISSN 2444-8478 • https://dx.doi.org/10.12795/IESTSCIENTIA.2024.i02.04

desarrollo de una IA segura, responsable y ética dentro de la UE. El documento establece unas líneas y propuestas para regular y fomentar la implementación de la IA en el marco de la Unión buscando un equilibrio entre la innovación y la protección de los derechos fundamentales, es decir, busca desplegar las líneas de actuación de la UE a seguir en esta materia. El documento pone de manifiesto, en primer lugar, las amenazas y temores que supone la incorporación de la IA centrándose en el problema de la opacidad de los algoritmos para la toma de decisiones, los posibles sesgos o la utilización de la IA con fines delictivos (De Hoyos Sancho, 2021). La mención en las líneas que nos preceden a la Directiva (UE) 2016/680, de 27 de abril de 2016, no es casual, pues el funcionamiento de muchos de los sistemas de IA, así como las acciones y decisiones que puedan adoptar depende, en gran medida, de los datos utilizados para el entrenamiento. En este sentido es necesario respetar los requisitos dirigidos a garantizar que la privacidad y los datos personales estarán adecuadamente protegidos en el uso de herramientas basadas en la IA, especialmente se habrán de tener en cuenta las indicaciones de la mencionada Directiva. Teniendo en cuenta las posibles amenazas y el equilibrio y debida protección de los derechos fundamentales se pretende crear un ecosistema de excelencia que impuse la innovación de la IA en Europa y otro ecosistema de confianza en el que los ciudadanos puedan estar seguros de que esta herramienta respetará sus derechos siguiendo unos principios éticos claros. Tales principios se resumen en la seguridad y ética, la transparencia en el sentido de que los sistemas de IA sean comprensibles y con decisiones explicables, la imparcialidad y no discriminación evitando todo sesgo en los datos de entrenamiento y, por último, la responsabilidad por cuanto los desarrolladores de estos sistemas deben operar con cautela asumiendo un rol de responsables en los efectos que la IA despliega en la sociedad.



Un año después, en 2021, se hace público el Reglamento 2021/694 del Parlamento Europeo y del Consejo de 29 de abril de 2021 por el que se establece el Programa Europa Digital y por el que se deroga la Decisión (UE) 2015/2240, este nuevo texto sienta las bases para regular de una manera más contundente el uso de la IA para con los datos personales de carácter electrónico insistiendo en su necesaria implementación respetando los pilares fundamentales sobre los que descansa la UE (Bueno de Mata, 2021). La finalidad general del Programa Europa Digital es apoyar y acelerar la transformación digital de la economía, sociedad e industria europeas. El art. 5 del mencionado Reglamento presenta como objetivo específico 2 el uso de la IA con el fin de desarrollar las capacidades y conocimientos básicos de esta inteligencia en la UE.

El punto álgido de la regulación europea en materia de IA lo encontramos en 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, el cual ya se conoce comúnmente como Reglamento de Inteligencia Artificial (RIA, en adelante). Este Reglamento nace en el marco de la Estrategia Europea de IA de la Comisión Europea, mediante el cual se buscar convertir a la UE en una referencia mundial para la IA garantizando que esta tenga un enfoque antropocéntrico, sostenible, seguro y fiable. El Reglamento incide en la regulación de diferentes niveles de riesgo en atención a las características y funcionalidades de los sistemas de IA clasificando los niveles en: riesgo inaceptable o prohibido –explotación de vulnerabilidades de grupos específicos, sistemas que permiten la clasificación por parte de gobiernos,

IUS ET SCIENTIA • 2024
Vol. 10 • Nº 2 • pp. 80-100
ISSN 2444-8478 • https://dx.doi.org/10.12795/IESTSCIENTIA.2024.i02.04

sistemas para evaluar o predecir el riesgo de que una persona física cometa un delito, sistemas que amplíen bases de datos a través del reconocimiento facial a través de una selección de imágenes no selectivas en Internet o circuitos cerrados, sistemas de categorización biométrica que clasifiquen de manera individual a personas tomando como base datos para deducir su raza, creencias religiosas o políticas, orientación sexual, sistemas de identificación biométrica en tiempo real y lugares públicos, etc.- Se consideran como sistemas de alto riesgo aquellos que pueden perjudicar a la seguridad, salud o derechos fundamentales de las personas, así conforme al art. 6.1 del RIA se establecen distintas categorías de alto riesgo –sistemas de IA utilizados como componentes de seguridad de productos que están cubiertos por la legislación europea de seguridad de productos, sistemas independientes con implicaciones en los derechos fundamentales como los de identificación biométrica remota o los usados para la gestión de la migración y control de fronteras–. En cuanto los sistemas de riesgo limitado son aquellos que presentan algunos problemas en relación con la falta de transparencia –sistemas de IA destinados a interactuar con personas físicas como los *chatbots*, los que generan contenidos de audio, video o imagen, o los sistemas de *deepfakes* en virtud de los cuales se manipula una imagen, vídeo o sonido de tal modo que se asemeja a personas físicas, lugares o sucesos reales. Por último, como sistemas de riesgo mínimo se consideran aquellos que libre uso de aplicaciones como los que se incluyen en videojuegos (Barrios Andrés, 2024). La clasificación de los niveles de riesgo comportará, de manera obligatoria, la adopción de sistemas de gestión riesgo, prácticas adecuadas en la gestión de datos, supervisión por parte de personas físicas y otorgamiento de un nivel adecuado de precisión y ciberseguridad (Ruiz Forns y Nicolás, 2024).



## 2.2. La regulación de las diligencias de investigación tecnológicas en el ordenamiento jurídico español

No es esta el año 2015 cuando se hace frente al desfase existente entre la LECrim y el auge de las nuevas tecnologías en la sociedad, es en este momento cuando entra en vigor la Ley Orgánica 13/2015, de 5 de octubre, de modificación de la Ley de Enjuiciamiento Criminal, para el fortalecimiento de las garantías procesales y la regulación de las medidas de investigación tecnológica (Alonso Salgado, 2015). Hasta esta reforma, de manera sucinta el art. 579.2 LECrim recogía la posibilidad de que la autoridad judicial acordase, a través de resolución motivada, la intervención de las comunicaciones telefónicas. La aprobación de la LO 13/2015, consciente del vacío legal existente en esta materia y de las dificultades de las Fuerzas y Cuerpos de Seguridad (FCS, en adelante) para investigar determinados hechos delictivos, supone un soplo de aire fresco con fines de innovación en la regulación de las diligencias de investigación (Richard González, 2016). Así la modificación incorpora en la LECrim una regulación más amplia de la interceptación de las comunicaciones y como novedad añade la utilización de dispositivos técnicos de seguimiento, localización y captación de la imagen, el registro de dispositivos de almacenamiento masivo, los registros remotos sobre equipos informáticos y el agente encubierto informático para actuar en canales cerrados de comunicación. Estas medidas de investigación pretenden así alcanzar el ritmo de los delincuentes en la comisión de hechos delictivos a través de tecnologías disruptivas (González Pulido, 2023).

IUS ET SCIENTIA • 2024
Vol. 10 • Nº 2 • pp. 80-100
ISSN 2444-8478 • https://dx.doi.org/10.12795/IESTSCIENTIA.2024.i02.04

La utilización de estas medidas no puede, sino, más que estar sujeta a la concurrencia de una serie de principios rectores que debe respetar dada la importante injerencia que su uso genera en determinados derechos fundamentales del investigado, como es el secreto de las comunicaciones (Art. 18.3 CE). Así la Circular 1/2019, de 6 de marzo, del Fiscal General del Estado, sobre disposiciones y medidas de aseguramiento de las diligencias de investigación tecnológicas en la Ley de Enjuiciamiento Criminal regula como principios rectores los siguientes: la especialidad, idoneidad, excepcionalidad y necesidad y proporcionalidad. El principio de especialidad implica que la intromisión en la intimidad del investigado no puede tener lugar «a ciegas», la medida debe estar relacionada con la investigación de un delito concreto de modo que, a la hora de solicitar su uso, se debe precisar el tipo de delito, aunque posteriormente pueda modificarse por la adición de otras particularidades penales[3]. En cuanto al principio de idoneidad, este implica que la medida resultará idónea cuando sea adecuada a los fines de la instrucción o permita continuar avanzando en la investigación (Bueno de Mata, 2019). La excepcionalidad y necesidad comportan que el recurso a la mismas debe ser la última opción y no considerarse de uso rutinario, es decir, si es posible emplear otra medida que permita esclarecer los hechos sin provocar una importante injerencia en los derechos fundamentales del investigado, la diligencia tecnológica quedará descartada. Por último, la proporcionalidad se refiere al equilibrio que debe operar en el sentido de que, en atención a las circunstancias del caso y el sacrificio de los intereses y derechos del investigado no puede ser superior al beneficio de que su adopción resulte para el interés público y de terceros.

La alusión al marco regulatorio de las diligencias de investigación tecnológicas no es casual, pues entendemos que la IA, la cual consideramos como una medida de este tipo, debe respetar de igual modo los principios rectores que operan para poder hacer uso de cualquier herramienta tecnológica en la fase de instrucción de una causa penal que así lo requiera. Téngase en cuenta que existen diversos derechos fundamentales y garantías procesales que se verán afectados por la injerencia de este instrumento y que, a continuación, serán expuestos.



## III. Derechos fundamentales y garantías procesales: Inteligencia Artificial ¿enemiga o aliada?

### 3.1. La incorporación de la Inteligencia Artificial al proceso

Como ya veníamos advirtiendo al comienzo de este trabajo, la esfera jurídica, en general, y el Derecho Procesal, en particular, no han quedado ajenos a las ventajas que pretenden ofrecer los sistemas de IA. El proceso, como instrumento de que dispone el Estado y mediante el cual la jurisdicción resuelve y decide sobre los conflictos intersubjetivos y sociales surgidos en una comunidad (Asencio Mellado, 2008) se construye bajo

3. Véase en este sentido la STS 393/2012, de 29 de mayo donde se recoge que «debe existir una especialidad de la materia a investigar porque no cabe decretar la intervención telefónica para propiciar el descubrimiento genérico de posibles infracciones penales, lo que supondría conceder autorizaciones en blanco»

IUS ET SCIENTIA • 2024
Vol. 10 • Nº 2 • pp. 80-100
ISSN 2444-8478 • https://dx.doi.org/10.12795/IESTSCIENTIA.2024.i02.04

toda una serie de principios, derechos y garantías reconocidas a las partes en aras de consagrar, en todo momento, la igualdad. De manera más concreta el proceso penal se erige como la vía que tiene por objeto la declaración del delito y la imposición de una pena, es decir, la herramienta que posee el Estado para aplicar el *ius puniendi* por cuanto esto se considera como una consecuencia de los postulados del Estado de Derecho (Goldschmidt, 2021).

Ahora bien, la concepción del proceso penal como garantista, sumado al difícil control del desarrollo tecnológico, plantea dificultades en la gestión de los posibles riesgos generados y en el mantenimiento de las garantías y derechos procesales fundamentales. La incorporación de la IA al proceso penal supone un auténtico reto en el mantenimiento del debido proceso (San Miguel Caso, 2023). El uso de sistemas de IA como herramienta de apoyo al ejercicio de la función jurisdiccional, especialmente en las fases de instrucción y enjuiciamiento, es una realidad asentada (De Hoyos Sancho, 2021). Teniendo en cuenta que su uso ha llegado para quedarse, la búsqueda de la eficacia y de la eficiencia a través del uso de sistemas de IA y, por tanto, el beneficio obtenido mediante su implementación no puede dar lugar a un sacrificio mayor que aquel de los derechos fundamentales de la persona. El entrenamiento de los datos a través de la técnica del *machine learning*, en virtud de la cual se produce un aprendizaje autónomo, es de difícil control provocando así que su uso dentro del proceso penal resulte complejo y cuya aplicación deba tener lugar con la máxima cautela. La utilización de esta tecnología puede poner en peligro los principios sobre los que descansa el proceso penal, de modo que será necesaria trazar debidamente el ámbito de aplicación de estas sistemas y marcar una serie de requisitos para que puedan ser empleados en las diferentes fases del proceso penal (De Miguel Beriain y Pérez Estrada, 2019; Noya Ferreiro, 2022).



## 3.2. La protección de los datos personales, la intimidad y la propia imagen frente al uso de la Inteligencia Artificial

El primer derecho fundamental afectado que se nos viene a la mente cuando hablamos de IA es, indudablemente, la protección de los datos personales. El art. 18 CE regula una limitación en el uso de la informática para garantizar el honor y la intimidad personal y familiar de los ciudadanos y es ahí, precisamente, donde se incardina la debida protección de los datos personales[4]. Nos referimos a un derecho de carácter instrumental en el sentido de que es la garantía de los derechos al honor y la intimidad y el disfrute de los mismos y, por otro lado, podría considerarse como un derecho autónomo porque su reconocimiento comporta un control del torrente de las informaciones de

4. El reconocimiento de la protección de datos como un derecho fundamental quedo recogido también en el art. 8 de la Carta de los Derechos Fundamentales de la Unión Europea (CDFUE), en el art. 8 del Convenio Europeo de los Derechos Humanos (CEDH) o en el Tratado de Funcionamiento de la Unión Europea en su art. 6 (TFUE). Asimismo, aunque no tiene carácter normativo esta protección se reconoce también como derecho en el apartado 1.III de la Carta de Derechos Digitales del Gobierno de España. Téngase en cuenta también, el Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, (Reglamento General de Protección de Datos) y la LO 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.

IUS ET SCIENTIA • 2024
Vol. 10 • Nº 2 • pp. 80-100
ISSN 2444-8478 • https://dx.doi.org/10.12795/IESTSCIENTIA.2024.i02.04

cada ciudadano (Gómez Abeja, 2022). La protección de datos en el seno del proceso penal también es esencial, la recolección, procesamiento, almacenamiento y uso de datos personales en este sentido plantea importantes desafíos dada la naturaleza sensible y la potencial trascendencia de esa información que incluye antecedentes penales, investigaciones en curso y datos biométricos.

En cuanto al uso de la IA y su posible injerencia en este derecho, debemos partir de la base de que estos sistemas reciben y son entrenados con una ingente cantidad de datos, muchos de ellos de carácter personal. Es más, una importante cantidad de herramientas basadas en IA recolectan de forma masiva datos que, a nuestros ojos, pasan inadvertidos (Fernández-Aller y Serrano Pérez, 2022). Los sistemas de IA necesitan todos esos datos para entrenar y mejorar sus algoritmos siendo aquellos, muy a menudo, datos personales y de naturaleza sensible. Las principales áreas de preocupación residen en la recopilación masiva, especialmente en sectores como la salud, la banca o el marketing. Por otro lado, encontramos la anonimización y reidentificación, aunque teóricamente los datos se deben anonimizar antes de ser utilizados en instrumentos de IA, en muchas ocasiones es posible reidentificar a la persona a través del cruce de datos anónimos con otras bases de datos. A ello debemos sumar que la recopilación y el almacenamiento de estas cantidades de datos hace que las bases de datos de origen sean susceptibles de ciberataques pudiendo exponer así información privada de millones de sujetos y pudiendo emplearse para entregar IA malintencionadamente.

En el ámbito del proceso penal, ya hemos indicado que la IA puede ofrecer grandes ventajas como la agilización de los procedimientos o la mejora en la precisión de las investigaciones, sin embargo, la injerencia en la protección de datos es también preocupante. El uso de la IA y la posible vulneración de la protección de datos se puede apreciar, por ejemplo, en el análisis de datos biométricos (huellas dactilares o reconocimiento facial), así como en comunicaciones privadas y datos de localización. En más de una ocasión, los sujetos sospechosos no son conscientes del alcance de la recopilación de sus datos y es probable que estos queden almacenados incluso si no se incoa el proceso o se obtiene, finalmente, una sentencia absolutoria. El empleo de sistemas de IA de reconocimiento facial, por ejemplo, en sistemas de videovigilancia comporta un tratamiento de datos biométricos, la existencia de un posible interés público –como es el caso de un delito de hurto en un supermercado que emplea estos sistemas– no legitima cualquier tipo de tratamiento de los datos personales[5]. De igual modo la obtención de datos a través de fuentes como redes sociales y otras plataformas digitales dificultan el control por parte de los titulares.



5. Véase en este sentido el Auto de la AP de Barcelona 72/2021, de 15 de febrero (Rec. 840/2021). En este auto se deniega a una cadena de supermercados utilizar medios automatizados de captación de datos biométricos de unos penados por delito de robo con la finalidad de detectar su entrada en cualquier establecimiento de esta cadena. Indica el Auto que no se está protegiendo aquí el interés público, sino los intereses privados de la empresa propietaria de la cadena y se estaría conculcando la protección de los derechos y libertades de los sujetos. Asimismo, de acuerdo con lo establecido en la LO 7/2021, de 26 de mayo, de protección de datos personales tratados para fines de prevención, detección, investigación y enjuiciamiento de infracciones penales y ejecución de sanciones penales, la instalación de sistemas de grabación de imágenes y sonido debe responder al principio de proporcionalidad.

IUS ET SCIENTIA • 2024
Vol. 10 • Nº 2 • pp. 80-100
ISSN 2444-8478 • https://dx.doi.org/10.12795/IESTSCIENTIA.2024.i02.04

Estos sistemas de vigilancia están en auge[6], forman parte de la expansión del derecho penal y de la solicitud, cada vez mayor, de prevención y seguridad en una sociedad que aboga por políticas de endurecimiento y estas herramientas suponen un control constante de todos los ciudadanos (Barona Vilar, 2022). Conocidos como Policía predictiva estos programas han traído un gran debate en torno a sus beneficios y costes en relación con su uso en el ámbito de la seguridad. Se considera que estos sistemas permiten optimizar la gestión favoreciendo la distribución de patrullas en zonas geográficas consideradas con mayor posibilidad de actividad delictiva y empleando menos recursos para hacerlos efectivos y, además, la recopilación de los datos permite una identificación de patrones, tendencias y relaciones que pueden ser usadas por las Fuerzas y Cuerpos de Seguridad (FCS, en adelante) para anticiparse al delito (Cinelli, 2019). A modo de ejemplo encontramos la herramienta *Predictive Policing* (PredPol), utilizada en ciudades de Estados Unidos. Este instrumento emplea datos de delitos ya cometidos y algoritmos para predecir cuáles son aquellas zonas donde es probable que se perpetren hechos delictivos. Los datos obtenidos permiten a las autoridades desplegar recursos en dichas áreas para disuadir o prevenir aquellos. Evidentemente ha sido objeto de críticas por la concurrencia de sesgos en sus predicciones ya que los datos históricos suelen estar influenciados por las desigualdades sociales a la par que se manejan datos sensibles como son los antecedentes penales. Del mismo modo, en 2020 la Policía Metropolitana de Londres comenzó a utilizar cámaras de reconocimiento facial en distintas áreas públicas de la ciudad. Estas se encuentran diseñadas para identificar a personas que en situación de búsqueda por la comisión de delitos graves comparando sus rostros con una lista de vigilancia en tiempo real[7] pudiendo hablar entonces de un «Gran Hermano» o *Big Brother Watch*[8]. Los sistemas de vigilancia masiva no suponen una novedad, ya que en la década de los 90 otros países –Canadá y Australia– emplearon sistemas conocidos como «Cinco ojos» (Peralta Gutiérrez, 2021). Estamos así ante un complejo régimen jurídico, donde hemos de tener claro que los sistemas de identificación son posibles siempre bajo la premisa de la excepcionalidad y siempre y cuando respondan a los principios de necesidad y proporcionalidad tal y como se contempla para el uso de diligencias de investigación tecnológicas (Cotino Hueso, 2023).



### 3.2.1. Uso de herramientas de monitorización en redes sociales

Vinculado a la protección de datos, queremos centrarnos en el uso de herramientas de monitorización y seguimiento de actividades en línea. Los sistemas de monitorización

6. En 2018 en Niza (Francia) se implementó un sistema de detección del comportamiento que analizaba las imágenes de las cámaras de seguridad en busca de comportamiento anómalos que podrían ser indicativos de la comisión de un delito. Aunque el sistema tenía una intención de mejorar la seguridad planteaba problemas en la privacidad, propia imagen e intimidad de las personas percibiéndose como una vigilancia invasiva.

7. Fuera del ámbito jurídico estas cámaras están siendo empeladas en el campo del marketing para desplegar publicidad personalizada. https://elpais.com/ciencia/2024-06-17/camaras-con-ia-en-el-metro-de-londres-captan-el-estado-emocional-de-los-viajeros.html

8. Véase la STEDH 58170/13 62322/14 24960/15, de 25 de mayo de 2021 (Asunto Big Brother y otros c. Reino Unido).

IUS ET SCIENTIA • 2024
Vol. 10 • Nº 2 • pp. 80-100
ISSN 2444-8478 • https://dx.doi.org/10.12795/IESTSCIENTIA.2024.i02.04

basados en IA son ya una realidad empleada en redes sociales utilizados esencialmente para analizar y recopilar información sobre el comportamiento en línea de los usuarios, identificando patrones, tendencias o potenciales amenazas. Las agencias de seguridad de diversos países han empleado estos sistemas en busca de palabras clave, imágenes o comportamientos asociados con actividades delictivas como el terrorismo, la violencia organizada o el tráfico de drogas. En Estados Unidos la Policía de Nueva York (NYPD) ha hecho uso de herramientas de monitoreo en redes para identificar a miembros de bandas organizadas, estos sistemas analizan publicaciones en las redes sociales en busca de símbolos o menciones que están vinculados a la actividad delictiva permitiendo a la policía anticiparse a posibles enfrentamientos entre bandas (Rodríguez Andrés y López-García, 2019).

Estos sistemas también han visto un nicho en la detección, prevención y seguimiento del discurso de odio. Se trata de un importante paso para proteger a los usuarios y prevenir la violencia motivada por ese sentimiento de animadversión frente a una persona o colectivo por motivos étnicos, racistas, orientación sexual e identidad de género o discapacidad, entre otros. En España, el Observatorio Español del Racismo y la Xenofobia (OBERAXE) ha lanzado el Proyecto Real-Up que tiene como objetivo mejorar las capacidades de las autoridades para analizar, supervisar y evaluar el discurso de odio en línea. El OBRAXE también ha colaborado con la Oficina Nacional de lucha contra los Delitos de Odio (ONDOD) para desarrollar investigación en la Generación Automática de Contranarrativas empleando grandes modelos de lenguaje[9]. En 2020 OBERAXE también utilizó la IA para identificar dicursos de odio en redes sociales contra personas asiáticas a raíz de la pandemia del COVID-19. Más recientemente, el pasado mes de septiembre de 2023, la LIGA presentó MOOD (Monitor para la Observación del Odio en el Deporte). Se trata de una herramienta de IA encargada de identificar y auditar el odio y el racismo que se genera en redes sociales en el ámbito del fútbol. Una herramienta externa que rastrea las plataformas para mostrar semanalmente las métricas registradas[10].



Si bien es cierto que estos mecanismos inteligentes son desarrollados con la mejor de las intenciones, también inciden directamente en la protección de los datos personales de los usuarios. Cuando los datos recopilados a través de herramientas de monitorización caen en manos de terceros, existe el riesgo de que la información personal de los usuarios sea explotada para fines que exceden el control del discurso de odio. Ello puede conducir, además, a un efecto de autocensura donde los usuarios se sienten vigilados y prefieren no expresar sus opiniones para evitar cualquier tipo de sanción. Téngase en cuenta que nos situamos en esa delgada línea que existe entre la opinión amparada por la libertad de expresión y la conducta típica que comporta el discurso de odio, lo cual será particularmente problemático en contextos de activismo o denuncias de injusticia social. Asimismo, los datos con los que se entrena el algoritmo dificultan la posibilidad de la tecnología de identificar la ironía o el contexto en que se interpreta ese mensaje, pudiendo calificarse entonces de discurso de odio una opinión que se enmarca en el ejercicio del derecho a la libertad de expresión. Así, existirán grupos que se

9. Información disponible en: https://blogs.ujaen.es/maite/?p=796

10. Más información sobre el sistema disponible en: https://www.laliga.com/noticias/laliga-presenta-mood-un-sistema-de-monitorizacion-del-odio-en-redes-sociales

IUS ET SCIENTIA • 2024
Vol. 10 • N° 2 • pp. 80-100
ISSN 2444-8478 • https://dx.doi.org/10.12795/IESTSCIENTIA.2024.i02.04

verán estigmatizados porque históricamente se han calificado de opresores aunque la situación actual sea completamente distinta.

## 3.3. El derecho a la tutela judicial efectiva

En el seno de una norma, como es la CE, caracterizada por su garantismo en lo que se refiere al tratamiento, desde la óptica constitucional, de los sistemas de protección de los derechos y libertades, el derecho a la tutela judicial efectiva cobra un especial protagonismo pues implica que, por primera vez, se reconocen un conjunto de derechos y garantías procesales cuyo ejercicio se circunscribe a los procedimientos que se ventilan en el seno de la jurisdicción ordinaria (Ruiz-Rico Ruiz y Carazo Liébana, 2013). El derecho a la tutela judicial efectiva es amplio y de contenido complejo que, de manera sucinta comprende el derecho de acceso a los órganos jurisdiccionales, a obtener una resolución motivada y fundada en derecho, a la efectividad de las resoluciones judiciales y el derecho al recurso (Picó i Junoy, 2012).

En las últimas décadas, se ha venido poniendo de manifiesto la falta de eficiencia en el seno de la administración de justicia lo que ha conducido a la necesidad de cambio en el derecho a la tutela judicial efectiva de modo que discurra, de forma paralela, a la transformación en el modelo de justicia (Martín Diz, 2019). En el contexto de la IA, la incorporación de esta tecnología en el sistema judicial plantea una serie de oportunidades y retos importantes en aras de asegurar el mantenimiento y garantismo del derecho a la tutela judicial efectiva, configurándose como sistemas con un potencial de mejora en este aspecto. En todo caso debemos tener presente que la eficiencia del sistema judicial a través de la incorporación de la IA resulta imprescindible para la cultura de paz (Fontestad Portalés, 2023). Así algunas de las aplicaciones de esta herramienta incluyen, en primer lugar, la automatización de trámites burocráticos facilitando la gestión de expedientes, asignación de casos y programación de audiencias sorteando la posible concurrencia de dilaciones indebidas (Nieva Fenoll, 2022). En segundo lugar, la asistencia en la toma de decisiones al permitir el análisis de grande volúmenes de datos y jurisprudencia relacionada con el asunto, lo que permitirá que los operadores jurídicos tengan un acceso a mayor cantidad de datos y, por ende, puedan tomar decisiones más fundamentadas. En tercer lugar, la IA ha sido empleada en el seno de la resolución extrajudicial de conflictos al existir sistemas de resolución alternativa de disputas como *chatbots* de mediación o conciliación.

La incorporación de sistemas predictivos y herramientas inteligentes, consideramos, deben ser empleados como un instrumento de carácter asistencial para los jueces y magistrados, si bien, defendemos la idea de que la función de juzgar y hacer ejecutar lo juzgado, es decir, la potestad jurisdiccional, debe ser desempeñada tal y como se recoge en la CE y en la LOPJ por jueces humanos no por jueces robot. Por tanto, la falta de capital material, humano y económico puede ser cubierta por una herramienta de IA, pero siendo plenamente conscientes de las garantías y derechos de las partes como es, precisamente, el derecho a la tutela judicial efectiva. La sustitución del juez natural conculcaría el concepto de función jurisdiccional y el reconocimiento del derecho a un juez imparcial y predeterminado por la ley (Bueno de Mata, 2020). En todo caso, la utilización de estos mecanismos como instrumento de asistencia al juez no puede realizarse



IUS ET SCIENTIA • 2024
Vol. 10 • Nº 2 • pp. 80-100
ISSN 2444-8478 • https://dx.doi.org/10.12795/IESTSCIENTIA.2024.i02.04

libremente y sin limitación alguna dados los desafíos éticos y jurídicos que plantean su uso. Es importante remarcar el problema de los sesgos algorítmicos, esto es, la IA se entrena a raíz de una serie –muy amplia– de datos históricos que pueden contener sesgos de índole racial, de género o socioeconómicos, de modo que la falta de eliminación de estos sesgos podría comportar la emisión de una recomendación que perpetúe desigualdades comprometiendo así la igualdad ante la ley y la imparcialidad del juzgador. Asimismo, puede concurrir una falta de transparencia y explicabilidad, dado que los algoritmos son complejos y difíciles de entender por los operadores de justicia y los propios ciudadanos, pues no dejan de crearse a través de difíciles fórmulas matemáticas[11]. La falta de entendimiento o el cuestionamiento de cómo una IA puede llegar una determinada conclusión sin motivación alguna vulneraría, nuevamente, el derecho a la tutela judicial efectiva y el derecho al debido proceso (Cancio Fernández, 2020)). El auxilio prestado por el sistema será un condicionante no determinante de la resolución del juez y en este sentido, la alegación de indefensión por falta de explicabilidad en los elementos en que se basa el algoritmo, podría ser equivalente a que el encargado de desarrollar el informe no solo alegara las razones por las que es expertos, sino todos y cada uno de los elementos que le han llevado a serlo (Castellanos Claramunt y Montero Caro, 2020). Téngase en cuenta, asimismo, que la función del juzgador no se limita exclusivamente a juzgar y hacer ejecutar lo juzgado, sino que en el marco de esa labor también deberá velar por el aseguramiento de las garantías y derechos de los partes durante todo el proceso pues la resolución final, ya sea condenatoria o absolutoria, no deja de ser el resultado un largo desarrollo de tareas que difícilmente pueden ser estandarizadas por una máquina.



### 3.4. La posible conculcación del derecho a la igualdad

Otro de los derechos que también puede ser víctima de la incorrecta configuración y utilización de sistemas de IA es el derecho a la igualdad. El papel clave del marco de los derechos humanos en la regulación ética y legal de la IA es una prioridad que centra su foco de atención en la igualdad y no discriminación en la tarea del aprendizaje automático de los algoritmos (Grigore, 2022).

De nuevo, hacemos mención del problema de los sesgos en los algoritmos, los cuales se nutren de una serie de datos que, en teoría, deberían ser neutrales. Téngase en cuenta que los seres humanos presentamos opiniones, ciertos valores y prejuicios que de una manera u otra pueden incorporarse en los sistemas de IA (Blanco García, 2024). Así los datos usados en el entrenamiento y que después se incorporan en la decisiones judiciales, como pueden ser estadísticas de criminalidad o perfiles de riesgo, pueden verse influenciados por factores de género, raciales, étnicos o socioeconómicos. Al utilizar posibles datos sesgados, la IA puede amplificar patrones discriminatorios y dar lugar a decisiones que afectan de manera desigual a ciertos grupos, como minorías

11. La dificultad en las operaciones de su desarrollo puede dar lugar a «cajas negras de datos» en las que el proceso de toma decisiones no sea comprensible. Es esencial que todas las partes comprendan cómo y por qué el algoritmo llegó a tomar esa decisión haciendo así efectivo el derecho a recurrir la decisión tomada.

IUS ET SCIENTIA • 2024
Vol. 10 • Nº 2 • pp. 80-100
ISSN 2444-8478 • https://dx.doi.org/10.12795/IESTSCIENTIA.2024.i02.04

raciales, migrantes o personas en riesgo de exclusión social. Algunos de los sistemas utilizados realizan evaluaciones de riesgo como, por ejemplo, la probabilidad de reincidencia o el riesgo de fuga. Para estos cálculos se suelen emplear variables que, aunque aparentemente pueden parecer neutrales –un nivel educativo, situación laboral o un simple código postal–, están correlacionadas con otros factores personales como la etnia, el nivel económico o la orientación sexual. Estas evaluaciones podrían conducir a una discriminación indirecta donde ciertas personas pertenecientes a uno u otro grupo sean consideradas de alto riesgo dadas sus características demográficas. Así, a la hora de adoptar una u otra medida cautelar ante un sujeto, la persona podría verse discriminada por residir en un barrio «de clase obrera» o en riesgo de exclusión social. Cuanto mayor sea el riesgo para los derechos humanos de la persona, más estrictos deberán ser los requisitos legales para el empleo de la IA[12]. También en la fase de instrucción, la IA empleada por las autoridades para analizar patrones de comportamiento o identificación de sospechosos pueden llevar a que se determine de manera desproporcionada a personas de ciertos grupos o áreas geográficas. Esta focalización en ciertos núcleos de población, por ejemplo, socava la igualdad al poner bajo sospecha grupos concretos por factores externos reforzando de manera injustificada su estigmatización.

Hemos de tener presente, de igual modo, la falta de transparencia y la posible dificultad para poder detectar la discriminación. La posible opacidad del algoritmo hace difícil la detección de cuándo y cómo se genera la discriminación. La falta de explicabilidad de las decisiones de IA complica que la persona afectada pueda entender por qué se ha tomado una decisión en su contra o argumentar que se trata de una discriminación. Ello incide directamente en el derecho a la igualdad, pues sin transparencia no es posible saber si la IA ha actuado de manera sesgada o si la decisión tomada por el juez es realmente imparcial y equitativa. A ello debemos unir la exclusión de grupos vulnerables en el acceso a la justicia. Hoy en día existen dificultades e importantes obstáculos para acceder a la justicia cuando estamos ante colectivos considerados especialmente vulnerables[13] al limitar y restringir el acceso a tales grupos a la vez que se ve coartado el ejercicio de los derechos de que son titulares (Carrizo González-Castell, 2019). La instauración de sistemas de IA en la administración justicia a menudo requiere un nivel básico de acceso y familiaridad con las tecnologías. No obstante, las personas que precisamente se hallan en situación de vulnerabilidad carecen o ven limitado el acceso a estos dispositivos. La brecha digital presenta, como principal consecuencia, la desigualdad en el acceso a la información y los servicios judiciales que están basados en IA, creando así una barrera que impide una plena participación en el proceso judicial. De nuevo las decisiones algorítmicas de IA pueden generar una discriminación en estos colectivos por ser considerados de alto riesgo sin que concurra razón válida alguna.



12. Véase en este sentido el Informe de Naciones Unidas del Alto Comisionado para los Derechos Humanos en relación con el derecho a la privacidad en la era digital, elaborado en Ginebra el 15 de septiembre de 2021 (A/HRC/48/31). Texto disponible en: https://www.ohchr.org/es/documents/thematic-reports/ahrc4831-right-privacy-digital-age-report-united-nations-high

13. Véase aquí el Documento de las 100 Reglas de Brasilia sobre el acceso a la justicia de las personas en condición de vulnerabilidad (actualizado en 2018) elaborado en el seno de la Cumbre Judicial Iberoamericana (CIJ). Texto disponible en: https://eurosocial.eu/biblioteca/doc/reglas-de-brasilia-sobre-acceso-a-la-justicia-de-las-personas-en-condicion-de-vulnerabilidad/

IUS ET SCIENTIA • 2024
Vol. 10 • Nº 2 • pp. 80-100
ISSN 2444-8478 • https://dx.doi.org/10.12795/IESTSCIENTIA.2024.i02.04

Esto, a su vez, genera desconfianza en estos grupos al percibir que el sistema no es imparcial o que no ofrece suficiente transparencia.

Tras los expuesto, y de acuerdo con las previsiones contenidas en el art. 23 de la Ley 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación, las administraciones que pongan en marcha mecanismos de algoritmos involucrados en la toma de decisiones deberán tener en cuenta los criterios de minimización de los sesgos, la transparencia y rendición de cuentas. La transparencia y la capacidad de interpretación en la toma de decisiones deberán operan como una prioridad en el diseño del algoritmo evitando cualquier conculcación de los derechos fundamentales y, especialmente, del derecho de igualdad

## 3.5. El compromiso de la presunción de inocencia

Todo sistema de justicia penal presenta como elemento clave la presunción de inocencia[14], pues el sometimiento de una persona como sujeto pasivo del proceso penal conduce a su señalamiento como sospechoso y, por ende, puede generar de manera automático un rechazo social hacia esa persona (Nieva Fenoll, 2016). Su reconocimiento en el art. 24.2 CE hace que opere en el proceso como una regla de juicio[15], esta despliega sus efectos en el momento en que tiene lugar la valoración de la prueba ligada tanto a la estructura del proceso como a la constatación del hecho probado. La presunción de inocencia supone, pues, que el investigado o encausado haya de ser tratado en todo momento como inocente hasta que exista sentencia firme que demuestre lo contrario, constituyendo entonces una garantía que ha de respetarse y estar presente durante todo el proceso penal y, por lo tanto, en todas las instancias.

Dicho esto, de nuevo, los parámetros del algoritmo de la IA pueden incorporar unas valoraciones que podría afectar a la presunción de inocencia (Schumann Barragán, 2021) y la implementación de esta tecnología no puede comportar restricciones y mucho menos menoscabar este derecho. En la actualidad existen diversas herramientas que pueden afectar a este derecho. A modo de ejemplo, encontramos las evaluaciones de riesgo de reincidencia, los sistemas de IA que se utilizan para evaluar el riesgo de reincidencia del investigado pueden influir en la toma decisiones sobre la adopción de medidas cautelares como la prisión provisional. Estas evaluaciones a menudo se basan en datos históricos que pueden reflejar sesgos geográficos o raciales. A modo de ejemplo en Estados Unidos existe el sistema COMPAS (*Correctional Offender Mangament for Alternative Sanctions*) un software que se dedica a evaluar el riesgo de reincidencia de la persona acusada y posteriormente aconseja al juzgador sobre el tipo de pena aplicable y su duración (Borges Blázquez, 2020). ProPublica realizó un estudio sobre



14. La importancia de este derecho se deriva de su reconocimiento en los instrumentos internacionales. Véase en este sentido el art. 11.1 de la Declaración Universal de Derechos Humanos (DUDH) en virtud de cual toda persona acusada de un delito tiene derecho a que se presuma su inocencia mientras no se demuestre su culpabilidad. Así, se recoge de igual modo en el CEDH, en el Pacto Internacional de Derechos Civiles y Políticos (PIDCP) o en el art. 48 de la Carta de los Derechos Fundamentales de la Unión Europea.

15. Véase la STC 128/1995, de 16 de julio, Rec. 993/1995.

IUS ET SCIENTIA • 2024
Vol. 10 • Nº 2 • pp. 80-100
ISSN 2444-8478 • https://dx.doi.org/10.12795/IESTSCIENTIA.2024.i02.04

COMPAS donde advirtió una posible discriminación algorítmica hacia personas afrodescendientes, lo que conllevaría un riesgo de violación de la presunción de inocencia (Roa Avella; Sanabria-Moyano y Dinas Hurtado, 2022). Estamos así frente a una herramientas que, en numerosas ocasiones, operan como instrumentos de «pre-criminalización» que no siempre reflejan la intención o culpabilidad real de un individuo.

Queda así demostrado que el uso de la IA en el proceso penal plantea significativos riesgos para la presunción de inocencia, un derecho clave en el sistema de administración de justicia penal. La naturaleza de la IA y las decisiones basadas en patrones y estadísticas pueden predisponer a las autoridades a percibir, incorrectamente, a individuos como culpables o peligrosos incluso sin pruebas concluyentes que permitan desvirtuar la citada presunción. Por ello es esencial que el uso de la IA en el sistema penal esté sujeto a un regulación estricta que garantice la transparencia y el acceso a la información. Siendo igualmente relevante la inclusión de mecanismos, siempre, de supervisión humana que permitan revisar y poner en duda los resultados algorítmicos, velando así por los derechos fundamentales y las garantías procesales del sujeto pasivo del proceso.

## 3.6. El derecho de defensa ante el uso de la Inteligencia Artificial

Por último, no debemos pasar por alto el riesgo que se deriva de utilizar IA en el proceso penal en relación con el derecho de defensa. Este es considerado como un derecho público constitucional que se reconoce a toda persona física a la que se imputa la comisión de un hecho delictivo y en virtud del cual se le concede la posibilidad de oponerse a la pretensión punitiva (Zafra Espinosa de los Montero, 2014). Estamos ante un derecho que, junto con la tutela judicial efectiva, comporta uno de los derechos básicos de la protección ciudadana y que está íntimamente relacionado con el Estado de Derecho. De acuerdo con el Proyecto de Ley Orgánica del Derecho de Defensa[16] que verá la luz en los próximos días, tras haber sido aprobado por el Congreso el 30 de octubre de 2024, el derecho de defensa comprende la asistencia letrada y el asesoramiento jurídico, el acceso a los tribunales de justicia, la ausencia de dilaciones indebidas y la obtención de una resolución fundada en derecho y, concretamente en el escenario del proceso penal el derecho a ser informado de los hechos que se atribuyen, a no declarar contra uno mismo, a no declararse culpable, a la presunción de inocencia y a la doble instancia.

Cuando hablamos del derecho de defensa, debemos pensar también en el principio de igualdad de armas, que debe mantenerse durante todo el proceso evitando cualquier resultado lesivo o pérdida del derecho que se reconoce al justiciable, ha de tener las mismas posibilidades que la parte contraria a la hora de presentar el material probatorio para su defensa y el acceso a los mismos recursos (Rodríguez del Blanco, 2024). La especial consideración e importancia de este derecho nos lleva a plantearnos cuáles son los riesgos que derivan del uso de la IA en el proceso penal. En primer lugar, pensamos en la dificultad que puede concurrir para impugnar lo que podríamos denominar como «evidencia algorítmica». Partiendo de la base del principio de contradicción, entendemos

16. Texto disponible en: https://www.congreso.es/public_oficiales/L15/CONG/BOCG/A/BOCG-15-A-6-4.PDF



IUS ET SCIENTIA • 2024
Vol. 10 • Nº 2 • pp. 80-100
ISSN 2444-8478 • https://dx.doi.org/10.12795/IESTSCIENTIA.2024.i02.04

que el uso de algoritmos para recolectar o analizar la evidencia en las redes sociales, dispositivos tecnológicos o en los sistemas de reconocimiento facial puede afectar a la capacidad de la defensa para su impugnación. Una incorrecta configuración del algoritmo puede dar lugar a la generación de falsos positivos en los que, si una adecuada supervisión de «inteligencia humana», la persona puede verse perjudicada. De nuevo, insistimos en que las tareas de reconocimiento facial existen importantes tasas de error que son más altas para ciertos grupos demográficos, al igual que ocurre con los sistemas de policía predictiva. En este sentido, si la defensa intenta cuestionar una identificación o una evidencia obtenida por la IA, puede encontrarse con importantes barreras técnicas y limitaciones que ponen en peligro el principio de igualdad de armas y, por ende, el derecho de defensa.

En segundo lugar y, vinculado a los expuesto en las líneas anteriores, el uso de la IA exige una comprensión técnica profunda para poder cuestionar la evidencia planteada *supra*. La defensa, en muchas ocasiones, puede no contar con los recursos necesarios suficientes para poder llegar a expertos en peritaje técnico dando lugar así a una desigualdad en el proceso, pues el Estado que ostenta los recursos para utilizar la IA se encontrará siempre en una posición superior en cuanto al acceso a los recursos. En tercer lugar, la percepción de que los algoritmos son infalibles puede llevar a que la autoridad judicial confíe ciegamente en sus resultados generándose así un «prejuicio automatizado». Así partimos de un sesgo institucional cuando se considera que el resultado de un algoritmo es más confiable que las declaraciones prestadas por testigos o por el propio investigado o posterior encausado. Este sitúa a la defensa letrada en una posición en la que sus argumentos son infravalorados frente a los resultados algorítmicos. En cuarto lugar, en el marco del derecho de defensa es esencial el reconocimiento y puesta en marcha del acceso a las actuaciones. Este derecho de acceso a las actuaciones del sospechoso se erige como una de las manifestaciones básicas del derecho a un proceso justo y del derecho de defensa[17]. Este derecho también comporta así la puesta en conocimiento de elementos vinculados a la investigación y de cualquier cambio que se produzca en relación con la misma y también el derecho a examinar las actuaciones. Este último aspecto debe entenderse como el acceso a todos los materiales, documentos, grabaciones, vídeos o fotografías (Muerza Esparza, 2023). Así pues, en caso de obtener elementos que se derivan del uso de algoritmos complejo y técnicas de *deep learning* puede surgir la posibilidad de que sean de difícil interpretación. Si la defensa no puede acceder a los datos y métodos que hay detrás de una decisión algorítmica, es complicado que se pueda a ver valer el derecho a cuestionar su pertinencia.



## IV. Una propuesta de buenas prácticas para la implementación de la Inteligencia Artificial en el proceso penal

Expuestos los riesgos que implica la utilización de la IA en los sistemas de justicia penal es necesario implementar estrategias y regulaciones que permitan asegurar la

17. Véase la Circular 3/2018, de 1 de junio, de la Fiscalía General del Estado sobre el derecho de información de los investigados en los procesos penales.

IUS ET SCIENTIA • 2024
Vol. 10 • Nº 2 • pp. 80-100
ISSN 2444-8478 • https://dx.doi.org/10.12795/IESTSCIENTIA.2024.i02.04

transparencia, imparcialidad y supervisión humana en cada fase del proceso. Así sera crucial que cualquier sistema de IA utilizado en el proceso penal sea transparente y explicable de modo que todos los involucrados, es decir, las partes, jueces y letrados, puedan comprender como se generan la decisiones. Igualmente será fundamental crear estándares de explicabilidad y transparencia que obliguen a las empresas y entidades que desarrollan el software a explicar el funcionamiento y *modus operandi* de sus algoritmos. Estos sistemas deben poder justificar cada resultado o recomendación de manera comprensible, especialmente en decisiones sensibles como puede ser la adopción de medidas cautelares y, más concretamente, en aquellas que limiten derechos fundamentales como es la libertad personal (Muñoz Rodríguez, 2020).

Para que la IA sea una herramienta justa y equitativa, la defensa debe poder acceder a la información y datos utilizados por el sistema. En este sentido, el legislador se enfrenta al reto de crear un marco normativo que permita este derecho, concretamente en relación con los datos y criterios que ha utilizado la IA. Se debe exigir entonces a los proveedores de IA en el sector público que cumplan con los requisitos de acceso a la información, permitiendo auditorías y otros análisis de técnicos expertos en la materia. Asimismo, los datos históricos con que se entran a los algoritmos deberán ser anonimizados para cumplir con las regulaciones sobre privacidad.

No menos importante es, insistimos, la configuración de herramientas de evaluación de impacto del riesgo y sesgos algorítmicos, los cuales afectan a los principios de equidad y no discriminación de las decisiones. Para ello consideramos que es necesaria la realización de evaluaciones de carácter periódico del impacto de la IA, detectando y eliminando los sesgos relacionados con la raza, la orientación sexual o identidad de género, la situación económica o las creencias políticas y religiosas. Evidentemente, estas evaluaciones deben realizarse antes de poner en marcha cualquier sistema de IA y, en todo caso, deberán ser públicas y accesibles. Quizá el legislador deba proponerse también la creación de una Agencia reguladora, similar a la Agencia Española de Protección de Datos, que se encargue de auditar los sistemas de IA de los que se hace uso en el sistema de administración de justicia del Estado con la finalidad de requerir mejorar o suspender, temporalmente, el uso de un algoritmo si se encuentra defectuoso o sesgado. Fomentar el uso de algoritmos de código abierto en el proceso permitirá a la defensa y otros actores revisar dicho código para verificar la concurrencia o no de sesgos o errores que puedan afectar a los derechos del acusado.



El gran manejo de una cantidad ingente de datos incide directamente en la necesidad de proteger la privacidad y derechos del acusado en todo momento. El establecimiento de reglas estrictas sobre el uso, almacenamiento y eliminación de datos personales recopilados es crucial. Estos datos deberán ser utilizados para el caso en cuestión y ser protegidos frente a posibles accesos no autorizados o usos indebidos. Para ello consideramos necesaria la colaboración de la Agencia Española de Protección de Datos y la elaboración de protocolos de anonimización de los datos que también sirva de guía al usuario para hacer efectivos sus derechos de acceso o eliminación en caso necesario. Por último, consideramos útil la creación de una comisión ética que se encargue de la supervisión de la AI en el sistema de justicia penal en aras de verificar y garantizar que su aplicación respeta los principios del proceso y del procedimiento, así como los derechos y garantías procesales. Hablaríamos entonces de un grupo compuestos

IUS ET SCIENTIA • 2024
Vol. 10 • N° 2 • pp. 80-100
ISSN 2444-8478 • https://dx.doi.org/10.12795/IESTSCIENTIA.2024.i02.04

multidisciplinar integrado por expertos en derecho, tecnología y ética con autoridad para revisar, aprobar y monitorear la utilización de los sistemas de AI emitiendo recomendaciones. Sobre este último aspecto, debemos tener presente, la Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas. En ella se destaca la necesidad de que los programas de IA y las tecnologías conexas respeten los principios de necesidad y proporcionalidad y que todas aquellas consideradas de alto riesgo deberán regirse siempre por principios éticos y estar generadas para respetar y permitir, en todo caso, la intervención humana y el control democrático de los Estados siendo posible la recuperación del control humanos cuando sea necesario (San Miguel Caso, 2021). Ello comporta, igualmente, la necesidad de seguir un enfoque basado en el riesgo y orientado al futuro, tal y como se plasma en el RIA de 2024. Con la finalidad de garantizar ese sistema de evaluación de riesgo será necesaria, entonces, una lista exhaustiva sobre los sectores de alto riesgo y de los fines del mismo.

## V. Conclusiones

El uso de la IA es ya una realidad, una tecnología que ha llegado para quedarse en todos los aspectos de nuestro entorno social, privado, laboral y jurídico. En un sistema con oficinas judiciales cada vez más saturadas, con falta de personal, se genera una especial desconfianza y decepción por parte de la ciudadanía. Su uso en la administración de justicia, en general, y en el proceso penal, en particular, ofrece un potencial más que significativo para incrementar la eficiencia del sistema judicial optimizando así los recursos. Las herramientas de IA como los sistemas predictivos de reincidencia o análisis masivo de datos que se encuentran en funcionamiento tienen como fin servir de apoyo a las autoridades judicial en la toma decisiones más informadas y coherentes. Ahora bien, la implementación de estos sistemas también supone importantes desafíos para los derechos fundamentales y las garantías procesales como se ha puesto de manifiesto y, más concretamente, en la tutela judicial efectiva, la igualdad o el derecho de defensa. La opaca naturaleza de los algoritmos utilizados, junto con el posible sesgo inherente en los datos, podría generar decisiones desproporcionadas y faltas de justificación que afecten de manera desigual a determinados grupos sociales.



Para poder lograr el equilibrio entre la eficiencia y las garantías procesales necesitaremos de marcos normativos sólidos que regulen el uso de la IA en el ámbito del sistema penal. La entrada en vigor del Reglamento de IA ha supuesto un importante avance, no obstante, la incidencia de estos sistemas –teniendo en cuenta el ordenamiento constitucional español– en los derechos fundamentales precisarán de una regulación a través de leyes orgánicas para poder preservar tales derechos. Insistimos en la necesidad de cumplir con los criterios de explicabilidad y transparencia y, sobre todo, en la concurrencia de supervisión humana. Por ello será necesario que los operadores comprendan tanto las capacidades como las limitaciones de esta tecnología de tal modo que nuevos programas de formación en esta materia comenzarán a asentarse.

En conclusión, la inteligencia artificial puede ser una herramienta poderosa en el proceso penal español, pero solo si su uso se ajusta a un enfoque ético y garantista

IUS ET SCIENTIA • 2024
Vol. 10 • N° 2 • pp. 80-100
ISSN 2444-8478 • https://dx.doi.org/10.12795/IESTSCIENTIA.2024.i02.04

que priorice los derechos fundamentales sobre la mera eficiencia. Un modelo regulado, transparente y supervisado de IA, en el que se valoren tanto los beneficios como los riesgos, permitiría alcanzar un equilibrio entre eficacia y garantías procesales. La integración de la IA en el ámbito penal debe estar guiada por el compromiso de mantener un sistema de justicia humano, donde el respeto por la dignidad, la igualdad y los derechos individuales prevalezca sobre cualquier avance tecnológico.

Necesitaremos, igualmente un compromiso de los Estados y también del poder legislativo, el cual será clave para el establecimiento de un marco normativo que regule el uso. El papel del legislador habrá de ir más alla del simple permiso a la hora de usar estas herramientas, es decir, habrá de establecer en todo momento las concretas salvaguardias de los derechos fundamentales y el aseguramiento de un uso ético y responsable. Solo mediante un marco claro y garantista se permitirá que la IA sea una herramienta judicial que no comprometa las garantías del sistema de justicia penal que tanto ha costado construir, pues un sistema que únicamente persiga la eficiencia y agilidad correrá el riesgo de perder su elemento más importante: la acción humana.

## Bibliografía

Alonso Salgado, C. (2015). «El largo camino hasta la Ley Orgánica 13/2015: algunos aspectos relevantes en relación a la interceptación de las comunicaciones», en F. Bueno de Mata (Coord.), *FODERTICS 4.0. Estudios sobre nuevas tecnologías y justicia*, Comares, pp. 95-105.

Asencio Mellado, J.M. (2008). *Introducción al Derecho Procesal (5ª edición)*, Tirant lo Blanch.

Barona Vilar, S. (2021). *Algoritmización del Derecho y de la Justicia. De la Inteligencia Artificial a la Smart Justice*, Tirant lo Blanch.

Barona Vilar, S. (2022). «Justicia algorítmica, ¿más o menos sostenible?», en P. Arrabal Platero (Dir), *Los objetivos de desarrollo sostenible y la Inteligencia Artificial en el proceso judicial*, Tirant lo Blanch, pp. 227-256.

Barrio Andrés, M. (2024). «Objeto, ámbito de aplicación y sentido del Reglamento Europeo de Inteligencia artificial», en M. Barrio Andrés (Dir.), *El Reglamento Europeo de Inteligencia Artificial*, Tirant lo Blanch, pp. 21-48.

Blanco García, A.I. (2024). «Retos para un inteligencia artificial inclusiva de los colectivos vulnerables», *Actualidad jurídica iberoamericana*, 21, pp. 360-383.

Borges Blázquez, R. (2020). «El sesgo de la máquina en la toma de decisiones en el proceso penal», *IUS ET SCIENTIA: Revista electrónica de Derecho y Ciencia,* (6)2, pp. 54-71.

Bueno de Mata, F. (2019). *Las diligencias de investigación penal en la cuarta revolución industrial. Principios teóricos y problemas prácticos*, Thomson Reuters Aranzadi.

Bueno de Mata, F. (2020). «Macrodatos, Inteligencia Artificial y proceso: luces y sombras», *Revista General de Derecho Procesal,* 51, pp. 1-31.

Bueno de Mata, F. (2021). «Protección de datos, investigación de infracciones penales e inteligencia artificial: novedades y desafíos a nivel nacional y europeo en la era postcovid», *La Ley Penal*, 150, pp. 1-20.

Cancio Fernández, R.C. (2020). ¿Sueñas los jueces con sentencias electrónicas?, *Revista Análisis Jurídico-Político*, (2),3, pp. 145-168.

Carrizo González-Castell, A. (2019). El acceso a la justicia de las personas en condición de vulnerabilidad. Un reto pendiente para los derechos humanos, en N. Sanz Mulas (Dir), *Los derechos humanos 70 años después de la Declaración Universal,* Tirant lo Blanch.



IUS ET SCIENTIA • 2024
Vol. 10 • Nº 2 • pp. 80-100
ISSN 2444-8478 • https://dx.doi.org/10.12795/IESTSCIENTIA.2024.i02.04

Castellanos Claramunt, J; Montero Caro, M.D. (2020). «Perspectiva constitucional de las garantías de aplicación de la inteligencia artificial: la ineludible protección de los derechos fundamentales», *IUS ET SCIENTIA: Revista electrónica de Derecho y Ciencia* (6)2, pp. 72-82.

Cinelli, V. (2019). «El uso de programas de análisis predictivo en la inteligencia policial: una comparativa europea», *Revista de Estudios en Seguridad Internacional*, 5(2), pp. 1-19.

Cotino Hueso, L. (2023). «Reconocimiento facial automatizado y sistemas de identificación biométrica bajo la regulación superpuesta de inteligencia artificial y protección de datos», en F. Balaguer Callejón (Coord.), *Derecho público de la inteligencia artificial*, Fundación Manuel Giménez Abad de Estudios Parlamentarios y del Estado Autonómico, pp. 347-402.

De Hoyos Sancho, M. (2021). «El uso jurisdiccional de los sistemas de Inteligencia Artificial y la necesidad de su armonización en el contexto de la Unión Europea», *Revista General de Derecho Procesal*, 55, pp.1-29.

De Hoyos Sancho, M. (2021), «Premisas y finalidades del Libro Blanco sobre Inteligencia Artificial de la Comisión Europea: perspectiva procesal del nuevo marco regulador», en S. Barona Vilar (Ed.), *Justicia algorítmica y neuroderecho: una mirada multidisciplinar*, Tirant lo Blanch.

De Miguel Beriain, I; Pérez Estrada, M.J. (2019), «La inteligencia artificial en el proceso penal español. Un análisis de su admisibilidad sobre la base de los derechos fundamentales implicados», 25, pp. 531-561.

Fernández-Aller, C; Serrano Pérez, M.M. (2022). «¿Es posible una inteligencia artificial respetuosa con la protección de datos?», *Doxa: Cuadernos de Filosofía del Derecho*, 45, pp. 307-336.

Fontestad Portalés, L. (2023). «Eficiencia procesal versus jurisdicción», *La Ley Actualidad Civil*, 11, pp. 1-12.

Gamero Casado, E. (2021), «El enfoque europeo de la Inteligencia Artificial», *Revista de Derecho Administrativo*, 20, pp. 268-289.

Goldschmidt, J. (2021). *Problemas jurídicos y políticos del proceso penal*, Ediciones Olejnik.

Gómez Abeja, L. (2022). «Inteligencia artificial y derechos fundamentales», en F.H. Llano Alonso, *Inteligencia artificial y Filosofía del Derecho*, Laborum, pp. 91-114.

González Pulido, I. (2023). «Perspectivas de futuro respecto a la obtención de pruebas electrónicas transfronterizas y a la cooperación con proveedores de servicios: investigación y prueba de los ciberdelitos graves en la Unión Europea», *Diario La Ley*, 10266, pp. 1-21.

Grigore, A.E. (2022), «Derechos humanos e inteligencia artificial», *IUS ET SCIENTIA: Revista electrónica de Derecho y Ciencia*, (8)1, pp. 164-175.

Martín Diz, F. (2019). «El derecho fundamental a justicia: Revisión integral e integradora del derecho a la tutela judicial efectiva» *Revista de Derecho Político*, 106, pp. 13-42.

Martín Diz, F. (2024), «Derechos y garantías procesales penales fundamentales: una lectura en clave tecnológica, *IUS ET SCIENTIA: Revista electrónica de Derecho y Ciencia*, (10) 1, pp. 52-81.

Muerza Esparza, J. (2023). «Algunas cuestiones sobre el derecho de información del investigado», 3, pp. 616-643.

Muñoz Rodríguez, A.N. (2020). «El impacto de la inteligencia artificial en el proceso penal», *Anuario de la Facultad de Derecho. Universidad de Extremadura*, 36, pp. 695-728.

Nieva Fenoll, J. (2016). «La razón de ser de la presunción de inocencia», *InDret: Revista para el Análisis del Derecho*, 1, pp. 1-23.

Nieva Fenoll, J. (2022) «Inteligencia artificial y proceso judicial: perspectivas tras un alto tecnológico en el camino» *Revista General de Derecho Procesal*, 57, pp. 1-21.

Noya Ferreiro, M.L. (2022). «Algunas consideraciones sobre Inteligencia Artificial, proceso penal y derechos fundamentales», en P. González Granda; J. Damián Moreno; M.J. Ariza Col-



IUS ET SCIENTIA • 2024
Vol. 10 • Nº 2 • pp. 80-100
ISSN 2444-8478 • https://dx.doi.org/10.12795/IESTSCIENTIA.2024.i02.04

menarejo (Dirs), *Variaciones sobre un tema: el ejercicio procesal de los derechos. Libro homenaje a Valentín Cortés Domínguez*, Colex.

Peralta Gutiérrez, A. (2021). «La necesaria regulación de la vigilancia masiva: Casos Quadrature du Net y Big Brother Watch», *Diario La Ley*, 9973, pp. 1-23.

Picó i Junoy, J. (2012), *Las garantías constitucionales del proceso*, Bosch Editor.

Richard González, M. (2016). «Conductas susceptibles de ser intervenidas por medidas de investigación electrónica. Presupuestos para su autorización», *Diario La Ley*, 8808, pp.1-18.

Roa Avella, M; Sanabria Moyano, J.E; Dinas hurtado, K. (2022). «Uso del algoritmo COMPAS en el proceso penal y los riesgos de los derechos humanos», *Revista Brasileira de Direito Processual Penal*, (8)1, pp. 275-310.

Rodríguez Andrés, R; López-García, J.M. (2019). «Aproximación al uso de las redes sociales por las fuerzas y cuerpos de seguridad en España en perspectiva internacional, *Index. Comunicación*, (9), 1, pp. 127-148.

Rodríguez del Blanco, A. (2024). «Detectando los riesgos de la Inteligencia Artificial en la instrucción penal», *Revista General de Derecho Procesal*, 64, pp. 1-66.

Ruiz Forns, A; Nicolás, A. (2024). «Nuevo Reglamento Europeo de Inteligencia Artificial», *Diario La Ley*, 10491, pp. 1-3.

Ruiz-Rico Ruiz, G; Carazo Liébana, M.J. (2013). *El derecho de la tutela judicial efectiva. Análisis jurisprudencial*, Tirant lo Blanch.

San Miguel Caso, C. (2021). «La aplicación de la Inteligencia Artificial en el proceso: ¿un nuevo reto para las garantías procesales?», *IUS ET SCIENTIA: Revista electrónica de Derecho y Ciencia*, (7)1, pp. 286-303.

San Miguel Caso, C. (2023). «Inteligencia Artificial y algoritmos: la controvertida evolución de la tutela judicial efectiva en el proceso penal», *Estudios Penales y Criminológicos*, 44(ext), pp. 1-23.

Schumann Barragán, G. (2021) «La inteligencia artificial aplicada al proceso penal desde la perspectiva de la UE», en S. Pereira Puigvert; F. Ordoñez Ponz (Dirs.), *Investigación y proceso penal en el Siglo XXI. Nuevas tecnologías y protección de datos*, Thomson Reuters Aranzadi, pp. 517-540.

Zafra Espinosa de los Monteros, R. (2014). «Sobre el derecho de defensa en la mediación penal», en V.C Guzmán Fluja; I. Flores Prada (Dirs), *Justicia penal y derecho de defensa*, Tirant lo Blanch.



IUS ET SCIENTIA • 2024
Vol. 10 • Nº 2 • pp. 80-100
ISSN 2444-8478 • https://dx.doi.org/10.12795/IESTSCIENTIA.2024.i02.04