

Monográfico: «Digitalización y algoritmización de la justicia» (coord.: F. Miró)

Justicia, inteligencia artificial y derecho de defensa

Elena de Luis García
Universitat de València

Fecha de recepción: junio 2023
Fecha de aceptación: octubre 2023
Fecha de publicación: noviembre 2023

Resumen

La inteligencia artificial es una realidad de la que el derecho procesal no puede escapar. Son indudables las ventajas que presentan en términos de coste y tiempo para la Administración de Justicia, e incluso en determinadas ocasiones pueden contribuir a alcanzar un resultado más objetivo. Sin embargo, tan visibles son sus beneficios como los potenciales riesgos que entrañan para los derechos y las garantías fundamentales del proceso. Características como la opacidad o la falibilidad pueden afectar irremediablemente al derecho de defensa y la contradicción. Ante esta situación, debemos plantearnos qué futuro queremos, es decir, si la justicia algorítmica es imparable, cómo queremos que se configure en el proceso. Esto solo podrá hacerse a través del establecimiento de un marco legal que identifique los riesgos existentes y recoja medidas para paliarlos. Todo ello con el objetivo de conjugar eficiencia y eficacia, de modo que podamos gozar de las ventajas que nos ofrecen estas nuevas herramientas, pero sin que ello conlleve una merma en los derechos fundamentales y garantías del proceso, pues estos son y deben ser un límite irrebutable en la utilización de sistemas de inteligencia artificial en el proceso.

Palabras clave

inteligencia artificial; algoritmos; justicia predictiva; derecho de defensa; contradicción

Justice, artificial intelligence and right of defence

Abstract

Artificial intelligence is a reality that procedural law cannot escape. The advantages they present in terms of cost and time for the Administration of Justice are undoubted, and even on certain occasions, they can contribute to achieving a more objective result. However, their benefits are as visible as the potential risks to the fundamental rights and safeguards of the process. Characteristics such as opacity or fallibility can irremediably affect the right of defense and contradiction. In this situation, we must consider what future we do want, that is if algorithmic justice is unstoppable, how we want it to be set up in the process. This can only be done by establishing a legal framework that identifies existing risks and proposes measures to alleviate them. All this with the aim of combining efficiency and effectiveness, so that we can enjoy the advantages these new tools offer us, but without this leading to a reduction in the fundamental rights and guarantees of the process, since these are and should be a limit never exceeded in the use of artificial intelligence systems in the process.

Keywords

artificial intelligence; algorithms; predictive justice; right of defence; contradiction

Introducción

La irrupción de la inteligencia artificial (en adelante, IA) en el proceso judicial no es un hecho producido de forma aislada, sino que forma parte del proceso de modernización de la justicia que lleva décadas produciéndose. Y es que, como señala Barona Vilar, «el retrato de la sociedad algorítmica integra igualmente el modelo de Justicia y sus valores, principios y estructuras guardan una estrecha simbiosis con la sociedad que vivimos», de modo que «la irrupción de la tecnología y la digitalización social empapa la Justicia» (Barona Vilar, 2022a, pág. 39).

Así pues, si hace unos pocos años se discutía sobre la admisibilidad de la prueba electrónica o las declaraciones por videoconferencia, ahora nos encontramos en un estadio posterior en la evolución digital (o, si se quiere, revolución digital) del proceso judicial. Tanto es así que las discusiones doctrinales ahora versan sobre si puede el juzgador adoptar una decisión con apoyo de una máquina, si podemos valorar riesgos de reincidencia a partir de algoritmos e incluso, como última instancia, si sería admisible la figura de una suerte de «juez robot». Podemos aquí adelantar que la doctrina mayoritaria se inclina por rechazar que una máquina pueda tomar la decisión de forma autónoma, para sostener que la IA debería mantener un papel asistencial

y no sustitutorio de la función judicial (Martín Diz, 2021b, págs. 80-83). Lo que no obsta para que en un futuro quizá pudiera plantearse esta función decisoria para tareas específicas, casos de menor complejidad o asuntos repetitivos, entre otros (Guzmán Fluja, 2021, pág. 33).

En cualquier caso, lo innegable es que esta clase de herramientas están irrumpiendo en el sistema de justicia y es labor de los juristas examinarlas y, por lo que concierne al derecho procesal, tratar de determinar de qué modo pueden verse afectados los principios y las garantías esenciales del proceso. Y es que, aun cuando la afectación no lo sea en sentido negativo, lo que es indudable es que algunos principios -tanto del proceso, como del procedimiento- sufrirán una importante transformación como consecuencia de la implementación de estos sistemas (Barona Vilar, 2022b, págs. 390-421).

A partir de tales premisas, el objeto del presente artículo es realizar una aproximación al impacto que el empleo de herramientas de IA en el sistema de justicia puede tener sobre el derecho de defensa y la contradicción, como pilares esenciales en la configuración del proceso judicial. Para ello, vamos a articular el estudio a partir de las siguientes ideas: en primer lugar, el impacto que la utilización de los sistemas de IA puede tener en los derechos y las garantías

procesales con carácter general y, en segundo lugar, cómo puede verse afectado el derecho de defensa en particular.

1. Inteligencia artificial y garantías procesales

1.1. La irrupción de la IA en la justicia

El proceso de irrupción y expansión de las tecnologías digitales en el sistema de justicia se lleva a cabo en dos niveles distintos. El primero, referido a la «tecnología instrumental» a través de herramientas, como el expediente judicial electrónico o las notificaciones telemáticas, entre otras, y, el segundo, referido a la «tecnología funcional», en la que tienen cabida las herramientas algorítmicas como complemento o sustitución de la función humana (Barona Vilar, 2022a, pág. 39).

Es en este segundo nivel donde se ubican las herramientas de IA. Proporcionar una definición de inteligencia artificial entraña grandes dificultades, pues no existe tampoco unanimidad doctrinal al respecto. Sin embargo, por lo que al presente concierne, adoptaremos la definición general contenida en las Directrices éticas para una IA fiable, elaboradas por un grupo de expertos de alto nivel en IA constituido por la Comisión Europea en 2018,¹ según la cual los sistemas de IA son: «sistemas de *software* (y en algunos casos también de *hardware*) diseñados por seres humanos que, dado un objetivo complejo, actúan en la dimensión física o digital mediante la percepción de su entorno a través de la obtención de datos, la interpretación de los datos estructurados o no estructurados que recopilan, el razonamiento sobre el conocimiento o el procesamiento de la información derivados de esos datos, y decidiendo la acción o acciones óptimas que deben llevar a cabo para lograr el objetivo establecido».

A partir de dicha concepción, es indudable el gran potencial que tiene no solo por la infinidad de posibles aplicaciones, sino también en términos de reducción de tiempo y coste para la Administración de Justicia. Tal y como señala Barona Vilar, existe «un componente economicista evidente, menos esfuerzo, menos gasto, mayor efectivismo» (Barona Vilar, 2019a, págs. 41-42). Además, tampoco puede desconocerse

que, en determinadas ocasiones, podrán contribuir a objetivar determinadas situaciones y, en consecuencia, obtener soluciones «más justas» (ello sin obviar los posibles sesgos de los algoritmos a los que muy sucintamente después nos referiremos).

Excedería con creces el objeto del presente tratar de abordar todas las herramientas existentes, pero sí conviene, al menos, mencionar algunas de las más relevantes. Así, a título de ejemplo, entre las herramientas de IA aplicadas al ámbito policial y judicial podemos encontrar desde aquellas que sirven para calcular el riesgo de reincidencia (como el famoso sistema COMPAS), pasando por aquellas cuya finalidad es detectar denuncias falsas (como VERIPOL, el sistema empleado en las comisarías españolas), hasta las que permiten predecir delitos (como serían las conocidas HART o PREDPOL), o incluso predecir el fallo de una sentencia (como JURIMETRÍA).²

Por lo que respecta al ámbito de la justicia predictiva, estos sistemas se han definido como aquellos que utilizan un gran número de datos de distinta naturaleza que, procesados mediante algoritmos, ofrecen información sobre el comportamiento futuro de una persona en distintos escenarios (De Hoyos Sancho, 2021, pág. 3). Huelga decir que tienen un evidente potencial en el campo de la predicción y prevención delictiva, esto es, el denominado *predictive policing*. Dentro de estas últimas, podemos hallar, a su vez, dos modalidades: las basadas en la persona y las basadas en la localización, siendo las primeras las que mayores problemas plantean, en la medida en que el algoritmo se alimenta con unos datos que pueden contener sesgos por razón de la raza o el género (Miró Llinares, 2020, pág. 4). Aunque de las basadas en la localización se ha criticado que podían conllevar la reducción de la vigilancia en otras zonas «menos conflictivas» (Harcourt, 2007, pág. 132).

Cobran también especial relevancia las herramientas de cálculo de riesgos, en la medida en que, a partir del cálculo del riesgo de reincidencia, pueden auxiliar al juez en la toma de decisiones relativas a la imposición de una pena privativa de libertad, la adopción de una medida de libertad vigilada o la sujeción a programas de corrección más intensos (Monahan y Skeem, 2015, págs. 493-493). Dentro de los sistemas de *risk assessment* el más conocido es,

1. Comisión Europea, Dirección General de Redes de Comunicación, Contenido y Tecnologías (2019). *Directrices éticas para una IA fiable*. Oficina de Publicaciones.
2. Puede verse más en Chanenson S. y Hyatt J., 2016, pág. 4.

sin duda alguna, COMPAS (*Correctional Offender Management Profiles for Alternative Sanctions*), empleado en Estados Unidos, y dado a conocer especialmente a partir del caso *State vs. Loomis* al que luego nos referiremos.

Junto a estas no podemos dejar de mencionar VIOGEN, el Sistema de Seguimiento Integral en los casos de Violencia de Género, aun cuando la doctrina considera que en sentido estricto no es IA (Miró Llinares, 2018, pág. 103), pero su relación y similitud con estas herramientas es indudable. Se trata de una herramienta informática que, entre otras funciones, evalúa el nivel de riesgo para las víctimas y, en función de esta predicción, realiza el seguimiento y protección de las víctimas en todo el territorio nacional, además de emitir avisos cuando se detecte alguna incidencia que pueda poner en peligro a la víctima (Montesinos García, 2021).

Finalmente, el ámbito de la tutela cautelar presenta muchas oportunidades para la implantación de sistemas de cálculo de riesgo en los ámbitos tanto civil como penal, ya que a través de estos podría calcularse el riesgo de mora procesal que integra uno de los presupuestos básicos para la adopción de una medida cautelar. Como recoge Montesinos García, estas herramientas permitirían objetivar los riesgos que integran el *periculum in mora* en función de la clase de proceso y medida solicitada, para arrojar un resultado respecto de la probabilidad que se produzca la situación de riesgo que permitiría al juzgador acordar o rechazar la adopción de la medida cautelar (Montesinos García, 2022, pág. 13; Planchadell Gargallo, 2021, págs. 399-402).

1.2. Riesgos derivados de su utilización

Resulta sencillo imaginar los riesgos que se derivan de la aplicación de herramientas de IA en el proceso judicial y el impacto que esto puede tener en los principios y garantías fundamentales de naturaleza procesal cuando el órgano judicial adopte una decisión con base en el resultado de la máquina. Dejando de lado el derecho de defensa y la contradicción, a los que luego nos referiremos con mayor detenimiento, la afectación podrá alcanzar también a los principios esenciales del proceso. Así pues, como pone de relieve la doctrina, en un contexto de toma integral de decisiones por una máquina, carecería de sentido hablar

de intermediación y oralidad, entre otras muchas instituciones procesales (Gómez Colomer, 2022, pág. 263).

También podrá tener un impacto sobre el principio de imparcialidad en la actuación jurisdiccional, en la medida en que el juzgador pueda prejuzgar el asunto debido a la decisión inicial de la máquina o incluso que su decisión adolezca de los sesgos que puedan existir en el algoritmo, tal y como antes adelantábamos. Todo ello en tanto que los algoritmos reflejan los datos con los que se los alimenta, por lo que, si esos datos incorporan sesgos, el algoritmo los reproducirá o incluso los potenciará (Martínez Garay, 2018, pág. 496; Angwin *et al.*, 2016).

La realidad es que no podemos desviar la mirada ante los riesgos que plantea la utilización de estas herramientas, pero tampoco adoptar una visión pesimista y de rechazo absoluto hacia estas, sino que el objetivo debe ser minimizar tales riesgos para poder también disfrutar de las innumerables ventajas que presenta. Por ello, identificar y reconocer las insuficiencias de estas herramientas, sus sesgos y sus riesgos, es compatible con tratar de corregirlas y regularlas (Miró Llinares, 2022, pág. 178). Esta corrección debe hacerse desde el más absoluto respeto a los derechos fundamentales, con una visión realista, que pondere adecuadamente los valores en juego e incorpore a la tecnología los valores éticos (Miró Llinares, 2022, pág. 178).

La vía para lograrlo es una regulación adecuada, que identifique los riesgos y contemple medidas para paliarlos. Sin embargo, esto no es todavía una realidad, ni en el ordenamiento nacional ni tampoco en el internacional. No obstante, en el ámbito europeo sí que se han desarrollado una serie de textos no vinculantes que arrojan pistas sobre cómo debería regularse la utilización de herramientas algorítmicas en la justicia, además de poner de relieve cómo pueden estos sistemas impactar en los derechos fundamentales. A ello se suma que desde el año 2021 se encuentra en tramitación la Propuesta de Reglamento de la Unión Europea sobre inteligencia artificial,³ la cual, una vez entre en vigor, será de aplicación directa en los ordenamientos de los Estados miembros.

El primero de dichos textos es la Carta Ética Europea sobre el uso de la inteligencia artificial en los sistemas judiciales

3. Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial (ley de inteligencia artificial) y se modifican determinados actos legislativos de la Unión (COM/2021/206 final).

y su entorno, adoptada en el seno del Consejo de Europa en el año 2018.⁴ Este texto recoge los principios que deben regir en la utilización de IA en la justicia y que se concretan en los siguientes: respeto por los derechos fundamentales; no discriminación; calidad y seguridad; transparencia; imparcialidad y justicia; y, por último, control del usuario. Respecto del primero de los principios, tiene como finalidad que el diseño e implementación de herramientas y servicios de IA sean compatibles con los derechos humanos. En particular, el texto se refiere a la necesidad de asegurar que la utilización de herramientas que ayuden en la toma de decisiones judiciales no menoscabe las garantías del derecho de acceso a la justicia y a un juicio justo.

A esta carta le siguieron, en 2019, las Directrices éticas para una IA fiable, referidas *supra*. Aun cuando este documento se refiere a un ámbito más amplio que el anteriormente referido, sí que contiene una mención expresa referida a los sistemas de justicia, cuando señala que las herramientas de IA deben incluir el compromiso «de asegurar el respeto de las garantías procesales y la igualdad ante la ley». En el mismo sentido, se afirma que «El respeto de los derechos fundamentales, dentro de un marco de democracia y estado de Derecho, proporciona la base más prometedora para identificar los principios y valores éticos abstractos que se pueden poner en práctica en el contexto de la IA».

En tercer lugar, no podemos tampoco dejar de referirnos, aunque sea muy brevemente, al *Libro Blanco sobre Inteligencia Artificial* de la Unión Europea,⁵ adoptado en el año 2020 por la Comisión Europea. Si bien no contiene ninguna referencia expresa a la utilización de sistemas algorítmicos en el ámbito judicial, sí que se resalta que las personas físicas y jurídicas podrán enfrentarse a dificultades en el acceso a la justicia frente a decisiones adoptadas con ayuda de herramientas de IA. Ello, según señala, debido a las características especiales de la IA -como la opacidad, complejidad, imprevisibilidad y autonomía-, que pueden conllevar que las partes afectadas desconozcan cómo se ha adoptado la decisión y si se han respetado las normas pertinentes.

Finalmente, llegamos a la Propuesta de Reglamento de la Unión Europea sobre inteligencia artificial. Por lo que respecta al ámbito judicial y policial, uno de los aspectos más relevantes de esta propuesta es que, tras la enmienda aprobada por el Parlamento Europeo en junio de 2023,⁶ queda prohibida la comercialización o utilización de sistemas de IA utilizados para llevar a cabo evaluaciones de riesgo de personas físicas o grupos de personas físicas, para determinar el riesgo de que cometan delitos o reincidan en su comisión, o con la misma finalidad, lleven a cabo la elaboración del perfil de personas físicas o la evaluación de rasgos de personalidad y características, en particular, su ubicación o conductas delictivas pasadas.

La importancia de dicha prohibición es mayúscula, pues de aprobarse dicho texto podría estar poniéndose veto a algunas de las herramientas que hemos mencionado, para cuyo cálculo se emplean precisamente datos personales relativos a la ubicación, antecedentes delictivos e incluso, como el caso de COMPAS, historial criminal de los familiares (Roa Avella *et al.*, 2022, págs. 283-284).

De todo lo expuesto es innegable la necesidad de contar con una regulación adecuada para la utilización de herramientas de IA en el proceso. Ello se traduce en la necesidad de disponer de un marco normativo en el cual se recojan los presupuestos, las condiciones para su utilización, las garantías esenciales y las consecuencias de su uso por las autoridades (De Hoyos Sancho, 2022, pág. 44).

Para ello, el primer paso debe ser la identificación de los aspectos más problemáticos que acarrear la implementación de esta clase de sistemas, para después determinar de qué modo pueden afectar a los derechos y garantías del proceso. Finalmente, se deberá diseñar una regulación que permita conjugar los beneficios que estas tecnologías ofrecen, sin que se produzca una merma de los derechos fundamentales de las partes en el proceso. En todo caso, las garantías deben primar sobre la eficiencia, pues esta se presume de la propia naturaleza de estas herramientas

4. *Carta Ética Europea sobre el uso de la inteligencia artificial en los sistemas judiciales y su entorno*, adoptada el 3 de diciembre de 2018 por la Comisión europea para la eficiencia de la Justicia (CEPEJ), del Consejo de Europa.
5. *Libro Blanco sobre la inteligencia artificial: un enfoque europeo orientado a la excelencia y la confianza*, adoptado el 19 de febrero de 2020 (COM (2020) 65 final).
6. Enmiendas aprobadas por el Parlamento Europeo el 14 de junio de 2023 sobre la propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de Inteligencia Artificial) y se modifican determinados actos legislativos de la Unión. Disponible en: https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_ES.html.

y sus perspectivas de expansión y aplicación (Martín Diz, 2019, pág. 819)

2. Contradicción y defensa ante este nuevo escenario

Sentado lo anterior, dedicaremos la segunda parte del presente trabajo a examinar más detenidamente cómo podrá verse afectada la contradicción y la defensa en el proceso ante este escenario que se vislumbra.

Partimos de que el principio de contradicción se erige, en palabras de Montero Aroca, como la «fuerza motriz del proceso, su garantía suprema» (Montero Aroca, 1997, pág. 139), de manera que el propio proceso se construye y se regula a partir de su vigencia. Ello tiene dos consecuencias lógicas. Primera, que una futura regulación sobre la utilización de herramientas de IA en la Administración de Justicia debe necesariamente velar por el respeto de la contradicción de las partes como núcleo esencial, no solo del proceso judicial, sino incluso de la fase policial previa. Segunda, que los jueces y cualesquiera otros operadores jurídicos que intervengan en el proceso deberán adoptar especiales cautelas, siempre que se tomen decisiones con apoyo de sistemas algorítmicos, con la finalidad de que el derecho de defensa de las partes sea respetado. Ahora bien, sí que debemos matizar que esta obligación alcanza relevancia constitucional cuando sea el órgano judicial quien la incumpla, pues la indefensión constitucionalmente prohibida debe ser imputable a la actuación judicial (Martí Mingarro, 2020, pág. 36).

Aun cuando las propias limitaciones de extensión del presente nos impiden llevar a cabo un análisis en profundidad sobre el derecho de defensa y el principio de contradicción, sí que conviene, cuanto menos, aproximarnos sucintamente a su concepto.⁷

Así pues, cabe recordar que el principio de contradicción implica que ambas partes deben conocer todos los elementos de hecho o de derecho que puedan influir en la resolución del proceso, a fin de poder rebatirlos y argumentar en su contra (Montero Aroca, 1997, págs. 141-142). Por su parte, el derecho de defensa es el derecho de las partes a ser oídas en el proceso, lo que se traduce en la

posibilidad de argumentar, alegar y probar, pudiendo tanto acusación como acusado aportar todos los hechos que estimen convenientes y los medios de prueba que se consideren pertinentes (Montero Aroca, 1997, págs. 141-142). Además, dichas garantías deben ser iguales para ambas partes, de modo que gocen de los mismos derechos y deberes en el proceso y la contienda se produzca en condiciones de igualdad, sin la atribución de privilegios para ninguna de ellas (Gómez Colomer, 2021, pág. 248).

Tomando en consideración dicha conceptualización, el derecho de defensa sea posiblemente el derecho que más gravemente puede verse afectado por el uso de sistemas algorítmicos de muy distinta naturaleza, tanto por las partes como por el órgano judicial.

A continuación, repasaremos qué circunstancias particulares de estos sistemas podrían poner en riesgo la defensa y la contradicción, en particular, cuando se utilicen para la toma de decisiones judiciales que impacten sobre los derechos de las personas.

2.1. Opacidad y falibilidad de los algoritmos

La característica de los sistemas de IA que más se critica es la opacidad de los algoritmos, pues puede conllevar que las partes desconozcan cómo se ha obtenido el resultado y, como consecuencia, no puedan impugnarlo adecuadamente. Si partimos de la afirmación de que nadie puede defenderse de aquello que desconoce, será necesario que la parte afectada por la resolución adoptada con el apoyo del algoritmo sepa cómo defenderse de este.

Esta precisamente es una de las cuestiones que se puso de relieve en el famoso caso *State vs. Loomis*, con relación al sistema COMPAS, es decir, el secretismo del algoritmo y la imposibilidad de conocer cómo se había calculado el riesgo de reincidencia con resultado de «alto riesgo» que llevó a la imposición de una pena privativa de libertad. A pesar de ello, el Tribunal Supremo de Wisconsin, en su polémica sentencia, rechazó el recurso afirmando, por un lado, que se basaba en datos individualizados y objetivos, que además habían sido proporcionados por el propio condenado en un cuestionario, y, por otro lado, que la decisión no se

7. Para un análisis más extenso sobre el derecho de defensa, nos remitimos al artículo de esta misma autora: De Luis García, E. (2018). «El derecho de defensa en el proceso penal: significado y manifestaciones en la jurisprudencia constitucional». *Diario La Ley*. n.º 9215.

había fundado exclusivamente en el resultado arrojado por COMPAS (Martínez Garay, 2018, págs. 491-492).⁸

Sin embargo, consideramos que la parte afectada debe poder cuestionar el diseño, el funcionamiento y los resultados del algoritmo (Martínez Garay, 2018, pág. 499). Por ello, deviene fundamental establecer mecanismos adecuados para poder valorar e impugnar los datos que se incorporen en el sistema, además de poder también rebatir los resultados que se alcancen (Barona Vilar, 2019b, pág. 14).

Ahora bien, como en el caso de COMPAS, la realidad es que estos algoritmos suelen protegerse amparados en secretos de empresa, por ello compartimos plenamente la opinión de Martínez Garay cuando afirma que «no debe permitirse que se haga negocio a costa de las garantías procesales de los acusados» (Martínez Garay, 2018, pág. 499). La solución pasaría por que el algoritmo fuese público y, considerando las ventajas que para la Administración de Justicia tienen estas herramientas en términos de coste y tiempo, las administraciones públicas podrían ser las primeras interesadas en desarrollarlas (Reifarth Muñoz, 2022, pág. 209).

A la opacidad algorítmica y los problemas que entraña, debemos sumar los errores que puede cometer el propio sistema, por lo que los esfuerzos deben encaminarse a aprovechar su utilidad, pero siendo conscientes de su falibilidad (Barona Vilar, 2022a, pág. 44).

2.2. Insuficiencia de motivación

Tal y como se acaba de exponer, la opacidad algorítmica puede lesionar el derecho de defensa en la medida en que el desconocimiento del funcionamiento del algoritmo impida la adecuada impugnación de este. Si, añadido a ello, la resolución judicial se limita a señalar que la decisión se adopta en virtud del resultado de la máquina, nos encontraremos ante una motivación insuficiente. Cabe en ese punto recordar que el deber de motivación de las resoluciones judiciales requiere que se justifique la decisión mediante una argumentación que recoja los fundamentos de las operaciones efectuadas por el juzgador (Fernández Entralgo, 1989, pág. 57).

La IA tiene el potencial para operar como asistente de la decisión judicial en tanto herramienta de valoración

de prueba y, en consecuencia, como asistente en la motivación de la sentencia (Nieva Fenoll, 2022, pág. 427). Por ello, en caso de querer impugnar la resolución, no se rebatiría únicamente el razonamiento esgrimido por el juzgador, sino que también debería poder atacarse el funcionamiento del algoritmo (Nieva Fenoll, 2021, pág. 1028).

Si pensamos en el caso *State vs. Loomis*, uno de los argumentos que esgrimió el Tribunal Supremo de Wisconsin para denegar el recurso fue que la decisión no se había basado exclusivamente en el resultado arrojado por COMPAS, sino que había tenido en consideración otros elementos. Pero ¿y si no hubiese sido así? ¿Y si el órgano judicial exclusivamente se hubiese limitado a recoger que, en tanto que el resultado de COMPAS era que el riesgo de reincidencia era elevado, procedía la imposición de la pena privativa de libertad? En ese caso, la resolución adolecería de una falta de motivación incompatible con las exigencias de la tutela judicial efectiva.

Así pues, como pone de relieve De Hoyos Sancho, existe el riesgo de que el juez acomode sus decisiones al resultado de la herramienta algorítmica, sin valoraciones adicionales (De Hoyos Sancho, 2020, pág. 36). Dicha carencia no solamente vulnerará el derecho a la tutela judicial efectiva, en su vertiente de derecho a una resolución motivada de fondo, sino que también provocará una situación de indefensión para la parte afectada negativamente por la decisión. En este sentido, compartimos el argumento esgrimido por Barona Vilar, en el sentido de que la valoración de la prueba por medios algorítmicos no puede ser un mero resultado calculado numéricamente, sino «el resultado del pensar jurídicamente», con independencia de si se ha apoyado en el primero (Barona Vilar, 2022b, págs. 602-603). Es más, como señala Bueno de Mata, si dejásemos que la decisión judicial fuese adoptada con base solamente en el resultado de la máquina, existiría el riesgo de caer en un «estadio de pensamiento único judicial», con un estancamiento de la doctrina jurisprudencial (Bueno de Mata, 2020, pág. 28), lo cual sería incompatible con la naturaleza dinámica del derecho. Es decir, del mismo modo que la asistencia de un sistema algorítmico puede mejorar la toma de decisiones judiciales, el elemento humano es fundamental, pues el razonamiento jurídico y la adaptación y evolución de la jurisprudencia y del dere-

8. *State vs. Loomis*, 881 N.W.2d 749 (Wisconsin, 2016). Disponible en: <https://law.justia.com/cases/wisconsin/supreme-court/2016/2015ap000157-cr.html>.

cho a la sociedad y sus constantes cambios requieren del pensar humano.

Además, deviene esencial determinar el grado de vinculación del juzgador con el resultado que arroje el algoritmo y si podría bastar con señalar en la resolución que se decide a partir del resultado de la herramienta, sin mayor argumentación (Ariza Colmenarejo, 2022, pág. 45). Por nuestra parte consideramos imprescindible que la persona que vaya a dictar la resolución pueda siempre separarse del resultado de la máquina y, además, no pueda limitarse a señalar si se acoge o no al resultado del algoritmo, sino que deba, en todo caso, razonar el motivo para haber actuado de un modo u otro. En definitiva, consideramos que estos sistemas nunca pueden ser vinculantes, sino que deben tener únicamente una función auxiliar al servicio del juez y es que, en palabras de Marchena Gómez «sustituir la decisión jurisdiccional por una resolución mecanizada que rinde culto a una supuesta precisión matemática, quebrantaría de modo irreparable las garantías del investigado, de forma especial su derecho de defensa» (Marchena Gómez, 2022, pág. 56).

Ante esta situación, no solamente será necesaria una regulación adecuada en el sentido expuesto, sino también la formación a los jueces en la utilización de estos sistemas, no solo para saber utilizarla correctamente, sino también para entenderla y razonar de forma adecuada las decisiones que se adopten con su apoyo (Ariza Colmenarejo, 2022, pág. 46).

2.3. Sesgo de automatización

Ubicamos el llamado «sesgo de automatización» inmediatamente después de la insuficiencia de motivación por su estrecha relación. El artículo 144 de la Propuesta de Reglamento sobre IA lo define como la «tendencia a confiar automáticamente o en exceso en la información de salida generada por un sistema de IA de alto riesgo (...), en particular con aquellos sistemas que se utilizan para aportar información o recomendaciones con el fin de que personas físicas adopten una decisión». Por ello, si las decisiones judiciales se basan exclusivamente en el resultado de la máquina, sin cuestionar si este es válido o no, estaremos ante el sesgo de automatización.

Para ilustrar este riesgo, nos sirve de nuevo en el caso Loomis, pues esa hipotética falta de motivación a la que aludíamos en el apartado anterior podría haber venido precisamente por esa «confianza ciega» en la herramien-

ta COMPAS y su resultado. Por lo tanto, cuando sea el juzgador quien incurra en el mismo, no solo se producirán las consecuencias expuestas *supra*, sino que además habrá una dejación en los deberes inherentes a la función de juzgar, tales como la valoración adecuada de la prueba. Ello en la medida en que asumir sin más el resultado ofrecido por la máquina supone que no ha habido un examen previo de la prueba aportada por las partes, ni un proceso de valoración lógica de esta. Pero, además, no solamente el juez puede caer en esta «confianza ciega» en el algoritmo, sino que también puede ocurrir que sean las partes. En este último caso, podría incluso ocurrir que el Ministerio Fiscal, con base en el resultado del algoritmo, modifique sus peticiones de modo que supongan una modificación de la acusación incompatible con el derecho fundamental a la defensa.

Por todo ello, como ya hemos apuntado *supra*, es fundamental determinar el grado de vinculación de los operadores jurídicos con el resultado que arroje el sistema de IA y tratar de reducir el sesgo de automatización de modo que, tras el examen y valoración de la prueba, pueda el juzgador decidir razonadamente si se aparta o no del resultado del algoritmo.

Conclusiones

Aparece ante nosotros un futuro en el cual las herramientas de IA no solo impregnarán nuestra vida cotidiana, como ya lo hacen, sino que alcanzarán irremediamente al sistema de justicia. Como hemos visto, algunas ya están comenzando a ponerse en práctica y poco a poco van revelando sus fallas y riesgos, pero también sus virtudes.

Si bien en España las experiencias son todavía limitadas, su generalización se vislumbra cada vez más próxima, sin perjuicio de que la posible prohibición de determinados sistemas en el texto definitivo del Reglamento de IA condicione su desarrollo. Por ello, experiencias como la de COMPAS en Estados Unidos son de enorme utilidad. Por un lado, para ver el impacto real que sobre el derecho de defensa y el debido proceso puede tener la adopción de decisiones con base en algoritmos. Y, por otro lado, porque, en el caso concreto *State vs. Loomis*, las razones esgrimidas por el Tribunal Supremo de Wisconsin para dotar de legitimidad a la decisión tomada con apoyo del algoritmo pueden ser las que algún día leamos de nuestros propios juzgados y tribunales.

Como decíamos al inicio del presente, no pretendemos ofrecer una visión pesimista, sino realista. Debemos ser conscientes de los beneficios que aportan en términos de reducción de tiempo y coste para la Administración de Justicia. Incluso en ciertas situaciones pueden contribuir a objetivar ciertas decisiones. Sin embargo, tampoco podemos desviar la mirada ante los riesgos que entrañan desde el punto de vista de los derechos y garantías procesales.

En particular, por lo que respecta al derecho de defensa y la contradicción, existen varias cuestiones que deben depurarse en una futura regulación sobre la utilización de herramientas de IA en los ámbitos judicial y policial. Y es que nos encontramos con distintos aspectos que pueden mermar un principio que, como hemos afirmado, es el motor del proceso y el pilar a partir del cual se construye. En otras palabras, no cabe proceso sin contradicción, de manera que la vigencia del principio de contradicción y el respeto del derecho de defensa se convierte un límite infranqueable.

Muchos de los obstáculos a los que nos hemos referido se podrían paliar mediante algunas medidas tales como la puesta a disposición pública de las herramientas de

IA o una adecuada formación a los operadores jurídicos sobre estos sistemas. Pero lo cierto es que lo que realmente urge es una regulación adecuada que identifique los distintos problemas que pueden surgir en la aplicación de estas herramientas al sistema y ofrezca un marco legal en el cual actuar para reducir su impacto.

Como decíamos al inicio, si hace unos años nos preguntábamos sobre la validez de una declaración por videoconferencia, ahora hemos dado un paso más en el desarrollo tecnológico y, del mismo modo que se hizo en su momento, debemos lograr ese equilibrio entre la eficiencia que aporta y el respeto de los principios y garantías básicos del proceso. En definitiva, se presenta la difícil tarea, cuando no imposible, de configurar el proceso judicial, de manera que podamos beneficiarnos de todas las ventajas que la IA ofrece sin que ello conlleve una merma de los derechos fundamentales de las partes.

Reconocimientos

Trabajo elaborado en el marco del proyecto Claves para una justicia digital y algorítmica con perspectiva de género (PID2021-1231700B-I00) del Ministerio de Ciencia e Innovación.

Referencias bibliográficas

- ANGWIN, J.; LARSON, J.; MATTU, S.; KIRCHNER, L. (2016). «Machine Bias. There's software used across the country to predict future criminals. And it's biased against blacks». *ProPublica* [en línea]. Disponible en: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
- ARIZA COLMENAREJO, M.J. (2022). «Impugnación de las decisiones judiciales dictadas con auxilio de Inteligencia Artificial». En: CALAZA LÓPEZ, S. (dir.) y LLORENTE SÁNCHEZ-ARJONA, M. (dir.). *Inteligencia artificial legal y Administración de Justicia*, págs. 29-54. Cizur Menor: Aranzadi.
- BARONA VILAR, S. (2022a). «La seductora algoritmización de la justicia. Hacia una justicia poshumanista (Justicia+) ¿utópica o distópica?». *El Cronista del Estado Social y Democrático de Derecho*, n.º 100, págs. 36-47.
- BARONA VILAR, S. (2022b). *Algoritmización del derecho y de la justicia. De la Inteligencia Artificial a la Smart Justice*. València: Tirant lo Blanch.
- BARONA VILAR, S. (2019a). «Inteligencia artificial o la algoritmización de la vida y de la justicia: ¿solución o problema?». *Revista Boliviana de Derecho*, n.º 28, págs. 18-49.
- BARONA VILAR, S. (2019b). «Cuarta revolución industrial (4.0) o ciberindustria en el proceso penal: revolución digital, inteligencia artificial y el camino hacia la robotización de la justicia». *Revista Jurídica Digital UANDES*, vol. 3, n.º 1, págs. 1-17. DOI: <https://doi.org/10.24822/rjduandes.03011>
- BUENO DE MATA, F. (2020). «Macrodatos, inteligencia artificial y proceso: luces y sombras». *Revista General de Derecho Procesal*, n.º 51, págs. 1-31.
- CHANENSON S.; HYATT J. (2016). *The Use of Risk Assessment at Sentencing: Implications for Research and Policy*. Working Paper Series. Villanova University Charles Widger School of Law.
- DE HOYOS SANCHO, M. (2022). «Garantías procesales en el uso jurisdiccional de sistemas de inteligencia artificial de alto riesgo». En: MARTÍN PASTOR, J. (dir.) y JUAN SÁNCHEZ, J. (dir.). *El Derecho procesal: entre la Academia y el Foro*, págs. 23-52. Barcelona: Atelier.
- 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*, n.º 55.
- DE HOYOS SANCHO, M. (2020). «El libro blanco sobre inteligencia artificial de la Comisión Europea. Reflexiones desde las garantías esenciales del proceso penal como "sector de riesgo"». *REDE. Revista española de derecho europeo*, n.º 76, págs. 9-43. DOI: https://doi.org/10.37417/REDE/num76_2020_534
- DE LUIS GARCÍA, E. (2018). «El derecho de defensa en el proceso penal: significado y manifestaciones en la jurisprudencia constitucional». *Diario La Ley*, n.º 9215.
- FERNÁNDEZ ENTRALGO, J. (1989). «La motivación de las resoluciones judiciales en la doctrina del Tribunal Constitucional». *Poder Judicial*, n.º 6, págs. 57-94.
- GÓMEZ COLOMER, J.L. (2022). «Derechos fundamentales, proceso e inteligencia artificial: una reflexión». En: CALAZA LÓPEZ, S. (dir.) y LLORENTE SÁNCHEZ-ARJONA, M. (dir.). *Inteligencia artificial legal y Administración de Justicia*, págs. 257-287. Cizur Menor: Aranzadi.
- GÓMEZ COLOMER, J.L. (2021). «Cuestiones Generales». En: GÓMEZ COLOMER, J.L. (coord.) y BARONA VILAR, S. (coord.). *Introducción al Derecho Procesal. Derecho Procesal I*, págs. 239-254. València: Tirant lo Blanch.
- GUZMÁN FLUJA, V. C. (2021). «Proceso penal y justicia automatizada». *Revista General de Derecho Procesal*, n.º 53.

- HARCOURT, B. E. (2017). *Against prediction: profiling, policing, and punishing in an actuarial age*. University of Chicago Press.
- MARCHENA GÓMEZ, M. (2022). «Inteligencia artificial y jurisdicción penal». Discurso leído por el Excmo. Sr. Dr. D. Manuel Marchena Gómez en el acto de su recepción como Académico de número de la Real Academia de Doctores de España y contestación del Excmo. Sr. D. Jorge Rodríguez-Zapata Pérez.
- MARTÍ MINGARRO, L. (2010). *Crisis del derecho de defensa*. Madrid: Marcial Pons. DOI: <https://doi.org/10.2307/j.ctv10rrbh4>
- MARTÍN DIZ, F. (2021b). «Modelos de aplicación de Inteligencia Artificial en justicia: asistencial o predictiva versus decisoria». En: BARONA VILAR, S. (ed.). *Justicia algorítmica y neuroderecho: una mirada multidisciplinar*, págs. 65-85. València: Tirant lo Blanch.
- MARTÍN DIZ, F. (2019). «Inteligencia artificial y proceso: Garantías frente a eficiencia en el entorno de los derechos procesales fundamentales». En: Jiménez Conde, F. (dir.) y Bellido Penadés, R. (dir.). *Justicia: ¿garantías versus eficiencia?*, págs. 815-827. València: Tirant lo Blanch.
- MARTÍNEZ GARAY, L. (2018). «Peligrosidad, algoritmos y due process: el caso State vs. Loomis». *Revista de Derecho Penal y Criminología*, n.º 20, págs. 485-502. DOI: <https://doi.org/10.5944/rdpc.20.2018.26484>
- MIRÓ LLINARES, F. (2022). «Inteligencia artificial, delito y control penal: nuevas reflexiones y algunas predicciones sobre su impacto en el derecho y la justicia penal». *El Cronista del Estado Social y Democrático de Derecho*, n.º 100 (ejemplar dedicado a: Inteligencia artificial y derecho), págs. 174-183.
- MIRÓ LLINARES, F. (2020). «Predictive policing: utopia or dystopia? On attitudes towards the use of big data algorithms for law enforcement». *IDP: revista de Internet, derecho y política*, n.º 30. DOI: <https://doi.org/10.7238/idp.v0i30.3223>
- MIRÓ LLINARES, F. (2018). «Inteligencia artificial y justicia penal: más allá de los resultados lesivos causados por robots». *Revista de Derecho Penal y Criminología*, n.º 20.
- MONAHAN, J.; SKEEM, J.R. (2015, diciembre). «Risk Assessment in Criminal Sentencing». *Annual Review of Clinical Psychology*.
- MONTERO AROCA, J. (1997). *Principios del proceso penal: una explicación basada en la razón*. València: Tirant lo Blanch.
- MONTESINOS GARCÍA, A. (2022). «Empleo de la inteligencia artificial en algunas fases del proceso judicial civil: prueba, medidas cautelares y sentencia». *Actualidad civil*, n.º 9.
- MONTESINOS GARCÍA, A. (2021). «Los algoritmos que valoran el riesgo de reincidencia: En especial, el sistema Viogen». *Revista de derecho y proceso penal*, n.º 64, págs. 19-55.
- NIEVA FENOLL, J. (2022). «Inteligencia artificial y proceso judicial: perspectivas ante un alto tecnológico en el camino». En: CALAZA LÓPEZ, S. (dir.) y LLORENTE SÁNCHEZ-ARJONA, M. (dir.). *Inteligencia artificial legal y Administración de Justicia*, págs. 417-437. Cizur Menor: Aranzadi.
- NIEVA FENOLL, J. (2021). «Tecnología y derechos fundamentales en el proceso judicial». En: MORENO CATENA, V. (dir.) y ROMERO PRADAS, M.I. (dir.). *Nuevos postulados de la cooperación judicial en la Unión Europea: libro homenaje a la Prof.ª Isabel González Cano*, págs. 1007-1033. València: Tirant lo Blanch.
- PLANCHADELL GARGALLO, A. (2021). «Inteligencia Artificial y medidas cautelares». En: BARONA VILAR, S. (ed.). *Justicia algorítmica y neuroderecho: una mirada multidisciplinar*, págs. 389-419. València: Tirant lo Blanch.

REIFARTH MUÑOZ, W. (2022). «El uso de la inteligencia artificial en el proceso judicial y los derechos fundamentales». En: BUENO DE MATA, F. (dir.). *El impacto de las tecnologías disruptivas en el derecho procesal*, págs. 199-216. València: Tirant lo Blanch.

ROA AVELLA, M.; SANABRIA-MOYANO, J.; DINAS-HURTADO, K. (2022). «Uso del algoritmo COMPAS en el proceso penal y los riesgos a los derechos humanos». *Revista Brasileira de Direito Processual Penal*, vol. 8, n.º 1. DOI: <https://doi.org/10.22197/rbdpp.v8i1.615>

Cita recomendada

LUIS GARCÍA, Elena (2023). «Justicia, inteligencia artificial y derecho de defensa». En: Miró, F. (coord.). «Digitalización y algoritmización de la justicia». *IDP. Revista de Internet, Derecho y Política*, núm. 39. UOC [Fecha de consulta: dd/mm/aa] <http://dx.doi.org/10.7238/idp.v0i39.417164>



Los textos publicados en esta revista están –si no se indica lo contrario– bajo una licencia Reconocimiento-Sin obras derivadas 3.0 España de Creative Commons. Puede copiarlos, distribuirlos y comunicarlos públicamente siempre que cite su autor y la revista y la institución que los publica (*IDP. Revista de Internet, Derecho y Política*; UOC); no haga con ellos obras derivadas. La licencia completa se puede consultar en: <http://creativecommons.org/licenses/by-nd/3.0/es/deed.es>.

Sobre la autoría

Elena de Luis García

Profesora Permanente Laboral en el Departamento de Derecho Administrativo y Procesal de la Universitat de València, Área de Derecho Procesal
elena.deluis@uv.es

Actualmente es miembro del Proyecto de investigación Claves para una justicia digital y algorítmica con perspectiva de género (PID2021-1231700B-I00) del Ministerio de Ciencia e Innovación y miembro del grupo de investigación MEDARB de la Universitat de València desde el año 2016. Su línea de investigación más reciente se centra en la utilización de herramientas de justicia predictiva en el proceso penal y su incidencia sobre los derechos fundamentales y las garantías procesales.