

LA APLICACIÓN DE LA INTELIGENCIA ARTIFICIAL A LA JUSTICIA¹

THE APPLICATION OF THE ARTIFICIAL INTELLIGENCE TO JUSTICE

■ **DRA. CRISTINA HERMIDA DEL LLANO**

Catedrática de Filosofía del Derecho,
Universidad Rey Juan Carlos, España

<https://orcid.org/0000-0001-8213-3840>

cristina.hermida@urjc.es



Resumen

La inteligencia artificial ha llegado para revolucionar la realidad social y, en particular, el ámbito del Derecho, con un amplio campo de acción en el espacio jurisdiccional. Es cierto que la inteligencia artificial puede constituir una valiosa herramienta de apoyo para la generación de soluciones jurídicas en un marco de productividad y aprendizaje continuo en la optimización de tareas rutinarias y de orden cuantitativo, y que puede llegar a reducir la mora judicial y la duración de los procesos. Sin embargo, aquí se examinan los graves riesgos que entraña un mal uso de esa herramienta en la administración de justicia. Se apela a mantenerse vigilantes ante posibles abusos o extralimitaciones, que puedan poner en entredicho las exigencias o los valores morales

- 1 Este texto se inscribe en el proyecto de investigación, del que la autora es investigadora principal, titulado «La libertad de expresión del juez en el marco del respeto a los principios de ética judicial en España», Proyecto PID2021-127122NB-I00 financiado por MCIN/AEI/10.13039/501100011033/ y por FEDER Una manera de hacer Europa. El texto fue defendido como ponencia en la Mesa Redonda «Inteligencia Artificial: Perspectiva Judicial y Perspectiva Antropológica» del curso Encuentro Comisión de Ética Judicial (EN24039), dirigido por don Antonio del Moral García, celebrado en el Tribunal Superior de Justicia de Granada entre el 22 y el 24 de abril de 2024.

básicos de la sociedad hasta el punto de provocar un cambio perverso de paradigma antropológico.

Palabras clave: Inteligencia artificial; ética; justicia; técnica; jueces.

Abstract

Artificial intelligence has arrived to revolutionize social reality and, in particular, the field of law, finding a wide field of action in the jurisdictional field. It is true that artificial intelligence can be a valuable support tool for the generation of legal solutions within a framework of productivity and continuous learning in the optimization of routine and quantitative tasks and can even reduce judicial delays and the duration of proceedings. However, here we examine the serious risks involved in the misuse of this tool in the administration of justice. We are called upon to remain vigilant in the face of possible abuses or overreach that could jeopardize the basic moral demands or values of society to the point of provoking a perverse change in the anthropological paradigm.

Keywords: Artificial Intelligence; Ethics; Justice; Technique; Judges.

Sumario

I. Introducción; II. Un acercamiento a la inteligencia artificial desde la iusfilosofía; 2.1. El *tsunami* del poder económico; 2.2. Surgimiento de un nuevo paradigma antropológico: Erradicación de la dimensión utópica del ser humano y de un posible renacimiento moral; 2.3. El uso de una racionalidad instrumental teñida de sesgos; 2.4. Una visión utilitarista del progreso humano y de la libertad en detrimento de la felicidad; 2.5. Una amenaza frente al humanismo o los derechos humanos fundamentales; 2.6. Jugando a ser Dios, desacralizando la naturaleza humana; 2.7. Trivialización y cosificación del ser humano; 2.8. Desnaturalización del poder judicial; III. Conclusiones; IV. Referencias.

I. INTRODUCCIÓN

«Ningún cambio en la manera en que [se piensa o actúa] podrá ser más decisivo que este: [debe ponerse] a las personas en el centro de

todo lo que [se haga]. No hay vocación más noble ni responsabilidad mayor que la de facilitar los medios a hombres, mujeres y niños, para que puedan vivir mejor» (Annan, 2000, p. 4). Esta frase de Kofi Annan retrata a la perfección la importancia del tema que aquí se aborda cuando se pregunta si la inteligencia artificial (IA) debería respetar el lugar que corresponde a los seres humanos. Además, anima a cuestionarse si, verdaderamente, el uso de la IA, aplicada a la justicia, permitirá vivir mejor y salvaguardar los derechos del ordenamiento jurídico.

Conviene comenzar aclarando que aquí se entiende por IA «el resultado del desarrollo de sistemas informáticos computacionales que simulan o pueden realizar tareas que normalmente requieren inteligencia humana, como la percepción, el razonamiento y el aprendizaje», siguiendo la definición recogida por la Comisión Iberoamericana de Ética Judicial (CIEJ) en su «XXIV dictamen sobre el uso ético de la inteligencia artificial en la labor jurisdiccional» (2023), cuyo ponente fue Sánchez Rodríguez. La pregunta que surge inmediatamente es si una máquina realmente percibe, razona y aprende y, en caso afirmativo, cómo lo hace.

Parece indiscutible que la IA ha llegado para revolucionar el ámbito del Derecho, en particular, con un amplio campo de acción en los procesos judiciales. A ello ha contribuido el avance imparable del fenómeno de la globalización, junto con las exigencias de una sociedad en constante crecimiento y desarrollo que está demandando sin cesar procesos de mayor eficiencia y celeridad. Como se ha precisado, en países como Argentina se han logrado elaborar dictámenes 200 % más rápido, en Reino Unido se han resuelto disputas por divorcio en línea y en China desde 2014 se cuenta con el primer asistente de IA en los juzgados, llamado Xiao Fa, que brinda asesoramiento legal a los justiciables en el tribunal de Beijing (Machuca, 2025; Zhabina, 2023).

Por no hablar del *software* de IA chino para el proceso penal, conocido como «sistema 206», el cual cuenta, entre sus principales iones, con

la constitución de estándares probatorios, que elimine o suprima toda manifestación de arbitrariedad que pueda exteriorizarse, a través del razonamiento interno, que deba realizar el tribunal a la hora de valorar la prueba, conforme a las reglas de la lógica, las máximas de experiencia y los conocimientos científicos,

aceptados o contenidos en informes periciales que hubieren podido mostrarse en el acto de juicio (arts. 50-59 LPPRPC). (Gil, 2021, pp. 357-358)

También, un caso curioso es el de Estonia, lugar donde ya se está apostando por jueces robot, como se ha resaltado en *The Technolawgist* (2019).

Ante la aseveración de que «una verdadera administración de justicia debe ser digital o no cumplirá con sus objetivos» (Simón, 2022, s. p.), el pensamiento iusfilosófico no puede quedar rezagado, sino que más bien al contrario debería asumir una actitud proactiva y permanecer alerta ante la inesperada dinámica social, resultado de los extraordinarios y vertiginosos avances que se están produciendo en el ámbito científico y tecnológico (Demetrio, 2017, pp. 2-3). Téngase en cuenta que estos están comenzando a tener una incidencia directa en los procesos judiciales y se debe asegurar que la función judicial continúe siendo una actividad humana, sin caer en la tentación de pensar que la IA es siempre la mejor opción para resolver un problema o dilema jurídico.

Aunque en no pocas ocasiones el Derecho vaya a remolque de la realidad social, aquel no solo puede ser un verdadero trampolín para el cambio social en este campo, sino que tiene la obligación de erigirse en él, a pesar de los vertiginosos e imparables avances tecnológicos. Ello, además, estaría en concordancia con la «Declaración sobre ética y protección de datos en el sector de la inteligencia artificial», aprobada en 2018 por las autoridades de protección de datos de la Unión Europea, el «Libro blanco sobre la inteligencia artificial» de 2020 de la Comisión Europea, el «Reglamento de inteligencia artificial de la Unión Europea» de 2024, o los principios contenidos en la «Carta ética europea sobre el uso de la inteligencia artificial», aprobada en 2018 por la Comisión Europea para la Eficiencia de la Justicia (CEPEJ), en el marco del Consejo de Europa.

De alguna manera, tanto el legislador nacional como europeo han de tratar de que la adaptación normativa sea capaz de seguir el ritmo del actual crecimiento tecnológico, para que no se generen vacíos en relación con las acciones que se puedan realizar, «tanto desde una estrategia preventiva como de otra estrategia de control y sanción para los usos indebidos», conforme a lo establecido por la CIEJ (2023) en el n.º 39 del «XXIV dictamen sobre el uso ético de la inteligencia artificial en la labor jurisdiccional».

La iusfilosofía como disciplina es fundamental a la hora de llamar la atención sobre los peligros que acechan frente al auge de un reduccionismo jurídico hasta la fecha desconocido, el cual encierra la contracción de la administración de justicia a la decisión adoptada por la IA. Probablemente, este nuevo reduccionismo pueda llegar a solapar o superponerse a otros que han sido debatidos en el campo de la Filosofía del Derecho durante siglos.

Piénsese, en primer lugar, en el reduccionismo del iusnaturalismo ontológico que considera Derecho solo al que se corresponde con los principios de justicia inmortales, eternos e imperecederos, la llamada o conocida como «Ley Natural», hasta el punto de destruir la juridicidad del Derecho que se considera injusto. El lector estará de acuerdo en que cuesta imaginar al personaje de Antígona en la gran tragedia de Sófocles, en el siglo va. C, invocando la decisión de una máquina como parámetro de justicia universal (Parodi, 2004, pp. 109-119). Y, en segundo orden, en el extremo opuesto, emerge el reduccionismo positivista ideológico radical, que encarna el filósofo político inglés Thomas Hobbes (1588-1679), cuando entiende que el Derecho es solo aquel del que ordena y manda en la sociedad, esto es, del soberano que concentra todos los poderes del Estado (Bobbio, 2004). Pues bien, el reduccionismo al que se está ahora seriamente expuestos en el siglo xxi sería el de considerar como Derecho justo aquel que emana del cauce de la IA.

II. UN ACERCAMIENTO A LA IA DESDE LA IUSFILOSOFÍA

El reduccionismo que deriva de una fe ciega en el uso de la IA aplicada a la administración de justicia plantea serios problemas, que aquí se abordarán, aunque sea brevemente, desde la iusfilosofía (Hermida, 2025).

2.1. EL TSUNAMI DEL PODER ECONÓMICO

Un dato meramente fáctico que ha subrayado, con sumo acierto Nemitz, es que la concentración de poder tecnológico, económico y político se encuentra

en manos de los “cinco temibles” (Amazon, Apple, Facebook, Google, Microsoft), que ocupan el centro de la escena del

desarrollo y uso comercial de la inteligencia artificial. [...] Esta acumulación de poder en manos de unos pocos debe ser observada de modo abarcativo: incluye el poder del dinero, el poder sobre la infraestructura democrática y el discurso público, el poder sobre los individuos a través de la elaboración de perfiles precisos y la supremacía en el campo de la inteligencia artificial. (2021, p. 131)

Esta aseveración obliga a reflexionar sobre si detrás de los expertos técnicos, apoyados por fuertes poderes económicos no debería estar siempre la mano humana de las personas juzgadoras cuando se hace uso de la IA para no dejar que la administración de la justicia se encuentre absolutamente dependiente de grandes corporaciones privadas o gigantes empresariales (Ponce, 2024). Como se ha resaltado: «A diferencia de internet, la inteligencia artificial no es producto de unos pocos académicos e idealistas, sino de las poderosísimas empresas de internet que la controlan» (Nemitz, 2021, p. 135); de tal manera, que los grandes Dioses decisorios pasan a ser unos pocos muy ricos por el poder económico que ostentan a nivel mundial y se crea una brecha todavía más grande entre ricos y pobres. Conforme ha subrayado la CIEJ en el «XXIV dictamen sobre el uso ético de la inteligencia artificial en la labor jurisdiccional»:

En numerosos casos, las tecnologías más avanzadas son desarrolladas por entidades privadas o externas a los Poderes judiciales, que podrían tener otro tipo de intereses, posiblemente más económicos que sociales, en el uso de los datos. Por lo general, las instituciones públicas no cuentan con personal profesional y técnico para conocer la forma en que operan estos sistemas, lo que genera una gran dependencia de proveedores externos. (2023, p. 7)

Para entender este riesgo de una forma todavía más clara y cercana piénsese en los graves problemas económicos que acechan al buscador de Google frente al auge imparable que está teniendo ChatGPT, en paulatino ascenso en el número de usuarios digitales. Según también ha precisado, con lucidez, Aretxaga (2020):

Se trata, como puede verse, de cuestiones de hondo calado que, como señalan los especialistas más sensatos, no pueden dejarse tan solo en manos de los científicos ni de los intereses económicos —u otros menos confesables—, especialmente

los de las grandes corporaciones. La importancia de lo que está en juego es de tal magnitud que se hace urgente que todos —científicos, políticos y opinión pública en general— nos hagamos eco de los retos y oportunidades que tales cuestiones representan, para que sepamos encontrar las respuestas más adecuadas con la seriedad y reflexión, pero también urgencia, que la situación demanda, evitando las posturas maximalistas, tanto del optimismo tecnófilo superficial e injustificado como del alarmismo catastrofista tecnófobo. Los riesgos y las oportunidades existen y, por ello, también las responsabilidades. (p. 231)

2.2. SURGIMIENTO DE UN NUEVO PARADIGMA ANTROPOLÓGICO: ERRADICACIÓN DE LA DIMENSIÓN UTÓPICA DEL SER HUMANO Y DE UN POSIBLE RENACIMIENTO MORAL

Si la IA funciona combinando grandes cantidades de datos con procesamiento rápido e iterativo y algoritmos inteligentes, que permiten al *software* aprender automáticamente de patrones o características en los datos, queda claro que la IA genera un nuevo paradigma a nivel antropológico. Con palabras de Aretxaga (2020), siguiendo a Beorlegui:

Este cambio de perspectiva —acaso de paradigma antropológico— supone un reto teórico y filosófico para los especialistas en estas materias, a la vez que conlleva profundas repercusiones en el ámbito de la ética, así como en el modo de organizar nuestras vidas y sociedades. (p. 230)

De hecho, en 2021, en el marco de las Naciones Unidas, la Unesco adoptó la *Recomendación sobre la ética de la inteligencia artificial*, que propone valores y principios que buscan, entre otras cuestiones, orientar a los Estados en la formulación de leyes, políticas u otros instrumentos relativos a la IA y las acciones para asegurar la incorporación de la ética en todas las etapas del ciclo de vida de los sistemas de IA (Unesco, 2021). Entre los valores que enumera esta *Recomendación* se encuentran: 1) Respeto, protección y promoción de los derechos humanos, las libertades fundamentales y la dignidad humana; 2) Prosperidad del medio ambiente y los ecosistemas; 3) Diversidad e inclusión; 4) Convivencia en sociedades pacíficas, justas e interconectadas. La misma *Recomendación* enuncia

los principios en los que se inspira, a saber: 1) Proporcionalidad e inocuidad; 2) Seguridad y protección; 3) Equidad y no discriminación; 4) Sostenibilidad; 5) Derecho a la intimidad y protección de datos; 6) Supervisión y decisión humanas; 7) Transparencia y explicabilidad; 8) Responsabilidad y rendición de cuentas; 9) Sensibilización y educación; 10) Gobernanza y colaboración adaptativas y de múltiples partes interesadas.

En lo que se refiere a la dimensión ética del ser humano, hay que tener en cuenta que la IA desconoce la visión a futuro, la dimensión utópica y anhelante, tan importante como inherente al ser humano. Nadie duda de que la IA puede llegar a ser un excelente agente probabilístico o predictivo, al asumir una función automática predictiva basada en patrones, pero lo que también parece evidente es que esta herramienta, con los patrones que juega, no mira al futuro, al trabajar solo con datos de un pasado histórico más o menos reciente. Se podría decir que la IA trabaja en función de una memoria histórica automatizada, que se va alimentando paulatinamente, incapaz de cambiar el curso previsible de los acontecimientos a futuro.

Tal y como señala la CIEJ en el XXIV dictamen (2023):

En el ámbito judicial, la IA puede ser utilizada, por ejemplo, para la automatización y mejora de los procesos, como pueden ser la identificación y comparación de pruebas, verificación y compilación de datos, programación de audiencias. Todas estas tareas pretenden una reducción de las cargas de trabajo de las personas juzgadoras y una reducción en los tiempos de tramitación de los casos. (p. 3)

En realidad, este dictamen remite al anterior de 2020 —el «IX dictamen sobre el uso por el juez de las nuevas tecnologías»—, en el que la CIEJ tuvo ocasión de explorar el empleo de la IA en el ejercicio de la función de juzgar.

Pues bien, a sabiendas de que la IA puede constituir una valiosa herramienta para contar con información más amplia y a mayor velocidad, así como para generar soluciones en un marco de productividad y aprendizaje continuo —dadas las ventajas que representa el uso de la IA en la optimización de tareas rutinarias y de orden cuantitativo—, resulta constatable una grave y palpable deficiencia, y es que esta no tiene una visión de futuro más que la que le dan los patrones del

pasado del quehacer humano, esto es, no goza de una dimensión utópica, lo que es un argumento a favor para reservar a las personas juzgadoras la tarea de adopción de las decisiones judiciales.

Además, a ello se suma otro problema al que se enfrenta la IA, estrechamente vinculado con lo anterior, y es que los seres humanos no siempre reflejan en el exterior, es decir, exteriorizan, sus preferencias, prejuicios, todos ellos cambiantes a lo largo de la vida.

No se puede, así, pasar por alto que la IA mutila o ignora otra dimensión importante del ser humano, como sujeto que es capaz de empezar desde cero en cualquier momento, sin volver sobre sus propios pasos, algo que, por cierto, la filósofa Hannah Arendt (2010) subrayó cuando hablaba de la importancia de la «vita activa» en el sujeto moral. Precisamente, la teórica política llamó la atención de que el mal radical del holocausto nazi tenía que ver con el hecho de «hacer superfluos a los seres humanos como seres humanos» (p. 209) y, curiosamente, de esa premisa se podía terminar derivando la tesis de que un juez bueno debería ser entonces aquel que combatiera la idea de lo superfluo para sí mismo, pero también para la sociedad. La dominación total del individuo pasaría, a juicio de Arendt, por distintas fases: en primer lugar, se mataba a las personas jurídicas de los hombres; en segundo lugar, venía el asesinato de la persona moral; en tercer lugar, la destrucción sistemática de la individualidad. Si «volver superfluos a los seres humanos equivale a erradicar las mismas condiciones que hacen posible la humanidad, destruir la pluralidad humana, la espontaneidad, la natalidad y la individualidad» (Bernstein, 2006, p. 244). *Sensu contrario*, podría sostenerse que una buena administración de justicia sería aquella que desde la bondad incondicional reconstruye la pluralidad humana, vivificando la espontaneidad, la natalidad y la individualidad, lo que no se puede conseguir a través de una máquina por compleja que sea esta.

2.3. EL USO DE UNA RACIONALIDAD INSTRUMENTAL TEÑIDA DE SEGOS

Aunque nadie dudaría en reconocer que el uso de la IA aplicada a la justicia puede facilitar la resolución de problemas o conflictos jurídicos en un plazo de tiempo más breve y con menor desgaste

o esfuerzo,² y que puede llegar a reducir la mora judicial y la duración de los procesos, también es cierto es que la IA se sirve de una racionalidad que no es finalista, sino meramente instrumental (Rayón, 2025). Precisamente por ello, aunque pudiera parecer que las emociones, a primera vista, no están tan presentes en las máquinas como en los seres humanos, en realidad, los sesgos pueden estar mucho más marcados y fuertes en las primeras, llegado el caso, y poner en peligro y entredicho la imparcialidad del juzgador. Téngase en cuenta que la decisión que emane de la IA estará en directa relación con los datos previamente introducidos en ella, los cuales

- 2 Ejemplos de usos de inteligencia artificial en América Latina, según el Instituto Belisario Domínguez, del Senado de la República de México (Rodríguez, 2022), serían los siguientes:

Argentina (Prometea): Su objetivo es la automatización de tareas reiterativas y la aplicación de IA para generar dictámenes judiciales en cuatro clases de amparos (habitacional, licencias para conducir taxi, ejecuciones fiscales y empleo público), para lo cual se basa en casos análogos para cuya solución ya existen precedente.

Colombia (PretorIA): Su objetivo es apoyar a la Corte Constitucional en la selección de tutelas judiciales de los derechos fundamentales que sean de su competencia.

México: En la actualidad no se implementa o desarrolla ningún procedimiento con IA en el sistema de justicia, pero existe evidencia de que se han desarrollado al menos tres proyectos relacionados con la IA: Justiniano, Expertius y Sistema Experto de Sentencias.

Justiniano: Fue desarrollado por el Instituto de Investigaciones Jurídicas de la Universidad Nacional Autónoma de México (IIJ-UNAM) y financiado por el CONACYT. Este sistema en materia de derechos humanos está basado en un modelo constructivista y busca resolver problemas consistentes en determinar hechos violatorios, derechos humanos vulnerados y el área de competencia a la cual debe ser asignado cada caso.

Expertius: Fue desarrollado por el Departamento de Inteligencia Artificial aplicada al Derecho, también del IIJ-UNAM y financiado por el CONACYT. Se basa en un modelo constructivista y de redes neuronales y tiene por objeto apoyar en la toma de decisiones judiciales en los procesos relacionados con pensiones alimenticias.

El Sistema Experto de Sentencias (SES), desarrollado por la Dra. María del Socorro Téllez Silva en su tesis doctoral: *Una aproximación a la representación del conocimiento legal en computadora asistido por inferencias difusas*. Fue un prototipo para apoyar a los jueces de lo familiar a emitir sentencias en materia de juicios de divorcio necesario, patria potestad y alimentos, cuya base de conocimiento se integra por los requisitos de forma y fondo de una sentencia de derecho familiar.

dependen, necesariamente, de decisiones humanas cargadas de valores, ideologías, condicionamientos de carácter cultural, económico, social, etcétera.

Sobre ello ya llamó la atención el Dictamen de la CIEJ de 2020, cuando retomaba, específicamente y respecto de la administración de justicia, los principios contenidos en la «Carta ética europea sobre el uso de la inteligencia artificial», mencionada *supra*. Allí se resaltó la repercusión que había tenido en diversos tribunales el uso de algoritmos en los ámbitos penal y administrativo, en cuanto al tratamiento privado de los datos, su interpretación y los fines perseguidos, así como la posibilidad de que se exacerbaban los sesgos ya presentes en los datos. Como ha recordado Ponce (2024), no se pueden ignorar

los sesgos cognitivos de los programadores, especialmente los sesgos de confirmación y disponibilidad, los de las personas supervisoras del sistema, especialmente el de automatización, y el de la persona decisoria humana, si la decisión no es automatizada, especialmente los de anclaje y automatización, se transfieren al sistema algorítmico o afectan a su desarrollo en la realidad. (p. 61)

Recientemente, el caso BOSCO ha llegado al Tribunal Supremo español, en julio de 2025, para convertirse en el primer gran pulso que se libra en los tribunales sobre la transparencia algorítmica en España. De hecho, este litigio se ha convertido en una batalla clave para la transparencia de las decisiones automatizadas aplicadas a la gestión de lo público. Concretamente, el asunto nació a raíz de la denegación de una solicitud de una ayuda esencial para poder pagar la factura de la luz (bono social). Fue Civio, una organización sin ánimo de lucro y a favor de la transparencia, preocupada por acercar la información pública a la ciudadanía, la que decidió pedir el código del programa informático BOSCO que gestionaba desde las empresas eléctricas el bono social para los más vulnerables. Debido a que el ministerio denegó este acceso, Civio se decidió a interponer una reclamación ante el Consejo de Transparencia, ya que había indicios de que este programa público, que había sido diseñado para ayudar, estaba cometiendo errores que afectaban, directamente, a más de dos millones de familias vulnerables. La Audiencia Nacional apoyó el argumento que dio en su momento el Gobierno, y que señalaba que compartir este código suponía un daño a la propiedad intelectual, la

protección de los datos personales que usaba, la seguridad de prácticamente todos los sistemas informáticos de la administración y los límites a la transparencia de la ley. Se consideró que liberar el código podría provocar y facilitar que se atacasen los sistemas informáticos de la Administración con todo tipo de ciberataques como *phishing*, extorsiones o incluso el minado de criptomonedas (Terán, 2024).

La batalla legal duró seis años y fue el Tribunal Supremo el encargado de cerrar el caso. La decisión no solo afecta al bono social, sino que sienta un precedente clave para todos los sistemas de IA y algoritmos que ya están decidiendo o influyendo sobre los derechos y recursos públicos de los españoles.

Este caso es relevante porque pone de manifiesto los riesgos de la opacidad algorítmica y propone soluciones para una IA pública transparente y responsable, desde la labor realizada por Civio. Aunque es cierto que la búsqueda de más eficiencia por parte de la Administración pública es siempre una tentación, la opacidad en este ámbito supone un riesgo inasumible, máxime si se trata de la administración de justicia:

Si no sabemos cómo funcionan estas «cajas negras» algorítmicas, ¿cómo podemos detectar errores o sesgos? ¿Cómo sabemos si los criterios que nos aplican son justos o si nos discriminan? Lo que parece evidente es que cuando no hay transparencia, hay indefensión. (Civio, 2025, s.p.)

Está en juego el derecho de los ciudadanos a entender cómo se toman las decisiones que les afectan directamente. El Artículo 9.3 de la Constitución española garantiza la seguridad jurídica para que las personas puedan conocer las normas y los criterios que las afectan, y, si en su caso, son justas o no. Si los algoritmos que automatizan estas decisiones son opacos, las personas afectadas quedan en total indefensión y no pueden cuestionar posibles errores o sesgos.

Además, la falta de transparencia pone en riesgo el control judicial de las decisiones automatizadas. De hecho, Civio apeló al Artículo 106.1 de la Constitución, por el que los tribunales deberían poder fiscalizar la legalidad de las actuaciones administrativas, al interpretar que, sin acceso al código fuente, los jueces no podían comprobar si BOSCO estaba funcionando conforme a la ley, lo cual dejaba en manos de la Administración un poder discrecional sin control externo. En el recurso planteado, se remitió al Derecho com-

parado, al alegar que el Consejo de Estado de Italia, en 2019, había dictaminado que el código fuente de los algoritmos públicos debía ser accesible para su revisión judicial.

Como, con razón, ha apuntado Nemitz:

a las personas debería asistirles el derecho legal de estar informadas sobre la funcionalidad exacta de la inteligencia artificial que están utilizando. ¿Qué lógicas sigue? ¿Cómo afecta su uso los propios intereses o los intereses de terceros? Esta obligación de proporcionar información también debería existir, aunque no se estén procesando datos personales. En este caso, ya existe la obligación legal en el [«Reglamento general de protección de datos»]. En este contexto, debe rechazarse la exigencia de los gigantes tecnológicos de mantener en secreto los modos de funcionamiento y los procesos de toma de decisiones de la inteligencia artificial. (2021, p. 294)

En este mismo sentido se ha pronunciado la CIEJ en el XXIV dictamen (2023), supracitado:

Como parte del deber de transparencia de la Administración pública, los desarrollos de la inteligencia artificial deben basarse en algoritmos de «cajas blancas», que eviten la opacidad, es decir deben guiarse por la explicabilidad y deben permitir la trazabilidad, lo que implica que a la persona usuaria se le informe y sepa cuándo interactúa con un sistema de IA. Por esa misma razón, para un servicio público de justicia en una sociedad democrática no se recomiendan los sistemas de «cajas negras», es decir, aquellos en los que se incluyen datos de los que no se conoce su procesamiento. (p. 11)

Todo deriva de la presunción de que los sesgos que pueda tener el juez, cuando resuelve un caso como sujeto humano, pueden haberse trasladado a la IA, en los mismos términos; sesgos como los de anclaje, reconocimiento o confirmación (Sancho, 2022), que pueden haber quedado en los algoritmos de la IA. En este sentido, la CIEJ (2023) ha precisado, con rotundidad:

Los programas de IA utilizan como materia prima la información estadística que recopilan de fuentes, hasta ahora, manejadas por seres humanos; esto implica que, si la información base presenta algún tipo de sesgo, la IA podrá reproducirlos e,

incluso, acrecentarlos. Por ejemplo, si en un país son frecuentes los arrestos de los ciudadanos de una población minoritaria, como sucede en muchas ocasiones con los migrantes, los programas podrían interpretar que las personas migrantes, en general, tienen mayor tendencia a delinquir, cuando puede ser que esto no coincida con la realidad. (p. 8)

Si la protección de los derechos fundamentales constituye una responsabilidad individual y colectiva, ¿cómo se puede considerar correcta la respuesta de la IA a la pregunta, por ejemplo, sobre si está justificado suprimir derechos de ciertos individuos en aras de preservar la identidad cultural del grupo? Se coincide con Husserl en que no se puede destruir, en ningún caso, la razón filosófica occidental (Benhabib, 2008, pp. 175-203).

Además, a ello se suma que el Derecho no es una ciencia exacta como las matemáticas y, en consecuencia, parece poco razonable que con la aritmética se pueda ir a prisión como ya está ocurriendo en Estados Unidos (Maybin, 2016). En una de sus enmiendas, el Parlamento Europeo estableció límites al uso de la IA en el ámbito judicial en 2023 y, a modo de principio, proclamó:

La utilización de herramientas de inteligencia artificial puede apoyar la toma de decisiones, pero no debe substituir el poder de toma de decisiones de los jueces o la independencia judicial, puesto que la toma de decisiones finales debe seguir siendo una actividad y una decisión de origen humano. (Parlamento Europeo, 2023)

Es más, no es una cuestión baladí que la IA pueda llegar a fabricar pruebas falsas. Como ha resaltado la CIEJ en el n.º 36 del «XXIV dictamen sobre el uso ético de la inteligencia artificial en la labor jurisdiccional» (2023):

Por ejemplo, ya se está utilizando para la creación de fotografías y videos donde aparecen personas que realmente no estuvieron presentes. También es posible la simulación de voces o la réplica de estilos y rasgos de escritura, lo que dificulta distinguir entre las pruebas artificiosas y las reales. (p. 8)

Todo ello obliga a pensar que la decisión final debería depender de un ser humano.

2.4. UNA VISIÓN UTILITARISTA DEL PROGRESO HUMANO Y DE LA LIBERTAD EN DETRIMENTO DE LA FELICIDAD

Otro problema que, igualmente, plantea la aplicación de la IA a la justicia es que se apoya en una visión utilitarista del progreso humano y de la libertad, al creer que los valores que de hecho han de imperar son la celeridad y la eficacia, y se olvida de un binomio imprescindible, que vincula el reino de la justicia con el de la felicidad, algo propio de los seres humanos. El uso poco ético o abusivo de la IA, aplicada a la administración de justicia, conllevaría a que las personas juzgadoras dejaran de cultivar la virtud de la justicia y todas las virtudes dianoéticas aristotélicas perderían su razón de ser, debido a que los jueces cesarían de reflexionar de forma crítica y deliberar sobre la decisión justa, al ser una máquina la que, parafraseando a Montesquieu, pasaría a convertirse en «la boca que pronuncia las palabras de la ley» (1964, p. 586). Del mismo modo, se olvidaría que la felicidad goza, en realidad, de una doble dimensión: como bienestar privado y como felicidad pública, la cual resulta inviable cuando los ciudadanos depositan las virtudes públicas en las máquinas.

Cuando en 2014, influido por la era posdarwinista (Soler, 2018), Bostrom sugirió que las máquinas inteligentes podían llegar a tener valores, con significado humano, desde los que poder motivar sus decisiones y acciones, ya estaba pensando en valores vinculados con la corriente utilitarista fundamentada en la maximización de la felicidad y la minimización del sufrimiento.

No cabe duda de que el transhumanismo acecha tras estos planteamientos, basados principalmente en la razón instrumental y el utilitarismo (Romañach, 2016, p. 11). Ello supone un distanciamiento de la defensa de una razón universal, que reconoce, en verdad, a los seres humanos como titulares de un «derecho a tener derechos», resultante de luchas políticas, movimientos sociales y procesos de aprendizaje entre —y a través de— clases, géneros, naciones, grupos étnicos y credos religiosos. Como se ha precisado, con acierto:

Este es el auténtico significado del universalismo: [...] experiencias de establecer una comunalidad a través de la diversidad, conflicto, división y lucha. El universalismo es una aspiración, un objetivo moral por el que pelear; no es un hecho, una descripción del modo en que el mundo es. (Benhabib, 2008, p. 191)

Con otras palabras, no se puede pretender crear un modelo de *deber ser* desde una realidad fáctica estática, como si fuera una fotografía fija, debido a que la realidad se construye de forma paulatina y, muchas veces, resulta imprevisible.

2.5. UNA AMENAZA FRENTE AL HUMANISMO O LOS DERECHOS HUMANOS FUNDAMENTALES

La aplicación de la IA a la justicia puede convertirse en una amenaza al humanismo o a los valores morales. Algunos escándalos demuestran que, en nombre de la ciencia, se ha explotado e instrumentalizado a personas especialmente vulnerables (Macip y Willmott, 2013, p. 197). Esta elocuente declaración, de alguna forma, advierte que argumentar a favor de la ciencia no garantiza en absoluto que la práctica que se lleve a cabo sea siempre respetuosa con los valores morales, esto es, con los derechos humanos. Es, por ello, imprescindible mantenerse vigilante ante posibles abusos o extralimitaciones de la IA que puedan poner en riesgo las exigencias o los valores morales básicos de la sociedad, los derechos humanos. La CIEJ (2023) ha recordado a este respecto:

En relación con el poder judicial, la *Recomendación* de la UNESCO señala la necesidad de prever mecanismos para vigilar el impacto social y económico de dichos sistemas y el deber de los Estados de reforzar la capacidad del Poder Judicial para adoptar decisiones relacionadas con IA, particularmente en lo relativo a la protección de los derechos humanos, el Estado de Derecho, la independencia judicial y el principio de supervisión humana, así como la fiabilidad de esos sistemas, su orientación al interés público y centrados en el ser humano. (p. 2)

De cierta manera, un mal uso de la IA aplicada a la justicia puede poner en entredicho los rasgos inherentes a los derechos humanos fundamentales, pues estos son inalienables, inviolables, irrenunciables, intransigibles, personalísimos, indisponibles, activa y pasivamente, en cuanto que no son alienables por el sujeto titular de los derechos, y tampoco pueden ser expropiados o limitados por otros sujetos, menos aún, una máquina. Como ha precisado Laporta, al referirse a la inalienabilidad de los derechos, en tanto rasgo esencial y característico de ellos «[...] nadie puede renunciar a tales derechos, ni siquiera en el uso o ejercicio de uno de ellos: el derecho de libertad»

(Laporta, 1993 y 1995, p. 84); o bien, como el autor advierte: «Los derechos humanos, en tanto en cuanto [sic] son «inalienables», se le adscriben al individuo al margen de su consentimiento, o contra él, y se le inmuniza moralmente incluso frente a su propia voluntad» (Laporta, 1987, p. 44).

Debería quedar claro que los derechos humanos fundamentales no pueden dejar de ser un límite no solo para los poderes constituidos, sino incluso para la propia autonomía de sus titulares (Ferrajoli, 2001, pp. 47-48). Desde esta perspectiva, cabe recordar el Artículo 3.2 de la «Carta de derechos fundamentales de la Unión Europea», cuando prohibía las prácticas eugenésicas, el tráfico de cuerpos humanos, o de partes de estos y la clonación reproductora de seres humanos, que llevó a establecer, en su momento, no tanto límites al poder del Estado como a la libertad de la sociedad civil (Rubio, 2002, pp. 9 y 11).

Asimismo, como ha puesto de relieve la CIEJ (2023) en su XXIV dictamen:

El uso de la IA afecta a [sic] la garantía de determinados derechos fundamentales de gran trascendencia en una sociedad democrática: la intimidad, la privacidad y la protección de los datos personales, lo que constituye un particular reto frente a la introducción de sistemas de IA, porque el procesamiento de la información de las causas judiciales podría llevarse a cabo mediante sistemas informáticos desarrollados, operados y/o pertenecientes a empresas privadas, tanto locales como multinacionales, sobre los cuales el Estado podría ejercer un bajo o un nulo control directo. Esto conduce a la necesidad de establecer pautas claras para el tratamiento de los datos y mecanismos de fiscalización. (p. 7)

Detrás de la aplicación de la IA a la justicia se puede estar escondiendo la exaltación de un individualismo radical pernicioso, el cual termina salpicando la propia comprensión de los derechos fundamentales, al perderse de vista que los derechos no pueden ser entendidos en abstracto individualmente. Hay que tener en cuenta que las decisiones humanas se toman siempre en un contexto personal, social, cultural, etc., a lo que de modo preciso trató de dar respuesta el proceso de especificación de los derechos fundamentales y que

costó tantas luchas a lo largo de la historia para la consolidación de aquellos. Sobre esto, la CIEJ (2023) también se ha pronunciado:

En suma y en el mundo actual la IA carece de capacidad de juicio moral, por lo que no puede comprender el contexto emocional o psicológico de una situación concreta o bien no puede interpretar a ciencia cierta el espíritu de las normas, lo que es esencial para la labor jurisdiccional, por ejemplo, en el manejo del margen de discrecionalidad inherente en la adopción de una decisión judicial. (p. 9)

Efectivamente, solo las personas juzgadoras son capaces de atender al espíritu de la norma y, en suma, responder por las decisiones adoptadas. Como con agudeza ha subrayado García (2022), en el Derecho queda reflejada la propia condición humana:

Un cuerpo material, las normas jurídicas, y otro esencial, que podemos llamar su alma o razón, y que es el que orienta los movimientos de quien aplica la norma y de la norma misma, de modo que podrán dirigirse a hacer efectiva la realización de la Justicia en el mundo o perseguir todo lo contrario. Un Derecho sin Ética es un cuerpo sin alma o sin razón: es decir, un monstruo. El verdadero Derecho habrá de unir bajo un mismo concepto norma positiva y norma moral: ley y honestidad. (p. 12)

Los jueces en su tarea de interpretar las leyes y aplicarlas se guían por una brújula moral, que es inherente a cada ser humano, esto es, una conciencia ética, que jamás podrá ser transferida a una máquina o computadora. Pensar lo contrario supone como, con acierto, ha apuntado Espinosa (2010), una posición materialista que niega cualquier residuo o dimensión espiritual del ser humano. Literalmente, el autor precisa: «En este sentido de rebasar cualquier compartimentación ontológica, los animales y las máquinas son considerados también sujetos y objetos morales» (pp. 583-615), lo que tiene graves consecuencias desde el punto de vista de la ética práctica. El filósofo Keekok Lee ha sostenido, con razón, que la tecnología ha servido para transformar «lo *natural* en *artefactual*» (1999).

Es más, tan importante como el contexto emocional, social o psicológico, es el contexto normativo, esto es, ninguna disposición jurídica se aplica aisladamente, sino como parte del sistema al que pertenece y el ordenamiento en su conjunto condiciona el tipo de

norma expresada por cualquier enunciado jurídico. En este sentido, Schauer (2004, p. 251) ha destacado que, cuando se está sujeto a un conjunto de reglas, lo que se sigue de ellas depende más del modo en que se relacionan que de lo que establecen particularmente. Ello invita a preguntarse cómo crea la IA ese tejido normativo y cómo contextualiza el sistema.

2.6. JUGANDO A SER DIOS, DESACRALIZANDO LA NATURALEZA HUMANA

A través del uso de la IA aplicada a la justicia se está jugando a ser Dios, pues se asume como premisa primera la tesis antinaturalista y se alega que se caería en la falacia naturalista, al considerar que, de un juicio fáctico —no se hizo anteriormente uso de la IA en la justicia— cabe derivar un juicio de valor de corrección moral.

La defensa del uso indiscriminado de la IA, aplicado a la justicia, no solamente se enfrentaría al humanismo cristiano renacentista o ilustrado, sino a cualquier forma de sacralización de la naturaleza humana, tal y como ha defendido, por ejemplo, More en la *Declaración transhumanista de los principios extropianos* (Llano, 2018, pp. 66-67). No les falta razón a los que opinan que desde esta desacralización de la naturaleza humana y sacralización de la IA «se juega a ser Dios» porque «la naturaleza no es sagrada, por lo que nada impide modificarla, mejorarla o aumentarla» (Llano, 2018, p. 66).

2.7. TRIVIALIZACIÓN Y COSIFICACIÓN DEL SER HUMANO

Decía el filósofo español Julián Marías que la vida constituye una lucha entre dos impulsos contrapuestos: «el afianzamiento y extensión de la zona personal, y la invasión y reducción de ésta por lo impersonal, en último extremo por la cosificación» (1993, p. 140). Conforme a ello, cabría decir que, con el uso de la IA aplicada a la justicia, se quiera o no, se alimenta una corriente de pensamiento con la que no se trata de humanizar las máquinas sino más bien, lo que es peor, de cosificar al ser humano, al dejar de ser un fin en sí mismo en el sentido propiamente kantiano.

Resulta sumamente peligroso y nocivo el hecho de que la técnica deje de ser un simple instrumento y pase a convertirse en un fin en sí mismo. Ortega tuvo la lucidez de afirmar que «la reforma

de la naturaleza o técnica, como todo cambio o mutación, es un movimiento con sus dos términos, *a quo* y *ad quem*» (2014, p. 95). Puesto que el término *a quo* es la naturaleza, según está ahí, para modificarla resulta imprescindible fijar el otro término, hacia el cual se la va a conformar. Pues bien, como precisó el filósofo español, «este término *ad quem* es el programa vital del hombre» (Ortega y Gasset, 2014, p. 95), algo irrenunciable. En definitiva, de todos y cada uno depende no olvidar el alegato orteguiano de que la técnica no puede dejar de ser «como un gigantesco aparato ortopédico» que los técnicos crean, a sabiendas de que «toda técnica tiene esta maravillosa y —como todo en el hombre— dramática tendencia y cualidad de ser una fabulosa y grande ortopedia» (Ortega y Gasset, 2014, p. 183). De ahí que la técnica sea lo opuesto de la adaptación del sujeto al medio: «la adaptación del medio al sujeto» (Ortega y Gasset, 2014, p. 67).

En esta misma dirección, no pueden ser más oportunas las palabras de Aranguren, a propósito de la filosofía orsiana, entendida como esfuerzo de integración y superación de lo que el positivismo había separado:

En la filosofía de Xenius, el ser humano es tan animal como ángel: la humanidad es el resultado de la tensión entre ambos. Por debajo, el hombre es animal; por encima, ángel. Pero no es ninguna de las dos cosas en estado puro, sino ambas a la vez. (Hermida, 2020, p. 274)

De ahí que, siguiendo al filósofo mencionado, se deba no solo huir, sino combatir cualquier intento de cosificación del ser humano.

2.8. DESNATURALIZACIÓN DEL PODER JUDICIAL

Un mal uso de la IA aplicada a la administración de justicia puede poner en riesgo los pilares del Estado de Derecho, concretamente, desde un poder intermedio, el económico-tecnológico, que, como diría Montesquieu, termina desequilibrando a los tres poderes, al desnaturalizarse el poder judicial. En esta línea, la CIEJ (2023) ha afirmado:

Si la IA sustituyera plenamente las decisiones humanas en el análisis jurídico y probatorio de casos concretos se desnaturalizaría el sistema judicial de inspiración democrática, basado

en la confianza de la sociedad en sus personas juzgadoras, que son llamadas a resolver litigios de manera independiente. El criterio humano es imprescindible ante discusiones jurídicas novedosas que presenten zonas grises, frente a las cuales la diversidad de opiniones (según los distintos contextos, formación, ideología jurídica y métodos de razonamiento) permite mantener la función judicial como un instrumento vivo, que evoluciona y se adapta a las necesidades de la sociedad. (p. 9)

Efectivamente, la IA puede poner en entredicho la independencia e imparcialidad de los jueces. Como ha señalado la CIEJ (2023) en otro pasaje del «XXIV dictamen sobre el uso ético de la inteligencia artificial en la labor jurisdiccional», son

de alto riesgo ciertos sistemas de IA destinados a la administración de justicia y los procesos democráticos, dado que pueden tener efectos potencialmente importantes para la democracia, el Estado de Derecho, las libertades individuales y el derecho a la tutela judicial efectiva y a un juez imparcial. (p. 2)

Es más, para que no disminuya la confianza de la ciudadanía en la justicia, la población debería tener suficiente formación y educación tecnológico-digital para comprender el desarrollo de un proceso judicial que se hace depender de la IA y, además, confiar en que los jueces también la tienen, al convertirse en supervisores de dicho proceso en aras de evitar cualquier funcionamiento indebido o defectuoso de la administración de justicia. Como ha precisado la CIEJ (2023): «[...] las personas justiciables tienen que comprender la forma en que se maneja esa información para poder confiar en la veracidad de los datos que por medio de ella se obtengan» (p. 10).

Ello indica la existencia de otro problema más: la falta de formación en el personal judicial; de ahí que se haya considerado imprescindible una formación judicial específica para evitar, por ejemplo, sesgos (*debiasing*); incluso, ello lleva a pensar en la necesidad de una formación en conocimientos transdisciplinarios que van más allá del Derecho (Ponce, 2024, p. 242). En esta línea, la CIEJ (2023) ha señalado:

Se recomienda promover una cultura, una capacitación y una sensibilización del personal judicial sobre los conceptos, [el] funcionamiento y apoyo de los ambientes y sistemas de IA en las labores cotidianas, enfatizando en el uso de estas herramientas para tareas repetitivas y de menor complejidad, para

así dedicar al personal a las tareas más complejas. Para ello, es clave insistir, como han hecho las *Normas Éticas Sobre Inteligencia Artificial* de la OCDE (2019) y de la UNESCO (2021), en la necesidad de colocar al ser humano como el centro y fin de la IA, evitando así los mitos y los temores sobre los desplazamientos y despidos de personal por la utilización de estas nuevas herramientas. Estos mitos y creencias impiden el desarrollo cuando, en realidad, es el personal judicial quien debe encargarse de entrenar y alimentar con datos estos sistemas y ambientes de IA. (p. 12)

Téngase en cuenta, también, que la IA no puede asumir las tareas propias de un juez cuando pondera bienes jurídicos en conflicto. Sobre ello la CIEJ (2023) se ha pronunciado expresamente:

En el momento tecnológico actual y para la utilización de sistemas de IA en el ámbito jurisdiccional se recomienda no encomendar a la IA razonamientos complejos ni el ejercicio de técnicas de ponderación en las decisiones judiciales sino tareas repetitivas que generen patrones de un mismo comportamiento y que así generen predicciones tanto en el funcionamiento como en la clasificación de gestiones, el cumplimiento de requisitos, la cumplimentación de formularios y otras tareas de similar naturaleza. (p. 11)

III. CONCLUSIONES

Son muchas las cuestiones que surgen al acercarse al tema de la IA aplicada a la administración de justicia: ¿Se lleva la ciencia más lejos de lo razonable solo por el placer de poder decir que se es capaz de hacerlo? ¿La IA exige una precisa y clara regulación normativa para evitar un cambio de paradigma antropológico que pervierta los derechos fundamentales? A ambas cuestiones debería contestarse con rotundidad que sí. Como ha resaltado Capella, «urge que la razón jurídica ilustrada actúe en el ámbito de los avances tecnológicos» (2000, p. 69), y, aunque no se diga expresamente en esta cita, esa razón jurídica ilustrada puede ser cada quien, si se permanece activos y en alerta frente a posibles extralimitaciones en el uso de la IA aplicada a la justicia, así como frente a ámbitos sin regular por el Derecho, en este campo, que puedan poner en peligro la democracia y el Estado de Derecho. Hay que tener claro que la IA y las tecnologías de

la información y la comunicación pueden crear ámbitos de justicia más eficaces, pero no más incluyentes y equitativos; de ahí que se considere conveniente la defensa de los siguientes postulados:

1. Apuesta por una regulación legal clara y precisa de la IA aplicada a la administración de justicia.

No se puede mantener el alegato de que la ley es incapaz de seguir el ritmo trepidante del desarrollo de la tecnología e, ingenuamente, tampoco se ha de pensar que esta actúa de forma neutral en su avance imparable. Con palabras de Nemitz, «[...] nuestra democracia debe volver a exponerse al riesgo de una tecnología omnipresente y decisiva, que puede tener efectos negativos sustanciales si no se la regula» (2021, p. 135). Sin diferenciar entre sector público y privado, cabe afirmar que la Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión Europea sobre normas de Derecho civil sobre robótica (2015/2103/inl), representa uno de los planes mejor desarrollados hasta la fecha para manejar la IA. Tal y como señala la introducción:

Considerando que, ahora que la humanidad se encuentra a las puertas de una era en la que robots, bots, androides y otras formas de inteligencia artificial cada vez más sofisticadas parecen dispuestas a desencadenar una nueva revolución industrial —que probablemente afecte a todos los estratos de la sociedad—, resulta de vital importancia que el legislador pondere las consecuencias jurídicas y éticas, sin obstaculizar con ello la innovación.

Lo que garantiza que la IA sirva al bien público es su regulación normativa, no su desregulación (Nemitz, 2021, p. 134), además de una profunda regulación ética. Por ello hay que valorar los pasos dados con el «Reglamento de inteligencia artificial de la Unión Europea» de 2024 (Parlamento Europeo y Consejo de Europa, 2024).

2. Aumento de la confianza en los códigos éticos que regulan el uso de la IA.

No se debería menospreciar el papel que juegan los códigos éticos o *soft law* a la hora de dar regulación a la IA; se piensa, erróneamente, que todo el peso ha de recaer en la regulación legal, como ha defendido, entre otros autores, Nemitz (2021, pp. 133-134). Más bien, por el contrario; aun cuando es cierto que la ética no pue-

de remplazar la ley como marco delimitador del comportamiento, por carecer aquella de la legitimidad democrática y la naturaleza vinculante y coercitiva con que los gobiernos y tribunales pueden exigir el cumplimiento normativo (Nemitz, 2021, p. 137), a diferencia del pensador alemán, ese considera imprescindible volver al punto primigenio que está en la base de la ética individual. Ello quiere decir que hay que conseguir que los operadores jurídicos tengan una sólida formación ética, al igual que la sociedad civil, para acercarse y aprehender la realidad social de una manera más certera. En este contexto, podría recordarse que en 2022 la Organización para la Cooperación y el Desarrollo Económicos (OCDE) elaboró un estudio sobre el «Uso estratégico y responsable de la inteligencia artificial en el sector público de América Latina y el Caribe» en el que recomendaba, precisamente, de conjunto con la aplicación de los Principios de la OCDE sobre IA, el desarrollo de marcos éticos a nivel nacional (OCDE/CAF, 2022). No por casualidad los principios propuestos fueron, entre otros, valores y equidad centrados en el ser humano (OCDE, 2019).

Es más, cuando la ética juegue un papel verdaderamente importante en el ámbito público, entonces, como ha resaltado el filósofo alemán Marcus Gabriel, se podrá luchar por la defensa de una nueva Ilustración, gracias a la cual todo el mundo tendrá derecho a recibir una educación ética capaz de afrontar los peligros que supone seguir a ciegas la ciencia y la técnica (Gabriel, 2020), resultado, a mi modo de ver, de la insatisfacción permanente del ser humano. Ortega y Gasset (2014) ya se ocupó de recordar este aspecto al afirmar que

el hombre es, esencialmente, un insatisfecho, y esto —la *insatisfacción*— es lo más alto que el hombre posee, precisamente porque se trata de una insatisfacción, porque desea tener cosas que no ha tenido nunca. [...] Esta insatisfacción es como un amor sin amada o como un dolor que siente en unos miembros que nunca ha tenido. (p. 182)

Ello demuestra que es, a todas luces, deseable y urgente que esta nueva Ilustración no pierda de vista que el verdadero humanismo es aquel que preserva la autonomía personal y la dignidad humana (Parens, 1998, pp. 1-28 y 2009, pp. 181-197).

3. Mayor transparencia en cuanto al proceso de toma de decisiones de la IA aplicada a la administración de justicia.

Teniendo en cuenta que la IA no es infalible, como ha hecho notar la CIEJ en el n.º 34 del «XXIV dictamen sobre el uso ético de la inteligencia artificial en la labor jurisdiccional» (2023), los programas

pueden dar información errónea que sería asumida como verdadera sin mayor cuestionamiento por los usuarios finales que pueden acostumbrarse a confiar ciegamente en los resultados que presenten estos sistemas, o bien, sin posibilidad de contrastarlos con otras fuentes que permitan verificar la información. (p. 8)

Conviene apostar por la transparencia en el proceso de toma de decisiones, precisamente, para que se puedan depurar responsabilidades en caso de que un sistema de IA realice un análisis incorrecto o tome una decisión equivocada que oriente en un sentido determinado la resolución judicial porque ¿quién es el responsable de un posible error?, ¿las personas que lo diseñaron, ¿las que le introdujeron los datos?, ¿aquel que se basó en sus resultados para decidir? o, por el contrario, ¿se debe a un error del sistema?

4. La justicia no puede quedar reducida a fuerza convencionalmente automatizada ni el sujeto moral cosificarse.

Si la justicia no puede quedar reducida a fuerza convencionalmente organizada por la IA, ello es porque resulta imposible encerrar lo que denominó Verdross «la conciencia del derecho positivo», en una máquina. Entre otras razones, porque esa conciencia se nutre de principios fundamentales constantes que solo pueden deducirse a partir de la luz natural de la razón humana (Hermida, 2013, p. 37). Asimismo, la exigencia del respeto a principios como, entre otros, el de proporcionalidad para las limitaciones a los derechos fundamentales (Hermida, 2013, p. 41) impide recurrir de forma automatizada a la IA.

5. Evitar la pérdida de la circularidad hermenéutica de la razón práctica.

La circularidad hermenéutica de la razón práctica, sobre la que había llamado la atención Aristóteles en su *Ética*, constituye un rasgo esencial de todo razonamiento moral y político: se debe asumir, siempre, cierta comprensión de la igualdad, la reciprocidad y la simetría para

ser capaces de acotar el modelo discursivo, en palabras de Benhabib (2008). Efectivamente, la libertad comunicativa es lo que hace posible la práctica de justificación normativa, ya que,

si los seres humanos no pueden asentir a, o rechazar las exigencias de cada uno en base a razones cuya validez puedan evaluar, entonces no puede haber tarea justificatoria en absoluto. Incluso si las razones que invocamos en esta práctica son utilitaristas o kantianas, nietzscheanas o cristianas, al hacerlo debemos siempre previamente presuponer la capacidad de nuestro interlocutor de asentir o de disentir de nuestras exigencias sobre la base de razones cuya validez ella comprende. (p. 193)

Con la IA esa libertad comunicativa que es capaz de crear un discurso justificatorio simétrico desaparece:

El «universalismo justificatorio» se encuentra en el corazón de la razón en tanto que tarea de dar-razón del mismo modo que lo es el reconocimiento del otro en tanto que un ser capaz de libertad comunicativa y del derecho a tener derechos. (Benhabib, 2008, pp. 192-193)

6. En el Derecho no solo hay reglas, sino también principios, lo que dificulta la aplicación mecánica de la IA a la administración de justicia.

Este postulado obliga a darle la razón a Marchena, cuando afirma que «una robotización de la justicia será un modelo que “tiende a fosilizar” la justicia» (2022, s. p.) y, además, ignoraría, según Aguiló (2007, pp. 666-668) y Atienza (2017), que «los nuevos *fenómenos* jurídicos *exigen* la formulación de un nuevo paradigma teórico» que acepte que «en el Derecho *hay* reglas y, en consecuencia, *hay* razonamientos subsuntivos. Pero también *hay* principios; y guiarse por o aplicar principios *exige* un tipo de razonamiento, la ponderación, que es distinto del razonamiento subsuntivo» (Aguiló, 2007, p. 671 y 2023, pp. 15-39). Como ha repetido una y otra vez Alexy (1986, p. 87), «toda norma es un principio o una regla». Mientras que las reglas son mandatos definitivos que se aplican a la manera de todo o nada, por el contrario, los principios son mandatos de optimización que pueden ser cumplidos en mayor o menor grado en función de los condicionantes fácticos y jurídicos; de ahí que pueda afirmarse que mientras las reglas se aplican por subsunción, los principios

se aplican por ponderación (Ruiz, 2005, pp. 343-344). Efectivamente, principios y reglas podrían distinguirse por el fundamento de su validez. Siguiendo a Dworkin (1977, pp. 65-66), cabe decir que las reglas son normas que se pueden identificar por su pedigrí o fuente social, mientras que los principios son válidos «por ser una exigencia de la justicia» (p. 72), es decir, porque permiten una solución con un mérito moral que no es posible alcanzar a partir de la regla.

Si, siguiendo a Aguiló, lo que determina, verdaderamente, que una norma sea una regla o un principio es la concepción del Derecho que cada jurista asuma y lo que hace que un ordenamiento jurídico esté más o menos principalizado es la concepción del Derecho dominante entre los juristas que elaboran, aplican o estudian sus normas, incluyendo los valores que justifican su preferencia por esa concepción, de ello se deriva, por tanto, que no hay normas jurídicas que sean reglas o principios como tales, sino juristas y jueces principalistas o legalistas que, en función de su concepción del Derecho y las razones por las que la promueven, las reconstruyen de un modo u otro.

El origen del debate entre legalistas/reglalización del Derecho y principalistas/principalización del Derecho ha hecho pensar que se está ante concepciones interpretativas rivales del Derecho, es decir, manifestaciones de lo que Stavropoulos (2016) calificó como interpretivismo jurídico (*legal interpretivism*). Esto se puede llegar a complicar todavía más si se piensa que jueces afines al principalismo pueden estar vinculados con esta concepción interpretativa del Derecho de una forma más fuerte o débil, en función de si deciden en un sentido distinto a una regla, conforme a principios que guardan relación con el Derecho (débil) o principios extrasistemáticos (fuerte). El principalismo fuerte abre la puerta a una intensa moralización del ordenamiento jurídico, entrando en colisión con los que defienden la autonomía del Derecho respecto de la moralidad.

Ello obliga a reflexionar sobre si la IA aplicada a la administración de justicia estará más cerca o más lejos del legalismo o del principalismo, bien sea este fuerte o débil. En todo caso, una IA principalista débil no debería considerarse siempre incompatible con una intensa moralización del Derecho. En esta línea, conforme ha explicado Atienza (2017, pp. 132, 141), el juez postpositivista podría asociarse al principalismo débil, por ser activo en la defensa de los valores, principios y derechos de la Constitución, y encontrar, generalmente,

en esta la solución moralmente óptima para los casos que ha de resolver.

¿Quién es entonces el que decide si la IA es garante de los derechos, principios y valores constitucionales y que no existe ninguna regla que tenga el peso necesario para excepcionar ese propósito?

Como vemos son muchas las preguntas de calado que asaltan y faltan las respuestas hoy en día a todos estos interrogantes. El tiempo dirá si el modelo de Huxley se ha hecho realidad cuando advertía que, con el tiempo, se podría llegar a imponer una forma de tiranía

mediante la sugestión y la seducción, haciendo que seamos inducidos a «amar nuestro sometimiento. [...] En palabras de Harari, *una vez que alguien [...] consiga la habilidad tecnológica para manipular el corazón humano —de forma fiable, barata y a escala—, la política democrática se convertirá en un espectáculo de guiñol emocional.* (Vallespín, 2018, s. p.)

IV. REFERENCIAS

- Aguiló Regla, J. (2007). Positivismo y postpositivismo: dos paradigmas jurídicos en pocas palabras. *Doxa. Cuadernos de Filosofía del Derecho*, (30), 665-675. <https://doi.org/10.14198/DOXA2007.30.55>
- Aguiló Regla, J. (2023). Son mandatos de ponderación. Breviario de teoría del Derecho en honor de Manuel Atienza. *Doxa. Cuadernos de Filosofía del Derecho*, (46), 15-39 (Ejemplar dedicado a: 30 años de argumentación jurídica en Alicante). <https://doi.org/10.14198/DOXA2023.46.01>
- Alexy, Robert (1986). *Teoría de los derechos fundamentales*. Centro de Estudios Constitucionales.
- Annan, K. (2000). *Nosotros los pueblos. La Función de las Naciones Unidas para el Siglo XXI*. Fundación Kofi Annan. <https://docs.un.org/es/a/54/2000>
- Arendt, H. (2010). *Lo que quiero es comprender. Sobre mi vida y mi obra* (trad. Abella, M. y López Lizaga, J. L.). Trotta.
- Aretxaga Burgos, R. (2020). Recensión de la obra de Carlos Beorlegui, *Humanos. Entre lo prehumano y lo pos- o transhumano*.

Revista de Hispanismo Filosófico. Historia del Pensamiento Iberoamericano, (25), 230-232.

Atienza, M. (2017). *Filosofía del Derecho y transformación social*. Trotta.

Benhabib, S. (Julio-diciembre, 2008). Otro universalismo: Sobre la unidad y diversidad de los derechos humanos. *ISEGORÍA*, (39), 175-203. <https://doi.org/10.3989/isegoria.2008.i39.627>

Bernstein, R. J. (2006). ¿Cambió Hannah Arendt de opinión? Del mal radical a la banalidad del mal. En Giner, S., Jonas, H., Lefort, C., Wellmer, A. y otros. *Hannah Arendt. El orgullo de pensar*, 235-257. Gedisa.

Bobbio, N. (2004). *El problema del positivismo jurídico*. Fontamara.

Bostrom, N. (2014). *Superintelligence, Paths, Dangers, Strategies*. Oxford University Press.

Capella, J. R. (Agosto, 2000). Las transformaciones de la función del jurista en nuestro tiempo. *Revista de Crítica Jurídica*, (17), 51-69. <https://revistas-colaboracion.juridicas.unam.mx/index.php/critica-juridica/article/download/3174/2974>

Comisión Iberoamericana de Ética Judicial. (2023). XXIV dictamen de la sobre el uso ético de la inteligencia artificial en la labor jurisdiccional. <https://ciej.juslapampa.gob.ar/files/CIEJ/FICHEROS/Version%20final%20XXIV%20Dictamen%20Inteligencia%20Artificial%20ponente%20Sanchez%20Rodriguez.pdf>

Civio (Enero 23, 2025). La transparencia de los algoritmos públicos, en juego: Civio presenta el recurso sobre BOSCO ante el Tribunal Supremo. <https://civio.es/novedades/2025/01/30/la-transparencia-de-los-algoritmos-publicos-en-juego-civio-presenta-el-recurso-sobre-bosco-ante-el-tribunal-supremo/>

Demetrio Crespo, E. (2017). *Fragmentos sobre neurociencias y Derecho penal*. B de F - Euros editores.

Dworkin, R. (1977). *Los derechos en serio*. Gedisa.

Espinosa Rubio, L. (2010). El desafío del posthumanismo (en relación a las nuevas tecnologías). En Aullón de Haro, P. (Coord.). *Teoría del humanismo*, 3, 583-615.

Ferrajoli, L. (2001). *Derechos y garantías. La ley del más débil* (2.^a ed.). Trotta.

- Gabriel, M. (Marzo 25, 2020). El orden mundial previo al virus era letal. *El País*. https://elpais.com/cultura/2020/03/21/babelia/1584809233_534841.html
- García Paz, D. (2022). *Entre la ley y la honestidad*. Literatura Abierta (Grupo Torre de Lis).
- Gil García, F. S. (2021). Inteligencia Artificial y proceso penal en China: especial referencia al sistema 206. En González Pulido, I. (Coord.) y Bueno de Mata, F. (Dir.). *Estudios sobre tecnologías disruptivas y justicia*, 347-360. Lorenzo Mateo Bujosa Vadell.
- Hermida del Llano, C. (2013). La universalidad racional de los derechos. *Bajo Palabra*, II(8), 33-45. <https://doi.org/10.15366/bp2013.8.002>
- Hermida del Llano, C. (2020). Recensión de la obra de Navarra Ordoño, Andreu, *La escritura y el poder. Vida y ambiciones de Eugenio D'Ors*, Tusquets, 2018, 543 pp. *Revista de Hispanismo Filosófico*, (25), 272-276. <https://ahf-filosofia.es/wp-content/uploads/Rev-n25.pdf>
- Hermida del Llano, C. (2025). Inteligencia artificial: Perspectiva judicial en clave antropológica (col. Cuadernos Digitales de Formación). Centro Nacional de Documentación Judicial. Consejo General del Poder Judicial de España.
- Lee, K. (1999). *The natural and the artefactual. The implications of deep science and deep technology for environmental philosophy*. Lexington Books.
- Laporta, F. (1993 y 1995). *Entre el Derecho y la moral*. Fontamara.
- Laporta, F. (1987). Sobre el concepto de derechos humanos. *Doxa. Cuadernos de Filosofía del Derecho*, (4), 23-46. <https://doi.org/10.14198/DOXA1987.4.01>
- Llano Alonso, F. H. (2018). *Homo Excelsior. Los límites ético-jurídicos del transhumanismo*. Tirant Lo Blanch.
- Macip, S. y Willmott, Ch. (2013). *Jugar a ser Dios. Los dilemas morales de la ciencia*. Cátedra de la Divulgació de la Ciència. Publicacions de la Universitat de València.
- Machuca Moreno, L. E. (Julio-agosto, 2025). La inteligencia artificial y la administración de justicia. *Ciencia Latina*, 9(4), 12256-12274. https://doi.org/10.37811/cl_rcm.v9i4.20019

- Marías, J. (1994). *Mapa del mundo personal* (1.^a reimp.). Alianza.
- Marchena, M. (Julio 13, 2022). Marchena, contra las «dictaduras digitales»: un robot no puede sustituir a un juez. *El Independiente*. <https://www.elindependiente.com/espana/tribunales/2022/07/13/marchena-contras-las-dictaduras-digitales-un-robot-no-puede-sustituir-a-un-juez/#:~:text=Manuel%20Marchena%20G%C3%B3mez%2C%20presidente%20de%20la%20Sala%20Segunda,rosto%20humano%22%2C%20subrayando%20que%20%22no%20todo%20es%20digitalizable%22>
- Maybin, S. (Octubre 17, 2016). ¿Cómo en Estados Unidos las matemáticas te pueden meter en prisión? *BBC News*. <https://www.bbc.com/mundo/noticias-37679463.amp>
- Montesquieu, Ch. L. S. (1964). *De l'esprit des lois (1748-1750). Œuvres complètes* (4 vol., vol. 1). Ed. du Seuil.
- Nemitz, P. (Julio-agosto, 2021). La democracia en la era de la inteligencia artificial. *Nueva Sociedad*, (294), 130-294. <https://share.google/otnTeL1pmEIQGAqHj>
- Organización para la Cooperación y el Desarrollo Económicos (2019). Recomendación del Consejo sobre inteligencia artificial. <https://legalinstruments.oecd.org/api/download/?uri=public/db5053b5-93e0-4cf5-a7cf-edce5ee6e8933.pdf>
- OCDE/CAF (Septiembre 14, 2022). Uso estratégico y responsable de la inteligencia artificial en el sector público de América Latina y el Caribe. En *Estudios de la OCDE sobre Gobernanza Pública*. OECD Publishing. <https://doi.org/10.1787/5b189cb4-es>
- Ortega y Gasset, J. (2014). *Ensimismamiento y alteración. Meditación de la técnica y otros ensayos*. Alianza.
- Parens, E. (1998). Is better always good? The Enhancement Project. En Parens, E. (Ed.). *Enhancing Human Traits*, 1-28. Georgetown University Press.
- Parens, E. (2009). Toward a More Fruitful Debate About Enhancement. En Savulescu, J. and Bostrom, N. (Eds.). *Human Enhancement*, 181-197. Oxford University Press.
- Parlamento Europeo. (Junio 14, 2023). Enmiendas aprobadas 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 (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)), P9_TA(2023)0236, A9-0188/2023. https://www.europarl.europa.eu/doceo/document/A-9-2023-0188_ES.html
- Parlamento Europeo. (2014-2019). Resolución, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión Europea sobre normas de Derecho civil sobre robótica (2015/2103/inl). En Normas de Derecho civil sobre robótica. https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_ES.docx
- Parlamento Europeo y Consejo de Europa. (Junio 13, 2024). Reglamento 2024/1689 por el que se establecen normas armonizadas en materia de inteligencia artificial y se modifican los Reglamentos (CE) n.º 300/2008, (UE) n.º 167/2013, (UE) n.º 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828 (Reglamento de Inteligencia Artificial) (Texto pertinente a efectos del EEE). <https://www.boe.es/doue/2024/1689/L00001-00144.pdf>
- Parodi Remón, C. (Junio, 2004). El Derecho en Antígona ¿Natural o positivo? *DIKAIOSYNE*, (12), 109-119. <https://www.saber.ula.ve/bitstream/handle/123456789/19050/?jsessionid=B3AC9C79F70DB132CD5D69BF3AA2F1AF?sequence=2>
- Ponce Solé, J. (2024). *El Reglamento de inteligencia artificial de la Unión Europea de 2024, el derecho a una buena administración digital y su control judicial en España*. Marcial Pons.
- Parlamento Europeo y Consejo de Europa. (2024). Reglamento de inteligencia artificial de la Unión Europea. Reglamento (UE) 2024/1689, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial y se modifican los Reglamentos (CE) n.º 300/2008, (UE) n.º 167/2013, (UE) n.º 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828 (Reglamento de inteligencia artificial). <https://www.boe.es/doue/2024/1689/L00001-00144.pdf>

- Rayón Jerez, A. (2025). El accidente del Columbia y el pensamiento racional. *Digital&Data* [newsletter]. <https://digitaldata.substack.com/p/223-powerpoint-el-accidente-del-columbia>
- Rodríguez Carrillo, J. M. (2022). Inteligencia artificial y su aplicación en los sistemas de justicia en América Latina. Instituto Belisario Domínguez, Senado de la República de México. <http://bibliodigitalbd.senado.gob.mx/handle/123456789/5594>
- Romañach Cabrero, J. (2016). Las propuestas éticas y sociales del transhumanismo y los derechos humanos, *Universitas*, (24), 2-38. <https://dx.doi.org/10.20318/universitas.2016.3175>
- Ruiz Manero, J. (2005). Principios, objetivos y derechos. Otra vuelta de tuerca. *Doxa. Cuadernos Electrónicos de Filosofía del Derecho*, (28), 341-365. <https://ruaua.es/server/api/core/bitstreams/d7156719-87b5-489d-8f43-809d120abc32/content>
- Rubio Llorente, F. (mayo, 2002). La Carta Europea de los Derechos. *Claves de razón práctica*, (122), 4-11.
- Sánchez Rodríguez, L. P. (Septiembre 8, 2023). *XXIV dictamen de la Comisión Iberoamericana de Ética Judicial sobre el uso ético de la inteligencia artificial en la labor jurisdiccional*. https://eticayvalores.poder-judicial-go.cr/images/CIEJ/Dictamen_CIEJ_24.pdf
- Sancho Gargallo, I. (2022). *El paradigma del buen juez*. Tirant Lo Blanch.
- Schauer, F. (2004). *Las reglas en juego*. Marcial Pons.
- Simón, P. (Junio 16, 2022). Por qué emplear la inteligencia artificial en la administración de justicia es un asunto delicado. *The Conversation*. <https://theconversation.com/por-que-emplear-la-inteligencia-artificial-en-la-administracion-de-justicia-es-un-asunto-delicado-173678>
- Soler, J. (Febrero 21, 2018). El imperio del placer. *El País*. Tribuna. https://elpais.com/elpais/2018/02/21/opinion/1519233035_395860.html
- Stavropoulos, N. (2016). Legal Interpretivism. *Problema*, 1(10), 23-61. <https://doi.org/10.22201/ijj.24487937e.2016.10.8194>

- Terán Haughey, M. (Mayo 13, 2024). La razón por la que el Gobierno no quiere mostrar el código del algoritmo que decide quién recibe el bono social para la luz. *El Economista*. <https://www.eleconomista.es/tecnologia/noticias/12812123/05/24/la-razon-por-la-que-el-gobierno-no-quiere-mostrar-el-codigo-del-algoritmo-que-decide-quien-recibe-el-bono-social-para-la-luz.html>
- The Technolawgist. (Junio 12, 2019). Estonia se prepara para tener «jueces robot» basados en inteligencia artificial. The Technolawgist. <https://thetechnolawgist.com/2019/06/12/estonia-se-prepara-para-tener-jueces-robot-basados-en-inteligencia-artificial/>
- Unesco. (Noviembre 23, 2021). Recomendación sobre la ética de la inteligencia artificial. <https://www.unesco.org/es/legal-affairs/recommendation-ethics-artificial-intelligence>
- Vallespín, F. (Octubre 7, 2018). La democracia es frágil. *El País*. https://elpais.com/internacional/2018/10/05/actualidad/1538751614_762756.html?outputType=amp
- Zhabina, A. (Enero 20, 2023). Las cortes de China ya utilizan inteligencia artificial para resolver casos. *El Deber*. <https://www.dw.com/es/las-cortes-de-china-ya-utilizan-inteligencia-artificial-para-resolver-casos/a-64471873>