

HURI-AGE

Red Tiempo de los Derechos



Papeles el tiempo de los derechos

RESPUESTA JURÍDICA, TUTELA JUDICIAL E INTELIGENCIA ARTIFICIAL

Miguel de Asís Pulido

Doctorando del Programa de Derecho y Ciencias Sociales de la Universidad Nacional de Educación a Distancia

Palabras Clave: Inteligencia artificial jurídica, tutela judicial, garantías procesales, acceso a la justicia.

Key Words: Legal artificial intelligence, fair trial, due process, Access to justice.

Número: 2 Año: 2024

ISSN: 1989-8797

Comité Evaluador de los Working Papers “El Tiempo de los Derechos”

María José Añón (Universidad de Valencia)
María del Carmen Barranco (Universidad Carlos III)
María José Bernuz (Universidad de Zaragoza)
Rafael de Asís (Universidad Carlos III)
Eusebio Fernández (Universidad Carlos III)
Andrés García Inda (Universidad de Zaragoza)
Cristina García Pascual (Universidad de Valencia)
Isabel Garrido (Universidad de Alcalá)
María José González Ordovás (Universidad de Zaragoza)
Jesús Ignacio Martínez García (Universidad of Cantabria)
Antonio E Pérez Luño (Universidad de Sevilla)
Miguel Revenga (Universidad de Cádiz)
Maria Eugenia Rodríguez Palop (Universidad Carlos III)
Eduardo Ruiz Vieytes (Universidad de Deusto)
Jaume Saura (Instituto de Derechos Humanos de Cataluña)

Respuesta jurídica, tutela judicial e inteligencia artificial¹

Legal response, fair trial and artificial intelligence

Miguel de Asís Pulido

Doctorando del Programa de Derecho y Ciencias sociales de la Universidad Nacional de
Educación a Distancia
mdeasis@der.uned.es

Palabras clave: Inteligencia artificial jurídica, tutela judicial, garantías procesales, acceso a la justicia

Key words: Legal artificial intelligence, fair trial, due process, Access to justice

El futuro ya está aquí: la inteligencia artificial ha llegado a los tribunales. Ejemplos de ello son COMPAS, el sistema de evaluación de riesgo de reincidencia utilizado en la toma de medidas cautelares en los juicios de los Estados Unidos; PROMETEA, una herramienta desarrollada por la Fiscalía de Buenos Aires para asistir al Fiscal General a través de la generación de informes legales en casos simples en los que se solicita amparo al Tribunal Superior de Justicia; o PretorIA, sistema de IA predictiva que desde 2020 establece prioridades entre los casos que llegan a la Corte Constitucional de Colombia. Pero también podemos mencionar ejemplos regulativos: en marzo de 2024 se aprobó en Europa el Reglamento de inteligencia artificial de la UE (RUEIA), que otorga la consideración de alto riesgo a ciertas herramientas de IA jurídica, y, antes, en diciembre de 2023, se aprobó en España el RD-Ley 6/2023, que, entre otras cosas, habla de decisiones automatizadas, asistidas y proactivas en el ámbito del procedimiento judicial.

Podemos decir, así, que la IA está cambiando la forma en que se administra la justicia en nuestras sociedades. Pero hoy todavía existe una dualidad en la justicia. Por un lado, tenemos la mencionada implementación gradual de nuevas herramientas tecnológicas de vanguardia, pero, por otro, también existe un amplio margen de mejora en lo que respecta a la digitalización. Este último hecho no debería llevarnos a cometer el error de ignorar las posibles consecuencias que podrían conllevar en la práctica jurídica lo digital y la IA, esta última con su particular automatización de las actividades humanas, o, al menos, con la revolución que provoca en dichas actividades a través de su función asistencial. Si bien el mencionado RUEIA aborda la cuestión desde una perspectiva general y de riesgo, creo que se hace preciso una regulación específica de la IA para la materia del Derecho.

En esta comunicación intentaré establecer los fundamentos para un derecho a la tutela judicial efectiva adaptado a la era tecnológica-algorítmica, una era en desarrollo que hará tambalear los

¹ Esta publicación es parte del proyecto de I+D+i “Inteligencia artificial jurídica y Estado de Derecho” [PID2022 – 139773OB-I00], financiado por MICIU/AEI/10.13039/501100011033 y por FEDER, UE.

cimientos mismos de la justicia. Tomaré como base el derecho a la tutela judicial efectiva del artículo 24 CE. Tras unas pocas palabras acerca de este derecho fundamental, esbozaré una vía para su posible adaptación a la era algorítmica, a fin de incluir nuevas garantías clave para la vigencia de los principios procesales en una justicia atravesada, como decimos, por la digitalización y la IA. Con esta adaptación podremos empezar a hablar de un verdadero derecho al debido proceso tecnológico.

Gran parte de la doctrina ha interpretado que el artículo 24 CE (que no reproduciremos aquí por cuestión de espacio) reconoce dos derechos distintos: el derecho de acceso a la justicia, es decir, el derecho de las partes litigantes a ser escuchadas y a obtener una decisión basada en derecho, sin posibilidad de que se produzca cualquier situación de indefensión; y el derecho a un proceso con las debidas garantías. Sin embargo, opino que el contenido del artículo obliga a considerarlo como reconocedor de un *único derecho*. Esto se debe a que, sin el debido respeto a las garantías procesales, inevitablemente se causaría una situación de indefensión a los intervenientes en el proceso y no se garantizaría un acceso real a la justicia. Para mí, este único derecho debería denominarse "derecho al debido proceso". "Debido" porque, por la dignidad que reviste cada *persona* en un OJ, es un *deber* lo que guía al juez y al operador jurídico en su función de aplicador del derecho, deber de justicia, y *esta justicia se puede resumir en un dar a cada uno lo que legítimamente le corresponde*. El mandato de deber en el "debido proceso" es la concretización moderna del *suum cuique tribuere* de Ulpiano.

Por otra parte, asumiendo una perspectiva realista sobre la IA -que deberíamos denominar IAs, pues cada arquitectura de esta tecnología da lugar a herramientas distintas que portan retos diferentes para el Derecho-, asumiendo también que la IA y lo digital forman ya parte -y cada vez más- de nuestra justicia, podemos afirmar que este derecho al debido proceso debe extenderse hacia el proceso tecnológico. Es así como, en esta era algorítmica, podemos pensar en un "derecho al debido proceso tecnológico".

Es momento de dilucidar, si quiera brevemente, en qué consistirá este derecho, aludiendo en primer lugar a algunas garantías procesales tradicionales y cómo se están viendo influenciadas por las nuevas tecnologías. El plato fuerte, digamos, llegará después: entonces hablaré de cuatro nuevas garantías que van más allá de las ya existentes, así como de cuatro principios generales que deben incorporarse al derecho al debido proceso en la era tecnológica.

En lo relativo al ya clásico *acceso a la justicia*, la virtualización del espacio público ha llevado a una reducción de los tiempos y las distancias, facilitando un acceso más eficiente al sistema público, incluyendo la justicia, que, por otra parte, aún no está completamente virtualizada, como ya hemos mencionado. Poco a poco se irán resolviendo ciertos problemas técnicos y de eficiencia, y lo digital será el canal principal de la justicia (vid. EJE, sede judicial electrónica, LexNet, Minerva NOJ, REAJ, etc.). Sin embargo, la resolución de problemas técnicos no conllevará la ausencia absoluta de problemas. Uno de ellos es que lo virtual puede considerarse como algo menos "real" que lo físico (por ejemplo, llevar a cabo audiencias judiciales a través de videoconferencia puede implicar una disminución en la calidad de la interacción humana). Además, la *brecha digital* puede generar desigualdades en el acceso a la justicia para aquellos que no saben o no pueden utilizar medios digitales. Más allá de la digitalización, la automatización de algunas actividades procesales puede ser positiva para el acceso a la justicia: herramientas de asistencia a personas con discapacidad, de creación de documentos legales o sistemas digitales de gestión de expedientes con IA incorporada pueden ayudar a los ciudadanos a acceder a la AJ. Pero la automatización también puede ser negativa: la implementación de sistemas predictivos de éxito procesal -basados en *machine learning* (ML)- por parte de despachos de abogados e incluso de la oficina judicial, por

ejemplo, plantea desafíos, ya que podrían resultar en la exclusión de algunos litigantes basándose en el resultado de predicciones algorítmicas.

Por su parte, las nuevas tecnologías también presentan una doble cara en lo que respecta al *derecho a la defensa*, que, de nuevo, puede ser fortalecido y debilitado con su implementación: las herramientas de asistencia a la argumentación jurídica proporcionan a los profesionales del derecho recursos adicionales para sustanciar sus casos; pero la implementación de sistemas de IA predictiva y de sistemas de automatización de la respuesta judicial podrían dejar al ciudadano indefenso. Incluso los propios sistemas de asistencia a la argumentación jurídica, en caso de ser utilizados por los jueces, podrían dar lugar a respuestas injustas cuando el juez incurriera en el conocido *sesgo de automatización* -prestando una confianza ciega a los resultados de la máquina-.

El *derecho a una sentencia motivada* está influenciado en mayor medida por las tecnologías de IA de ML, que en algunos casos pueden llevar a resultados basados en predicciones carentes de una explicación clara dada la complejidad del algoritmo (el problema de la caja negra o *blackbox*). Esto plantea interrogantes sobre la transparencia y la comprensión del razonamiento detrás de las decisiones judiciales. Pero también presentan un riesgo para este derecho las tecnologías de otra IA distinta, que podríamos denominar IA lógica, formada por una serie de reglas deductivas que a veces pueden resultar excesivamente simples.

El *derecho a practicar la prueba* se transforma con la introducción de nuevo material probatorio basado en IA -biometría, predicciones, etc.-. Además, la IA también impulsará sistemas automatizados de examen de pruebas, sistemas que podrán cometer errores. Aunque el humano también los cometa, la IA carece del imperativo subjetivo de justicia inherente al razonamiento jurídico humano, que debemos suponer presente en los jueces si verdaderamente confiamos en el fenómeno jurídico². Además, será necesario dotar a los tribunales de tecnologías de detección de IA para evitar riesgos como los *deep fakes* probatorios. Por otro lado, los sistemas de traducción automática, un campo avanzado dentro de la IA gracias a los modelos generativos de *procesamiento de lenguaje natural*, también contribuirán a fortalecer otras garantías, como el *derecho a un intérprete*.

En fin, en lo referido a la *presunción de inocencia*, existe hoy en día el peligro de la extensión de la justicia actuarial, un campo basado en modelos predictivos y estadísticos de perfilado, que generalmente toman el riesgo como criterio identificativo y que plantean desafíos éticos y jurídicos de gran calibre (véase el caso COMPAS).

Todas estas garantías, así, deberán ampliar su significado y sus efectos para enfrentar los desafíos de las nuevas tecnologías. Recordemos que estos serán diferentes dependiendo del tipo de IA del que estemos hablando, así como del tipo de problema jurídico -o microsistema-, del tipo de proceso e incluso del tipo de decisión dentro del mismo proceso.

No obstante, creo que la singularidad de los desafíos asociados a la IA hace necesarias una serie de nuevas garantías y principios. Podemos dividirlos en dos grupos:

- a) Cuatro garantías, que serán las *garantías procesales tecnológicas específicas*. Estas tomarán su fundamento del mandato general de prohibición de la indefensión -como todas las demás garantías procesales tradicionales- y tendrán que ver con la ética de la IA: garantía

² Con ello no se descarta la problematización de este aserto: de hecho, una de las principales conclusiones a la que se llega en el estudio de la implementación de la IA es que esta implementación ha de venir aparejada con una mayor exigencia de la ética judicial, de su formación iusfilosófica y deontológica. En caso contrario el humano no sacará ventaja al masivo procesamiento de información que traen consigo las inteligencias artificiales.

de transparencia y explicabilidad; garantía de debido diseño y exigencia de responsabilidad; garantía de igualdad de armas tecnológicas; y garantía de supervisión y revisión humana.

b) Cuatro *principios generales del proceso tecnológico*, que vendrán de la misma naturaleza del Derecho y la defensa de la dignidad humana, base de todo fenómeno jurídico del presente. Estos cuatro principios serán: el respeto a los límites impuestos por la naturaleza de la respuesta jurídica y los microsistemas del Derecho; el respeto de la reserva de humanidad; el principio de acceso y libertad de sometimiento, y el principio de humanidad en el encuentro procesal -o humanidad en el trato-.

La *garantía de transparencia y explicabilidad* puede resumirse como la exigencia de comprensibilidad de los “actos” de la máquina (los cálculos que realice su algoritmo) cuando se utilicen herramientas de IA en los procesos judiciales. En su interpretación ingenua, esta garantía conllevará la transparencia del código fuente. Sin embargo, dadas las dificultades para comprender el lenguaje de programación, una interpretación realista pondrá mayor énfasis en la explicabilidad que en la transparencia, dando entrada al concepto de *Inteligencia Artificial Explicable* (XAI). Podemos tomar como base el derecho a la información significativa del artículo 15 del RUEIA, interpretándolo como un derecho a la explicación. La garantía en cuestión obligará a que todo sistema de IA utilizado por la Administración de Justicia o que afecte a las estrategias procesales -abogados, fiscal, etc.- puede proporcionar una explicación técnica *accesible* al público, permitiendo, en última instancia, el seguimiento y la trazabilidad de sus resultados. Como podemos imaginar, los sistemas de la IA que hemos denominado IA lógica tendrán menos problemas para cumplir esta exigencia que ciertos sistemas de ML -redes neuronales o SVM-, por lo que estos últimos deberán adoptar estrategias de explicabilidad efectivas o, de lo contrario, quedarán fuera del proceso judicial. Se requerirá en cualquier caso un registro operativo similar al propuesto para los sistemas de alto riesgo por el artículo 12 del RUEIA.

Además de la explicabilidad, deberán garantizarse otros puntos fundamentales de todo sistema de IA y digital, máxime en su aplicación procesal y, en especial, en su aplicación por parte del juez y del fiscal: *precisión, solidez, imparcialidad, seguridad del sistema, calidad de los datos, protección de los mismos y no discriminación*. Una forma de garantizarlos podría ser la creación de un sistema de registro público y certificaciones de calidad como el descrito en el Capítulo V del Título III del RUEIA, siendo quizá oportuno un registro específico para los sistemas algorítmicos utilizados en el contexto judicial, donde se incluirá información relativa a los desarrolladores, la entidad implementadora, los datos con los que se ha entrenado al sistema -en caso de ML-, procesos de validación, metodologías de programación, información sobre las capacidades y limitaciones del sistema, tareas para los que ha sido creado, el grado de implicación humana que se prevé en el momento de la implementación, etc. Quizá sea conveniente la evaluación desde el diseño de estos sistemas por un comité ético multidisciplinario, como sugiere la Recomendación de la UNESCO sobre la ética de la Inteligencia Artificial -art. 123-, que incluya juristas, filósofos éticos e informáticos. Además, se requerirán garantías de seguridad, como protocolos para prevenir y gestionar ataques informáticos. Otro punto fundamental serán los mecanismos para establecer responsabilidades, aspectos complejos en el ámbito de la Inteligencia Artificial, especialmente cuando se utiliza por la Administración de Justicia, donde la responsabilidad del juez se encauzará, generalmente, a través de figuras como el *error judicial* y el sistema de recursos, dando lugar, en los casos más extremos, a responsabilidad disciplinaria. Estas garantías, protocolos y mecanismos podrán explorarse en entornos cerrados: los famosos “Sandbox regulatorios”, que facilitan el desarrollo, las pruebas y la validación de los sistemas algorítmicos para cumplir con la normativa vigente.

Con el advenimiento de las tecnologías en el proceso será imprescindible, por otro lado, garantizar una nueva dimensión del derecho a la paridad de armas procesales, denominada *paridad de armas tecnológicas*, que puede considerarse una *garantía hermana* (pero no idéntica) respecto a aquella paridad tradicional: el mandato de igualdad de recursos tecnológicos implicará la extensión del bien jurídico protegido en la garantía de igualdad procesal. Incluso en el caso de que las armas tecnológicas se incluyesen en lo que se ha considerado tradicionalmente como armas procesales, la garantía de igualdad de armas -tradicional y tecnológica- deberá explicitarse como una nueva garantía en el texto constitucional, sin limitarse a ser una consecuencia implícita de la cláusula general de la prohibición de indefensión. Con ello se establecería la prohibición categórica de que la brecha y disparidad digital pueda resultar en una violación de aquella cláusula.

Como última garantía tecnológica específica, se tendrá que blindar la *supervisión y revisión humana* en todos los problemas y decisiones jurídicas en los que se implementen herramientas de inteligencia artificial, especialmente en las actividades de la oficina judicial, pero no solo: también se deberá aplicar a los servicios legales proporcionados por abogados y procuradores, así como a las tareas de la Fiscalía. Se habrá de reconocer, en línea con los documentos éticos de la UE, el derecho del juez a distanciarse de los resultados del programa, consecuencia fundamental del respeto por la autonomía humana -y por la humanidad de la justicia-. En términos generales, esta garantía está incluida en el artículo 14 del RUEIA para todo sistema de alto riesgo, y lo que se trata de proteger con ello es que la decisión final siempre sea tomada por un ser humano. Sin embargo, por la propia naturaleza de la respuesta jurídica, esta *decisión final* puede anticiparse en ciertos microsistemas cerrados del Derecho al momento de traducir la ley al código informático que automatice la respuesta. En cualquier caso, la supervisión exigirá también la puesta en conocimiento al usuario de la justicia sobre el uso de la IA en la resolución de su litigio -ya sea que la IA haya automatizado la respuesta o que el juez se haya visto asistido por el sistema algorítmico-, así como exigirá también la presencia humana en los recursos judiciales -revisión-.

En lo que respecta a los cuatro principios generales, es preciso comenzar afirmando que un Estado eficiente y sostenible que no respeta los Derechos y Libertades Fundamentales no cumple con su función democrática, ni siquiera con su primordial función jurídica de administrar y lograr la *justicia*. De este axioma se derivan los siguientes cuatro principios generales aseguradores de la eficacia del debido proceso -y, con ello, la justicia- en la Era algorítmica.

Para entender los principios primero y segundo, relativos al *respeto a (1) los límites impuestos por la naturaleza de la respuesta jurídica y los microsistemas del Derecho y a (2) la reserva de humanidad*, debemos aclarar que las cuatro garantías específicas mencionadas, que imponen controles o requisitos, se aplican tanto a los sistemas de apoyo de los operadores legales como a aquellos sistemas que los sustituyen en determinadas tareas y decisiones del proceso judicial -automatización³. Es el momento de introducir, si quiera brevemente por la limitación de espacio, una distinción de los tipos de problemas jurídicos que va más allá de las teorías de los casos sencillos y casos complejos de Hart, Fuller, McCormick o Gardner. Hablaremos entonces en términos de microsistemas -o universos de casos y soluciones relativos a un ámbito jurídico- y diremos que en algunos de estos microsistemas el razonamiento del Derecho contenido en la respuesta jurídica puede reducirse a reglas lógicas en algún momento de la derivación de las consecuencias jurídicas y de la determinación y clasificación de los hechos, y, con ello, la respuesta

³ Mientras que las *garantías específicas* se aplican a todo tipo de tareas -instrumentales, procesales y de decisión- de la oficina judicial y fiscal y de los despachos, en la aplicación de los *principios generales* se tendrá que distinción entre tareas meramente instrumentales -donde no se aplicarán los dos primeros principios- y tareas procesales y decisorias, a las cuales se aplicarán todos los principios.

podrá automatizarse. Llamaremos al conjunto de estos microsistemas “microsistemas cerrados o parcialmente cerrados”. Los casos que pertenecen a este conjunto corresponderían, en mayor o menor medida⁴, a controversias en las que las normas se aplican al caso como si se tratara de un silogismo deductivo -a modo del Juez Júpiter descrito por F. Ost-, ya que las disposiciones contenidas en la norma subsumen el hecho particular de manera casi mecánica. En ellos, los sistemas de automatización consistirán en IA sin ML, IA lógica, y solo en ellos la respuesta jurídica podrá realizarse sin la supervisión del juez o funcionario. Esto no significa la no intervención del juez, sino su intervención previa (diseño) y posterior (recursos), en base a la garantía específica de supervisión humana.

Por otra parte, los microsistemas abiertos (M4) son aquellos que implican razonamiento complejo y no se cierran en un sentido estrictamente lógico. Cuando involucran principios jurídicos superiores que requieren un equilibrio de derechos y principios fundamentales, teniendo en cuenta las circunstancias particulares del caso, se denominan M4b; si no involucran estos principios se denominarán M4a. En los M4 la respuesta no puede ser automatizada, pero en los M4a se puede usar la IA como asistente. Sin embargo, las decisiones en los procesos de microsistemas M4b estarían prohibidas para los sistemas de Inteligencia Artificial.

A partir de lo dicho, podríamos entender mejor nuestro primer principio general (1): la automatización, en el sentido de una aplicación del derecho sin intervención humana en el mismo momento de la respuesta solo podría ocurrir en el conjunto de los microsistemas cerrados o parcialmente cerrados (solo en casos simples). Los sistemas de IA asistenciales podrían aplicarse, a primera vista, en la respuesta a los problemas jurídicos de todos los microsistemas -cerrados y abiertos-, y cualquier uso de IA deberá cumplir con las garantías procesales específicas mencionadas.

Sin embargo, de este primer principio se deriva otro: el respeto a la reserva de humanidad (2). En general, esta reserva consistiría en prohibir la intervención de sistemas algorítmicos de automatización de respuestas en ciertos sectores de la sociedad, especialmente aquellos relacionados con los Derechos Humanos, vinculándose estrechamente con el mandato constitucional básico del artículo 10 CE. En base a ella el derecho al debido proceso tecnológico deberá mencionar que el uso de sistemas algorítmicos en el proceso judicial no podrá extenderse a los microsistemas M4b. Esta reserva podrá aminorarse en un futuro si se garantizase el uso de sistemas de IA asistenciales en tareas procesales como elementos *absolutamente secundarios en el razonamiento del juez*.

En los microsistemas en los que sea posible aplicar la IA y la digitalización, será efectivo un tercer principio: el de *acceso y libertad de sometimiento*, principio bifronte que blinda, por una parte, el acceso universal a los procesos automatizados y, por otra, la posibilidad de rechazar la participación en ellos. En cuanto al acceso, este principio se refiere en un primer sentido a la garantía de *acceso a la justicia*, pero a la luz de la existencia de la "brecha digital": por lo tanto, este principio eleva la importancia de la misma, y deberá provocar la inclusión de una cláusula explícita en relación con la garantía de *acceso tecnológico a la justicia*. En otro sentido, el principio de acceso subraya la importancia de evitar una confianza excesiva en las herramientas predictivas de éxito, destacando la necesidad de asegurar el acceso a la justicia para todos aquellos con un interés legítimo. *La libertad de sometimiento*, por otro lado, enfatiza la capacidad de rechazar la participación en procesos automatizados y recurrir a un proceso tradicional y humano en los procedimientos automatizados

⁴ Podemos mencionar los M1 o microsistemas completamente cerrados, M2 o microsistemas cerrados en la derivación de las consecuencias, M3 o microsistemas abiertos con casos sencillos -parcialmente cerrados-.

por sistemas algorítmicos, principio recogido en el artículo 22 del RGPD y recientemente en el art. 14 RUEIA.

El último principio general será la *exigencia de un trato humano* durante todo el proceso. Este principio se diferencia de la garantía específica de supervisión humana en que esta última adopta un sentido particular de control de los resultados de la máquina, mientras que el principio que ahora se trata se refiere a la presencia general de operadores jurídicos humanos que acompañen a la persona a lo largo del proceso judicial. *Automatizar no podrá significar deshumanizar*. Aunque pueda resultar una afirmación paradójica, lo que hace este principio es dar cuenta de la importancia del factor humano en las interacciones sociales sin renunciar al avance tecnológico. Siendo el proceso judicial una red ordenada de interacciones entre individuos, entre personas que participan en él como partes procesales que a menudo enfrentan situaciones estresantes y críticas, se hacen preciso, *acompañando* en ocasiones a la tecnología -procesos automatizables- o *acompañados* en otras por ella -procesos con IA asistencial-, jueces y operadores jurídicos que estén presentes, activos y disponibles durante todo el proceso. El operador jurídico debe aspirar a la excelencia en el ejercicio de su profesión, una profesión que tiene como objetivo fundamental la búsqueda de la justicia, y que no puede ignorar a la persona concreta involucrada en el proceso.