

**CONSTRUYENDO EL MARCO LEGAL EUROPEO PARA EL DISEÑO Y USO
DE HERRAMIENTAS DE INTELIGENCIA ARTIFICIAL EN EL ENTORNO DE
LA JUSTICIA CIVIL-ADMINISTRATIVA**

Fernando ESTEBAN DE LA ROSA*

Publicado en el libro

El derecho y la inteligencia artificial, coord. por Francisco Herrera
Triguero, Alfonso Peralta Gutiérrez y Leopoldo Salvador Torres López
Editorial Universidad de Granada, 2022,

ISBN 978-84-338-7049-0

*Catedrático de Derecho internacional privado.

Facultad de Derecho.

Universidad de Granada.

Email: festeban@ugr.es

F. Esteban de la Rosa, “Construyendo el marco legal europeo para el diseño y uso de herramientas de inteligencia artificial en el entorno de la justicia civil-administrativa”, en el libro F. Herrera Triguero, A. Peralta Gutiérrez y L.S. Torres López, Universidad de Granada, pp. 423-451.

CONSTRUYENDO EL MARCO LEGAL EUROPEO PARA EL DISEÑO Y USO DE HERRAMIENTAS DE INTELIGENCIA ARTIFICIAL EN EL ENTORNO DE LA JUSTICIA CIVIL-ADMINISTRATIVA

Fernando Esteban de la Rosa

Catedrático de Derecho internacional privado

**Director científico de la Unidad de Excelencia de Investigación Sociedad Digital:
Seguridad y Protección de Derechos**

Universidad de Granada

Resumen: Mientras que la idea del juez robot queda por ahora en el ámbito de la ciencia ficción, la realidad muestra un uso cada vez más extendido de las herramientas de IA en el ámbito jurídico, en los servicios legales e incluso por los propios sistemas de justicia. El marco legal actual, que reposa principalmente sobre el Reglamento General de Protección de Datos (RGPD), no resulta suficiente para establecer una protección eficaz frente a las necesidades de transparencia y de no discriminación con miras a excluir los riesgos de opacidad y de sesgos. Para conseguir la implantación de una IA fiable el trabajo afirma la necesidad previa de tener en cuenta la distinta naturaleza, posibilidades y límites de la tecnología de IA utilizada, aplaude que el marco europeo propuesto parta de la distinción entre diferentes contextos de uso de la IA en el sector de la justicia y pasa revista al impacto del RGPD sobre los diferentes usos de las herramientas de IA en el ámbito jurídico. Para esta finalidad el trabajo profundiza en el concepto central de intervención humana significativa.

I. Introducción

Aunque la ciencia ficción nos ha preparado para la idea de jueces-robot, basados en Inteligencia Artificial (IA), capaces de conocer y decidir sin los sesgos y las limitaciones propios de los seres humanos lo cierto es que esta clase de juez robot no resulta viable contando con la tecnología disponible en la actualidad.¹ Cuestión distinta es la utilización de máquinas de IA para la realización de tareas concretas. En este terreno las experiencias están comenzando a andar en los países miembros de la Unión Europea contando exclusivamente con el marco de protección de datos, un marco cuyas respuestas a las necesidades de la IA resultan insuficientes.² Para la regulación de la IA resultan de utilidad los principios generales contenidos en el artículo 5 RGPD (licitud,

¹ Véase R.W. Campbell, “Artificial Intelligence in the Courtroom: the delivery of Justice in the age of Machine Learning”, 18, COLO.TECH. L.J. 323, 2020, pp. 323-349; R.C. Cancio Fernández, “¿Sueñan los jueces con sentencias electrónicas? Do judges dream of electronic sentences?” Análisis Jurídico - Político Enero – Junio de 2020, Vol.2 Núm.3, Sección temática: inteligencia artificial. DOI: <https://doi.org/10.22490/26655489.3854>

² Véase A. Soriano Aranz, “Decisiones automatizadas : problemas y soluciones jurídicas. Más allá de la protección de datos”. Revista de Derecho público, Marcial Pons, 2021 doi: https://doi.org/10.37417/RPD/vol_3_2021_5352021, p. 108.

F. Esteban de la Rosa, “Construyendo el marco legal europeo para el diseño y uso de herramientas de inteligencia artificial en el entorno de la justicia civil-administrativa”, en el libro F. Herrera Triguero, A. Peralta Gutiérrez y L.S. Torres López, Universidad de Granada, pp. 423-451.

lealtad, transparencia, limitación de la finalidad y del plazo de conservación, minimización de datos, exactitud, integridad y confidencialidad), las prohibiciones de procesamiento de determinadas categorías de datos a que se refieren los artículos 9 y 10 RGPD, que son de interés para combatir las discriminaciones y los sesgos, y la proclamación del derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado cuando pueda producir efectos significativos sobre la persona (art. 22 RGPD). En cambio, no hay en el RGPD una regulación apta para hacer frente a la falta de transparencia, los sesgos, las discriminaciones o directamente las decisiones que provienen de la caja negra o *black box*. El énfasis de la regulación europea de protección de datos ha sido puesto en la protección de los datos de entrada al sistema y no tanto en los datos de salida, es decir, en las inferencias que realiza el sistema y que dan lugar a resultados.

Ante la falta de regulación hoy nos movemos en el ámbito del *soft law* para determinar cuáles deberían ser las garantías exigibles a los algoritmos que se utilizan en el ámbito de la administración de justicia y su entorno. De los documentos oficiales europeos, la Carta Ética Europea sobre el uso de la Inteligencia Artificial en sistemas judiciales y su entorno, aprobada el 3-4 de diciembre de 2018 por la Comisión Europea para la Eficiencia de la Justicia (CEPEJ) del Consejo de Europa,³ es quizás el que ha descrito de forma más clara los riesgos del uso de la IA en este ámbito. Según la Carta Ética los sistemas de IA ponen en juego el principio de primacía del derecho y pueden incentivar una nueva forma de normatividad que signifique límites a la discrecionalidad soberana del juez y conducir, a largo plazo, a una estandarización de las decisiones judiciales que podrían llegar a basarse en un cálculo puramente estadístico vinculado, por ejemplo, a la media de las indemnizaciones concedidas anteriormente por otros tribunales. Se pone también en duda la compatibilidad de soluciones basadas en IA con derechos individuales consagrados en el CEDH como el derecho a un juicio justo, el derecho al juez natural establecido por la ley, el derecho a un tribunal independiente e imparcial, así como a la igualdad de armas en los procedimientos judiciales. Junto a ello se duda sobre la posibilidad de conceder una protección adecuada a los datos y, como consecuencia, a la vida privada y familiar. El informe que acompaña a la Carta Ética concluye sugiriendo el gran potencial de la IA para ayudar a los profesionales del derecho en su trabajo, así como la ayuda que el sistema judicial podrá encontrar en la IA con aplicaciones destinadas a mejorar la investigación jurídica.⁴ A pesar de esta visión la Carta Ética acuñó cinco principios sobre el uso de la IA en los sistemas judiciales y su entorno a modo de principios para su gobernanza y uso. Entre ellos se encuentran el principio de respeto a los derechos fundamentales; el principio de no discriminación; el principio de calidad y seguridad; el principio de transparencia, imparcialidad y equidad, a fin de hacer el procesamiento de datos accesible y explicable, contando con auditorías externas; así como el principio bajo el control del usuario, con miras a excluir una aproximación prescriptiva y garantizar que los usuarios son actores informados y tienen el control sobre sus opciones.

³ Disponible en el siguiente enlace: <https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c>

⁴ European Ethical Charter, *cit*, p.15.

F. Esteban de la Rosa, “Construyendo el marco legal europeo para el diseño y uso de herramientas de inteligencia artificial en el entorno de la justicia civil-administrativa”, en el libro F. Herrera Triguero, A. Peralta Gutiérrez y L.S. Torres López, Universidad de Granada, pp. 423-451.

La Unión Europea lleva varios años trabajando en un marco para la gobernanza de la IA que se caracteriza por haber puesto al ser humano y a sus necesidades en el centro de atención. Como hitos en la construcción de este marco destacan la Comunicación de la Comisión Inteligencia artificial para Europa,⁵ el Libro Blanco sobre la inteligencia artificial - un enfoque europeo orientado a la excelencia y la confianza,⁶ y la Resolución del Parlamento Europeo (2020) relativa a un Marco de los aspectos éticos de la IA, la robótica y las tecnologías relacionadas.⁷ Por su gran impacto potencial en la actualidad resulta obligado tener en cuenta la propuesta de Reglamento de la Unión Europea de armonización de las normas sobre Inteligencia Artificial.⁸ La propuesta parte de la distinción entre prácticas de IA prohibidas, prácticas consideradas de alto riesgo y prácticas no consideradas de alto riesgo. En lo que interesa a la construcción del marco legal para las herramientas de IA que son utilizadas en el ámbito de la justicia civil-administrativa el Reglamento no incluye a la utilización de esas herramientas entre los usos prohibidos. Sin embargo, la propuesta de Reglamento sí traza una distinción entre usos de alto riesgo y usos de no alto riesgo que resulta de interés para el desarrollo de la IA en este sector. Según el apartado 40 del preámbulo “deben considerarse 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. En particular, a fin de evitar el riesgo de posibles sesgos, errores y opacidades, *procede considerar de alto riesgo aquellos sistemas de IA cuyo objetivo es ayudar a las autoridades judiciales a investigar e interpretar los hechos y el Derecho y a aplicar la ley a unos hechos concretos*. No obstante, dicha clasificación no debe hacerse extensiva a los sistemas de IA destinados a actividades administrativas meramente accesorias que no afectan a la administración de justicia en casos concretos, como la anonimización o seudonimización de las resoluciones judiciales, documentos o datos; la comunicación entre los miembros del personal; tareas administrativas, o la asignación de recursos”.

La propuesta de Reglamento establece en su capítulo tercero (artículos 16 a 29) las obligaciones a que quedan sometidos los proveedores de sistemas de IA de alto riesgo. En ese marco se contienen obligaciones de respeto a los valores y normas europeas para datos de entrenamiento; de documentación y trazabilidad de datos para garantizar la intervención regulatoria; de transparencia y suministro de información; de solidez y precisión a fin de conseguir una intervención ex ante adecuada; de supervisión humana para conseguir una IA centrada en el ser humano. La estrategia trazada por la Comisión ofrece una ventaja competitiva para la industria europea de la IA pues se contempla que las disposiciones del nuevo marco regulatorio sean aplicables a los proveedores, productores, importadores, distribuidores y usuarios de los sistemas de IA que se usen o produzcan esta tecnología en territorio europeo. El marco regulatorio, por tanto, se

⁵ Documento COM (2018) 237 final.

⁶ Documento COM (2020) 65 final, 19.2.2020.

⁷ Documento 2020/2012 (INL).

⁸ Documento COM (2021) 206 final.

F. Esteban de la Rosa, “Construyendo el marco legal europeo para el diseño y uso de herramientas de inteligencia artificial en el entorno de la justicia civil-administrativa”, en el libro F. Herrera Triguero, A. Peralta Gutiérrez y L.S. Torres López, Universidad de Granada, pp. 423-451.

muestra exigente con las empresas que hayan desarrollado los sistemas de IA en otra parte del mundo.

Para los sistemas de IA que no son de alto riesgo la propuesta de Reglamento pretende simplemente fomentar que los proveedores cumplan de manera voluntaria los requisitos que son obligatorios para los sistemas de IA de alto riesgo. La propuesta de Reglamento lanza la posibilidad de que los proveedores de sistemas de IA que no sean de alto riesgo puedan crear y aplicar sus propios códigos de conducta, los cuales podrían incluir compromisos voluntarios relativos, por ejemplo, a la sostenibilidad medioambiental, la accesibilidad para las personas con discapacidad, la participación de las partes interesadas en el diseño y el desarrollo de sistemas de IA, y la diversidad de los equipos de desarrollo.⁹ Se quiere así fomentar un ámbito de autorregulación cuyo éxito dependerá del compromiso de los propios proveedores y del efecto sobre los clientes de la existencia de sellos de calidad que garanticen un cumplimiento normativo al que no quedarían obligados.

La distinción de contextos de uso de los sistemas de IA es una aportación muy significativa de la propuesta pues se muestra capaz de definir niveles de exigencia y de garantía diferentes en dependencia de la potencial afectación de valores jurídicos y abre la posibilidad de no someter a exigencias desproporcionadas a aquellos sistemas de IA que son de riesgo bajo o nulo. El establecimiento de regímenes de gobernanza diferentes para los distintos contextos de uso es una propuesta que ya había sido mantenida por la doctrina.¹⁰ Como consecuencia, una misma herramienta de IA podrá quedar sometida a uno u otro régimen de gobernanza dependiendo de si se destina a un uso de alto riesgo (por ejemplo, asistir a los jueces a la hora de dictar una sentencia) o a un uso que no sea de alto riesgo (asistir a un profesional sobre cuál podrá ser el resultado de un litigio). Con arreglo al tenor actual de la propuesta de Reglamento el criterio a tener en cuenta para determinar el sometimiento del proveedor de IA a las obligaciones especiales que derivan de los sistemas de alto riesgo es que la finalidad de la herramienta sea ayudar a las autoridades judiciales a investigar e interpretar los hechos y el Derecho y a aplicar la ley a unos hechos concretos. Si la herramienta de IA sirve a otras funciones o tiene otros usuarios distintos de las autoridades judiciales (usuarios de la justicia, profesionales) podríamos salir de la caracterización de las herramientas de alto riesgo.

Pero el marco de gobernanza no puede ser construido de espaldas al conocimiento suficiente de las posibilidades y límites de las diferentes tecnologías de IA, a fin de

⁹⁹ Para un análisis crítico del texto puede verse L. Cotino Hueso, J. A. Castillo Parrilla, I. Salazar, R. Benjamins, M. Cumbreñas, A. M. Esteban, “Un análisis crítico constructivo de la Propuesta de Reglamento de la Unión Europea por el que se establecen normas armonizadas sobre la Inteligencia Artificial (Artificial Intelligence Act)”, Diario La Ley, 2 de julio de 2021. En: <https://diariolaley.laleynext.es/Content/Documento.aspx?params=H4sIAAAAAAAAAEAMtMSbF1C TEAAmMjS0NTS7WY1KLizPw8WYMDI0MDM2MDtbz8INQQF2fb0ryU1LTMvNQUkJLMtEqX OSQyoJU27TEEnOJUtdSk PxsFJPIYSYAAG0FMtpjAAAAWKE>

¹⁰ Para el ámbito de la responsabilidad aparejada al uso de sistemas inteligentes véanse las propuestas de J. Valls Prieto, *Inteligencia Artificial, Derechos Humanos y bienes jurídicos*, Aranzadi, Cizur Menor, 2021, pp. 107-109.

F. Esteban de la Rosa, “Construyendo el marco legal europeo para el diseño y uso de herramientas de inteligencia artificial en el entorno de la justicia civil-administrativa”, en el libro F. Herrera Triguero, A. Peralta Gutiérrez y L.S. Torres López, Universidad de Granada, pp. 423-451.

detectar en qué grado cumplen con las funciones que se le encomiendan. Por esta razón el siguiente epígrafe contiene un recordatorio de las diversas configuraciones de la tecnología de IA con indicación de sus posibilidades e indicación de sus límites en el ámbito objeto de este trabajo. Esta exposición servirá también para realizar una aproximación al modo en que las obligaciones del previsible futuro marco de gobernanza europeo podrán ser cumplidas teniendo en cuenta las posibilidades que ofrecen las diferentes tecnologías de IA para satisfacer los requisitos de transparencia y explicabilidad.¹¹ En la última parte se analizan los parámetros legales que en la actualidad acotan la posibilidad de admitir las decisiones automatizadas procedentes de IA en el ámbito de la justicia civil-administrativa. Para ello se llevará a cabo una aproximación al concepto de intervención humana significativa.

II. Posibilidades y límites de las diferentes tecnologías de IA en el entorno de la administración de justicia

1. Sistemas Expertos

Por sistema de inteligencia artificial (IA) se entiende el software, desarrollado a través de técnicas distintas, que permite generar información de salida como contenidos, predicciones, recomendaciones o decisiones con capacidad para influir en el entorno con el que interactúa de forma similar a como lo haría un ser humano. En esencia cabe distinguir entre los Sistemas Expertos, o la denominada IA simbólica, basados en reglas y los sistemas de IA basados en datos, grupos en los que se insertan otras subcategorías de sistemas de IA.¹²

En el enfoque basado en reglas el conocimiento de un dominio legal específico se representa como una colección de reglas de la forma SI <condición(s)> ENTONCES acción/conclusión. Este tipo de método pertenece a la primera ola de IA, caracterizada por el método deductivo y por el hecho de que puede ser programada por humanos.¹³ La codificación de reglas ha mostrado éxito en sectores particulares.¹⁴ La creación de árboles de decisión puede ser de utilidad para calificar la demanda y para situar a los usuarios ante posibles soluciones de su caso, sobre todo en entornos jurídicos

¹¹ En este sentido véase R.C. Cancio Fernández, “¿Sueñan los jueces con sentencias electrónicas? Do judges dream of electronic sentences?” Análisis Jurídico – Político, p. 149.

¹² La propuesta de Reglamento hace alusión a las siguientes técnicas y estrategias de Inteligencia artificial: a) Estrategias de aprendizaje automático, incluidos el aprendizaje supervisado, el no supervisado y el realizado por refuerzo, que emplean una amplia variedad de métodos, entre ellos el aprendizaje profundo. b) Estrategias basadas en la lógica y el conocimiento, especialmente la representación del conocimiento, la programación (lógica) inductiva, las bases de conocimiento, los motores de inferencia y deducción, los sistemas expertos y de razonamiento (simbólico); c) Estrategias estadísticas, estimación bayesiana, métodos de búsqueda y optimización.

¹³ Ejemplos de ello son la conducción bajo los efectos del alcohol en Victoria, Australia, y las Directrices de Sentencia de los Estados Unidos [USSG 18 USCS Appx.] Las Directrices para la imposición de penas de EE.UU. han sido modeladas por un sistema informático (basado en reglas).

¹⁴ Véase Marek J Sergot, Fariba Sadri, R.A. Kowalski, Frank R. Kriwaczek, Peter Hammond and H.T. Cory, “The British Nationality Act as a logic program”, (1986), Communications of the ACM.

F. Esteban de la Rosa, "Construyendo el marco legal europeo para el diseño y uso de herramientas de inteligencia artificial en el entorno de la justicia civil-administrativa", en el libro F. Herrera Triguero, A. Peralta Gutiérrez y L.S. Torres López, Universidad de Granada, pp. 423-451.

específicos.¹⁵ Como limitaciones para la construcción de Sistemas Expertos se encuentran las siguientes:

1. No son adecuados para ámbitos jurídicos caracterizados por el protagonismo de la discrecionalidad.¹⁶ No obstante, la posibilidad tecnológica de modelar reglas y árboles de decisión en ámbitos caracterizados por un fuerte componente de discrecionalidad ha sido estudiada y afirmada.¹⁷ En su evolución los Sistemas Expertos están siendo utilizados en ámbitos judiciales, como veremos, para codificar escenarios típicos que quedan asociados, mediante la navegación web, con alternativas de actuación o solución.

2. Esta metodología no resulta satisfactoria si los árboles de decisión resultan excesivamente complejos por lo que su utilización es recomendada en ámbitos jurídicos con variables limitadas.

3. El fallo en la programación puede provocar disfunciones si la máquina se revela incapaz de tener en cuenta las distintas circunstancias de la realidad.¹⁸ Si al fallo de programación se le añade el carácter automático de la decisión sin supervisión humana los resultados pueden llegar a ser devastadores tal y como pasó en el Centrelink "robodebt" de Australia. Los pagos de la asistencia social realizados sobre la base de los ingresos quincenales declarados por los propios beneficiarios se cruzaron con una estimación de los ingresos quincenales, tomada como media simple de los ingresos anuales declarados a la Oficina de Impuestos de Australia, y fueron utilizados para generar automáticamente avisos de deuda sin ningún otro escrutinio o supervisión humana. Esta suposición no concordaba con la forma en que se paga realmente a la mano de obra altamente eventual en Australia. Por ejemplo, a una diseñadora gráfica que no pudo encontrar trabajo durante nueve meses del año fiscal, pero que ganó 12.000 dólares australianos en los tres meses anteriores a junio, se le generó una deuda de forma automática. Todo ello a pesar de que no había tenido lugar ningún fraude y de que esta situación constituye exactamente el tipo de dificultad que Centrelink pretendía resolver.¹⁹

El desarrollo de máquinas de IA de este tipo requiere la "extracción de conocimiento", normalmente mediante la realización de entrevistas a expertos, a lo que suele seguir su codificación e integración en un sistema informático.²⁰ Un ejemplo de utilización de

¹⁵ R. Susskind, *Expert Systems in Law*, Clarendon Press, Oxford, 1987.

¹⁶ Véase Ettore Battelli, "La decisión robótica: algoritmos, interpretación y justicia predictiva", (2020) 38, *Revista de Derecho Privado* 45, 55-58.

¹⁷ Véase Ruth Kannai, Uri J. Schild and John Zeleznikow, "Modeling the Evolution of Legal Discretion - An Artificial Intelligence Approach", (2007) Vol. 20, Issue 4, *Ratio Juris*, pp. 530-558.

¹⁸ Véase [Guido Governatori, Jeffrey Barnes, John Zeleznikow, Louis de Koker, Marta Poblet, Mustafa Hashmi and Pompeu Casanovas Romeu, 2020 'Rules as Code' will let computers apply laws and regulations. But over-rigid interpretations would undermine our freedoms, https://theconversation.com/rules-as-code-will-let-computers-apply-laws-and-regulations-but-over-rigid-interpretations-would-undermine-our-freedoms-149992](https://theconversation.com/rules-as-code-will-let-computers-apply-laws-and-regulations-but-over-rigid-interpretations-would-undermine-our-freedoms-149992)

¹⁹ Véase la noticia en la siguientes web <https://tinyurl.com/y3dqe6mg>.

²⁰ Darin Thompson, "Creating New Pathways to Justice Using Simple Artificial Intelligence and Online Dispute Resolution", 2015, vol 1 (2) *International Journal of Online Dispute Resolution*, 4, 13.

F. Esteban de la Rosa, "Construyendo el marco legal europeo para el diseño y uso de herramientas de inteligencia artificial en el entorno de la justicia civil-administrativa", en el libro F. Herrera Triguero, A. Peralta Gutiérrez y L.S. Torres López, Universidad de Granada, pp. 423-451.

Sistemas Expertos en el ámbito judicial se encuentra en el Civil Resolution Tribunal (CRT) de la Columbia Británica, creado en Canadá en 2016. Se trata de un Tribunal online que permite a los usuarios obtener un diagnóstico de su caso utilizando la máquina denominada *Solution Explorer*, una máquina que ha sido preparada para ofrecer información sobre cada uno de los ámbitos sobre los que se extiende la jurisdicción del CRT. El *Solution Explorer* ofrece a los usuarios la posibilidad de escribir una carta formal a la parte reclamada proporcionándole la plantilla para esta finalidad y la información sobre las vías legales para la presentación de una demanda, guiando al usuario durante este proceso para evitar que se comentan errores. Para el diseño y construcción de la máquina se requirió solamente un equipo compuesto por dos personas: un ingeniero de conocimiento, responsable de la adquisición y recopilación del conocimiento de los expertos sobre una concreta materia y de hacer la estructura lógica interna de la máquina; y un especialista en contenidos, que se encarga de grabar la información, redactar los recursos informativos y asegurarse de que todo queda escrito en un lenguaje sencillo. Para la selección de los expertos los responsables del CRT prefirieron contar con personas que trabajaban directamente con el público pues el enfoque del *Solution Explorer* es ser utilizado sin contar con la asistencia de un abogado. El CRT cuenta con una persona de mantenimiento que se ocupa de actualizar su contenido en caso de necesidad, algo que es relativamente sencillo. Si los cambios a introducir afectan a la base de conocimientos y a la estructura lógica del árbol de decisiones las modificaciones pueden llevar más tiempo e implica la necesidad de una nueva actuación de los expertos con la finalidad de revisar el sistema.

2. Minería de datos

La minería de datos es una metodología de resolución de problemas que se basa en el descubrimiento de patrones y regularidades en un conjunto de datos.²¹ El programa evalúa el grado de repetición de determinados factores y, mediante un análisis iterativo, detecta las relaciones que podrían no ser apreciables para el análisis humano.²² Para resolver problemas, las tecnologías basadas en datos no aplican reglas ni tampoco la lógica humana. Se ocupan únicamente de procesar de forma matemática una gran cantidad de datos de los que se nutren para identificar asociaciones o correlaciones entre ellos. Los programas por tanto ni saben ni tampoco explican por qué existen esas relaciones, simplemente las identifican. La existencia y disponibilidad de una gran cantidad de datos ha incrementado el poder de las máquinas basadas en Machine Learning. A su vez, la técnica de las redes neuronales,²³ basada en una colección de unidades conectadas o nodos llamados neuronas artificiales que asemejan las neuronas de un cerebro biológico,²⁴ han conseguido imitar el modo en que los seres humanos

²¹ Véase Usama Fayyad, Gregory Piatetsky-Shapiro, and Padhraic Smyth, "From Data Mining to Knowledge Discovery in Databases", (1996), Volume 17, Number 3, AI Magazine.

²² Véase Quinlan, J. Ross. "Induction of decision trees." *Machine learning* 1.1 (1986): 81-106.

²³ Véase Andrew Stranieri, John Zeleznikow, Mark Gawler, and Bryn Lewis. "A hybrid rule–neural approach for the automation of legal reasoning in the discretionary domain of family law in Australia." *Artificial intelligence and Law* 7, no. 2 (1999): 153-183.

²⁴ Andrew Stranieri and John Zeleznikow, *Knowledge Discovery from Legal Databases* (Springer 2005).

F. Esteban de la Rosa, “Construyendo el marco legal europeo para el diseño y uso de herramientas de inteligencia artificial en el entorno de la justicia civil-administrativa”, en el libro F. Herrera Triguero, A. Peralta Gutiérrez y L.S. Torres López, Universidad de Granada, pp. 423-451.

procesan la información, permitiendo dar mayor peso a unos datos sobre otros. Los sistemas de IA de este tipo se corresponden con sistemas de autonomía decisional o autónomos cuya actividad consiste en valorar el supuesto de hecho concreto en función de lo que han ido aprendiendo e irán ajustando los resultados a medida que cambie el contexto.

Para el ámbito de la administración de justicia han sido identificadas las siguientes limitaciones para los sistemas de IA basados en datos:

1. El proceso de aprendizaje debe ser apoyado con un conjunto de datos de entrenamiento específico y bien definido en términos de formato, lenguaje y resto de características. La inexistencia de datos suficientes, o de datos no preparados, impide el establecimiento de relaciones estadísticamente válidas.²⁵ Cualquier proyecto de creación de una máquina basada en datos para el ámbito judicial debe contemplar de qué datos se nutrirá la máquina y el modo en que se consiguen datos válidos y adaptados a las necesidades de la máquina. Deberá tener en cuenta, además, cuáles son los parámetros principales de funcionamiento del algoritmo. Este criterio puede encuadrarse tanto en una interpretación actualizada (art. 3.1 CC) de la obligatoriedad de motivar las sentencias (art. 120.3 CE) como en el respeto al principio de transparencia en el tratamiento de los datos (art. 5.1.a RGPD) en la medida en que es el algoritmo el que ejecuta los criterios de dicho tratamiento de datos.

2. El modo en que el sistema se nutre de datos debe respetar las exigencias provenientes de la normativa de protección de datos, prestando especial atención al principio de limitación de la finalidad. Si se trata de datos extraídos para otras finalidades es necesario atender a la base de legitimación originaria y comprobar que el análisis que se va a realizar es un segundo uso compatible, ya que el hecho de que un dato sea de público acceso no significa *per sé* que pueda utilizarse para cualquier finalidad.²⁶ Cuando se trate de datos procedentes de resoluciones judiciales es necesario proceder a la anonimización de los datos, salvo, en nuestro país, las sentencias del Tribunal Constitucional, cuyos datos son públicos por mandato constitucional (art. 164 CE) con únicamente ciertas excepciones.²⁷ Si los datos se recaban directamente del titular, se le

²⁵ See Ray Worthy Campbell, “Artificial Intelligence in the Courtroom: the delivery of Justice in the age of Machine Learning”, Colorado Tech Law Journal, vol. 18.2, 2020, pp. 323-350.

²⁶ Un ejemplo reciente de usos ulteriores de datos incompatibles con la finalidad original para la que fueron recabados lo encontramos en la Resolución de la AEPD de 26 de abril de 2021, donde se multa a una entidad de análisis de solvencia patrimonial por utilizar datos de acceso público (BOE) para finalidades ulteriores no compatibles. En esta Resolución, la AEPD recuerda que (1) cualquier uso secundario de datos debe ser compatible con la finalidad original con que fueron recabados (principio de limitación de finalidad del tratamiento de datos, artículo 5.1.b RGPD), (2) debe tener una base de legitimación propia (no es suficiente alegar que son datos de fuentes accesibles al público), y que (3) debe notificarse al interesado el uso secundario de sus datos. El texto íntegro de la Resolución puede consultarse en: <https://www.aepd.es/es/documento/ps-00240-2019.pdf>

²⁷ No obstante, véase el Acuerdo de 23 de julio de 2015, del Pleno del Tribunal Constitucional, por el que se regula la exclusión de los datos de identidad personal en la publicación de las resoluciones jurisdiccionales, disponible en la siguiente web:

F. Esteban de la Rosa, “Construyendo el marco legal europeo para el diseño y uso de herramientas de inteligencia artificial en el entorno de la justicia civil-administrativa”, en el libro F. Herrera Triguero, A. Peralta Gutiérrez y L.S. Torres López, Universidad de Granada, pp. 423-451.

deberá informar de forma clara, en cumplimiento de los principios de transparencia y limitación de finalidad (art. 5.1.a y b RGPD), además de argumentar convenientemente la base de legitimación en que se apoya el tratamiento de datos.

El RGPD se aplica al tratamiento de datos personales por parte de los tribunales, incluidos los datos personales incluidos en las sentencias. Se considera dato personal “toda información sobre una persona identificada o identificable” (art. 4.1 RGPD) y se considera persona identificable toda aquella cuya identidad pueda determinarse directa o indirectamente (art. 4.2 RGPD). Para determinar el carácter identificable de una persona deben tenerse en cuenta todos los medios de que disponga el responsable del tratamiento o cualquier otra persona para dicha identificación, y que exista una probabilidad razonable de que se utilicen tales medios teniendo en cuenta el coste y tiempo necesarios para la identificación y la tecnología disponible en el momento en que ésta tenga lugar (art. 4.1 y Cons. 26 RGPD). Como consecuencia, son también datos personales las inferencias y aquellos datos no suficientemente anonimizados o agregados que permitan una identificación del sujeto relativamente sencilla. La evolución tecnológica y el aumento ingente de datos, así como las capacidades de analítica de dichos datos, hacen cada vez más sencilla la identificación o reidentificación de titulares de datos a partir de datos agregados o anonimizados, ensanchando en la práctica la vis expansiva del concepto de datos personales y, por ende, la aplicación del RGPD. El tratamiento de datos personales viene definido en el art. 4.2 RGPD como cualquier operación o conjunto de operaciones realizadas sobre datos personales o conjuntos de datos personales, incluyendo actividades como las siguientes: “recogida, registro, organización, estructuración, conservación, adaptación o modificación, extracción, consulta, utilización, comunicación por transmisión, difusión o cualquier otra forma de habilitación de acceso, cotejo, o interconexión, limitación, supresión o destrucción (de datos personales)”.

La publicación en línea de las sentencias sintoniza con la transparencia del poder judicial. A la hora de evaluar qué datos deben hacerse públicos es necesario conseguir un equilibrio entre el derecho a la protección de datos y el derecho a la publicidad de las resoluciones judiciales. En este punto, un informe de impacto relativo a la publicación de resoluciones judiciales debe prestar particular atención al juicio de proporcionalidad de las operaciones de tratamiento con respecto a su finalidad, así como a los riesgos que comporta para los derechos y libertades de los interesados (art. 35.7.b y c RGPD). Debe ser valorado si la finalidad de dar publicidad a las resoluciones judiciales justifica

<https://www.cijc.org/es/cuadernos/Sentencias/Dossier%20de%20sentencias%20constitucionales%20sobre%20anonimizaci%C3%B3n.pdf>. Según el Acuerdo el Tribunal Constitucional preservará de oficio el anonimato de los menores y personas que requieran un especial deber de tutela, de las víctimas de delitos de cuya difusión se deriven especiales perjuicios y de las personas que no estén constituidas en parte en el proceso constitucional. El Tribunal, en los demás casos, podrá excepcionar, de oficio o a instancia de parte, la exigencia constitucional de publicidad de sus resoluciones (art. 164 CE), en lo relativo a los datos de identidad y situación personal de las partes intervinientes en el proceso. A tal fin, si una parte estimase necesario que en un asunto sometido al conocimiento del Tribunal no se divulgue públicamente su identidad o situación personal, deberá solicitarlo en el momento de formular la demanda o en el de su personación, exponiendo los motivos de su petición.

F. Esteban de la Rosa, “Construyendo el marco legal europeo para el diseño y uso de herramientas de inteligencia artificial en el entorno de la justicia civil-administrativa”, en el libro F. Herrera Triguero, A. Peralta Gutiérrez y L.S. Torres López, Universidad de Granada, pp. 423-451.

la publicación de ciertos detalles del caso que podrían facilitar una reidentificación posterior de los implicados que pusiera en riesgo sus derechos y libertades.

En ciertas ocasiones, la relevancia del caso puede influir sobre el mencionado juicio de proporcionalidad dando como resultado una menor necesidad de preocuparse por ocultar detalles jurídicamente poco relevantes (ya conocidos por la prensa, por ejemplo), o por ocultar la identidad de los protagonistas, bien porque ésta sea ya conocida o porque su relevancia como personas públicas justifique la necesidad de que estos datos no se oculten. Este es el caso, en particular, cuando existe un interés público predominante que justifica la divulgación de esos datos (art. 17.3.d RGPD). Pongamos por ejemplo un algoritmo capaz de seudonimizar los datos de resoluciones judiciales hasta un nivel “n”. Tras pasar por el proceso de seudonimización, los datos personales sólo son identificables con cierta facilidad si el caso adquiere gran notoriedad social. Estos casos de gran notoriedad social no necesitan tanto cuidado en la ocultación de ciertos datos en la medida en que o bien ya se conocen, o su conocimiento es necesario o conveniente por cuestiones de interés público o histórico (art. 17.3.d RGPD). Este algoritmo, pues, habría alcanzado un nivel de seudonimización de datos óptimo desde el punto de vista de la normativa de protección de datos. Por tanto, en las operaciones de seudonimización llevadas a cabo por algoritmos de IA el sistema debe ser orientado hacia el logro de la mayor dificultad para la identificación de las personas dentro del contexto de relevancia social de cada caso si fuera necesario tener en cuenta este criterio en los términos que acabamos de expresar. Hay que tener en cuenta que los datos producidos por el poder judicial también se rigen por la normativa de la UE sobre datos abiertos y la reutilización de la información del sector público en todo lo que no afecta a datos personales. La disponibilidad de las decisiones judiciales en un formato legible, bajo metadatos, es clave para conseguir un sistema de justicia amigable con los algoritmos.

3. Se afirma que los sistemas de Machine Learning se proyectan sobre el tiempo pasado. Puesto que estos sistemas están basados en conjuntos relativamente fijos de datos del pasado, se puede producir una tendencia hacia la osificación del derecho lo que resulta incompatible con la creatividad que caracteriza a la actividad del juez. No obstante, hoy existe tecnología capaz de remediar este defecto. Con una gran cantidad de datos y distintas ponderaciones asignadas a cada uno, los algoritmos del aprendizaje automático podrían generar resultados creativos o nuevos. Además, mientras las máquinas sigan resolviendo casos difíciles, estos conjuntos de datos se actualizarían de manera habitual.²⁸ La utilización de datos sintéticos puede ser también un remedio para permitir la actualización de los datos de la máquina de IA.

4. La IA basada en datos plantea dificultades para escapar de los sesgos inherentes a los datos procesados. Los sistemas basados en datos pueden ser considerados tan buenos como los datos en que se basan y que el algoritmo procesa. Ello conlleva el riesgo de que ciertos algoritmos refuercen o consoliden prejuicios o situaciones de

²⁸ R. Susskind, *Tribunales online y la justicia del futuro*, Walter Kluwers, 2020, p. 333.

F. Esteban de la Rosa, "Construyendo el marco legal europeo para el diseño y uso de herramientas de inteligencia artificial en el entorno de la justicia civil-administrativa", en el libro F. Herrera Triguero, A. Peralta Gutiérrez y L.S. Torres López, Universidad de Granada, pp. 423-451.

marginación de grupos vulnerables.²⁹ Un ejemplo de este tipo de situaciones fue puesto de relieve por el relator especial sobre la extrema pobreza y los derechos humanos de Naciones Unidas en relación con el algoritmo SyRI, con el que el Gobierno holandés preseleccionaba a ciudadanos que con mayor probabilidad podrían cometer fraude a la Seguridad Social de acuerdo con ciertos parámetros automatizados. Los ciudadanos preseleccionados sí eran ya objeto de una investigación humana, por lo que no podía hablarse en este caso de un incumplimiento de las previsiones del art. 22 RGPD. Sin embargo, el caso fue llevado ante la Corte de Distrito de La Haya, que resolvió considerando al algoritmo como incompatible con el artículo 8 del Convenio Europeo de Derechos Humanos (derecho a la intimidad).³⁰ En el contexto de este procedimiento, el informe del relator especial Philp Alston afirmaba que el algoritmo poseía un efecto discriminatorio y estigmatizador debido, entre otras razones, a la autoridad que se le concedía.³¹

5. Aunque los algoritmos son capaces de procesar enormes cantidades de datos, es muy difícil llegar a operacionalizar todos los aspectos relevantes de una realidad social, es decir, convertirlos en variables medibles. En la operacionalización de los diferentes elementos que pueden considerarse relevantes las personas encargadas de la programación pueden obviar algunas variables, introducir sesgos en el sistema o escoger bases de datos que ofrecen una representación parcial de la realidad.³²

Uno de los defectos más importantes achacables a los algoritmos de Machine Learning es su escasa transparencia y la dificultad de responder al criterio de explicabilidad. Sin embargo, desde la doctrina se impugna la misma posibilidad de hacer exigible esta clase de garantía a los algoritmos de Machine Learning. Para Richard Susskind no resulta claro, ni técnica ni filosóficamente, qué se quiere decir cuando se propone que la ética debería estar integrada en el aprendizaje automático o Machine Learning. Tampoco está claro qué se quiere decir cuando se exige que los ingenieros de software programen sistemas de aprendizaje automático para que proporcionen explicaciones inteligentes. Este pensamiento, según el autor, no lograría entender la diferencia entre los procesos inductivos inherentes al aprendizaje automático y el tipo de argumento que esperamos

²⁹ Véase C. O'Neil, "Armas de destrucción matemática", Madrid, Capitán Swing, 2017.

³⁰ Véase la sentencia del Tribunal de Distrito de la Haya de 5 de febrero de 2020 (Caso SyRI). Véase también la sentencia del Tribunal Ordinario de Bolonia de 31 de diciembre de 2020 (Caso Deliveroo). El algoritmo de la plataforma Deliveroo discriminaba a ciertos trabajadores en cuanto a las condiciones de acceso a la plataforma, forzando de facto su disponibilidad ineludible a horas intempestivas. Sobre este caso véase J. A. Castillo Parrilla, "Sentencia del Tribunal Ordinario de Bolonia de 31 de diciembre de 2020 (Caso Deliveroo) ¿Discriminación algorítmica o discriminación a través de un algoritmo?", Revista Derecho Digital e Innovación, nº 7, 2020.

³¹ En detalle sobre el caso, G. Lazcoz Moratinos y J. A. Castillo Parrilla, "Valoración algorítmica ante los derechos humanos y el Reglamento General de Protección de Datos: el caso SyRI", Revista Chilena de Derecho y Tecnología, Vol. 9, núm. 1 (2020), en: <https://rchdt.uchile.cl/index.php/RCHDT/article/view/56843>).

³² Véase A. Soriano Arranz, « Decisiones automatizadas : problemas y soluciones jurídicas. Más allá de la protección de datos », Revista de Derecho público, Marcial Pons, doi: https://doi.org/10.37417/RPD/vol_3_2021_5352021.

F. Esteban de la Rosa, “Construyendo el marco legal europeo para el diseño y uso de herramientas de inteligencia artificial en el entorno de la justicia civil-administrativa”, en el libro F. Herrera Triguero, A. Peralta Gutiérrez y L.S. Torres López, Universidad de Granada, pp. 423-451.

cuando pedimos una explicación.³³ En la doctrina existen voces a favor de hacer evolucionar el control en estos casos hacia otra clase de estrategias.³⁴

III. Impacto del marco legal sobre las diferentes herramientas de IA y sus diferentes usos en el ámbito legal y en el entorno de la justicia civil-administrativa

1. La IA en los servicios legales

La utilización de la IA ha comenzado en el ámbito de los servicios legales, siendo utilizado principalmente por despachos profesionales y compañías de seguros. Tomando en consideración las posibilidades que hoy en día plantea la IA José Ignacio Solar Cayón ha distinguido hasta siete sectores de la actividad jurídica en los que se están aplicando herramientas de IA, algunos específicos para el ámbito de los profesionales del Derecho al margen del sistema judicial. Entre ellos se encuentran la creación de sistemas expertos para la realización de tareas de *compliance*; los sistemas de lectura automática y análisis de contratos; los sistemas de recuperación de la información para la investigación jurídica así como las plataformas de búsqueda y análisis de la información jurídica; los sistemas de “búsqueda de respuestas” o *question answering*; los sistemas de análisis y predicción de las decisiones judiciales; los sistemas de elaboración automática de documentos jurídicos mediante aplicativos web; los sistemas de codificación predictiva para la realización de tareas de e-discovery; y la utilización de interfaces para obtener la solución automática en línea de una disputa.³⁵

Debido a la regulación sobre la intervención preceptiva de abogado, en algunos de los casos de uso que se describen a continuación la IA no puede sustituir a la intervención de un abogado, pero sí que sirve como una herramienta que estos pueden utilizar en su ejercicio. Las herramientas de IA que se describen pueden contribuir a la transformación de la actividad que llevan a cabo los abogados así como a la reestructuración de sus recursos humanos.³⁶ En ciertos casos la IA puede llegar a sustituir a la actuación de los profesionales del Derecho.

Los algoritmos de aprendizaje automático pueden revisar gran cantidad de información electrónica e incluso seleccionar la más relevante para un litigio.³⁷ La utilización de

³³ Richard E. Susskind, *Online courts and the future of justice*. Oxford University Press, 2019.

³⁴ J.I. Solar Cayón, “Inteligencia artificial en la justicia penal: los sistemas algorítmicos de evaluación de riesgos”, en José Ignacio Solar (ed.), *Dimensiones éticas y jurídicas de la Inteligencia Artificial en el marco del Estado de Derecho*. Madrid, Universidad de Alcalá de Henares, 2020, p. 153.

³⁵ Véase J.I. Solar Cayón, “La Inteligencia artificial jurídica : el impacto de la innovación tecnológica en la práctica del Derecho y el mercado de los servicios jurídicos”, Cizur Menor Navarra : Aranzadi Thomson Reuters, 2019.

³⁶ Sobre esta transformación véase A. Janssen y T.J. Venmanns, “The Effects of Technology on Legal Practice: From Punch Card to Artificial Intelligence?”, en L.A. Dimatteo, A. Janseen, P. Ortolani, F. de Elizalde y M. Cannarsa, *The Cambridge Handbook of Lawyering in the Digital Age*, Cambridge University Press, 2021, pp. 38-56.

³⁷ J.I. Solar Cayón, (2018). “La codificación predictiva: inteligencia artificial en la averiguación procesal de los hechos relevantes”, *Anuario Facultad de Derecho - Universidad de Alcalá XI* , pp. 75-105; p. 75 para la cita.

F. Esteban de la Rosa, “Construyendo el marco legal europeo para el diseño y uso de herramientas de inteligencia artificial en el entorno de la justicia civil-administrativa”, en el libro F. Herrera Triguero, A. Peralta Gutiérrez y L.S. Torres López, Universidad de Granada, pp. 423-451.

sistemas conocidos indistintamente como TAR (*Technology-Assisted Revision*) o *Predictive coding* permite clasificar los documentos según su relevancia para el caso y dirigir la revisión hacia los considerados pertinentes. Entre los productos de predicción se encuentran Lex Machina, Ravel Law, LexPredict, Premonition, Ogayar Co, y PROMETEA. En España destacan Jurimetría, vLex Analytics y Tirant Analytics. La utilidad de esta clase de herramientas para los profesionales es clara pues les permite calcular sus opciones de éxito antes incluso de hacerse cargo de un caso.

La IA es utilizada por programas como Iurisfy con el propósito de servir a la obtención de acuerdos de separación o divorcio,³⁸ y puede prestar ayuda para analizar diferentes tipos de contratos extrayendo información que ayuda en la gestión de contratos y de riesgos. Algunos de los productos más conocidos son Catalyst, eBrevia, iManage Ravn, Kira Systems, Lawgeex, Legal Robot o Leverton, ThoughtRiver, Dail o Luminance. La IA presta también servicio en el sector del compliance, ayudando a la identificación de incumplimientos normativos mediante la utilización de sistemas expertos, procesamiento natural del lenguaje y aprendizaje automático. Algunos de las empresas que ofrecen productos con esta finalidad son Neota Logic, Oracle e Intraspection, Compliance HR, Microsoft 365 Compliance, Foley & Lardner LLP o Global Risk Clinic.

Los Marketplaces jurídicos y redes de abogados utilizan la IA para extraer datos sobre abogados, jueces, litigantes y de las áreas del Derecho procedentes del análisis de millones de resoluciones judiciales. Su objetivo es ayudar en la contratación de un abogado para el diseño de estrategias. Entre los productos con esta finalidad se encuentran Lexdir o elAbogado. El uso de la IA en este caso cuenta con la dificultad de llevar a cabo un cumplimiento escrupuloso de las normas de protección de los datos de los que se nutre la máquina dada la dificultad de contar con el consentimiento de todos los involucrados.

Uno de los sectores que más desarrollo ha tenido es el de la automatización de reclamaciones. Estos programas suelen ser utilizados por empresas de abogados que, una vez determinado que el particular tiene derecho a la estimación de su reclamación, se ocupan de plantear de forma masiva los casos ante las instancias (administrativa o judicial) de reclamación. La finalidad de la máquina consiste en detectar, a través de los datos introducidos, si el usuario tiene o no derecho a que su reclamación sea estimada. La interacción con la aplicación es por escrito y oral a través de un chatbot.

Entre los productos de este tipo se encuentran DoNotPay, que es ofrecido para ayudar a los particulares a impugnar las multas de tráfico en el Reino Unido y en algunos estados de EEUU y Canadá. Este mismo programa ha sido utilizado para promover demandas en tribunales de menor cuantía a empresas como Uber, AT&T, Verizon, United Airlines y Frontier Communications. Quarande es una herramienta orientada hacia la reclamación de despidos. Para utilizarla el usuario debe descargar la app y registrarse. Después puede subir la documentación requerida y quedará informado en todo momento por el abogado asignado al caso. Si pierde, todos los gastos generados

38

Véase <https://www.derechopractico.es/guialegaltech/iurisfy/#:~:text=Iurisfy%20es%20una%20app%20que,de%20Iurisfy%20por%20380%20%E2%82%AC.>

F. Esteban de la Rosa, “Construyendo el marco legal europeo para el diseño y uso de herramientas de inteligencia artificial en el entorno de la justicia civil-administrativa”, en el libro F. Herrera Triguero, A. Peralta Gutiérrez y L.S. Torres López, Universidad de Granada, pp. 423-451.

los asume Quarande. Y si gana los honorarios son un 15% (mínimo de 250 €). Reclamación de Vuelos es una herramienta para resolver reclamaciones por incidencias en vuelos (cancelación, retraso, overbooking o equipaje). Su bot Harvei (Herramienta Asistente de Reclamaciones Virtual E Inteligente) analiza la documentación presentada online, decide si el asunto es reclamable y calcula la indemnización que procede.

Respecto del marco legal, habida cuenta que las máquinas a que se refiere esta sección no están orientadas a asistir al juez en la interpretación o aplicación de la ley a un caso concreto las herramientas indicadas pueden ser consideradas como de no alto riesgo de acuerdo con la clasificación de sistemas de IA al que se refiere la propuesta de Reglamento. Por otra parte, las aplicaciones utilizadas por los profesionales deben atemperar su funcionamiento a los principios establecidos por el artículo 5 RGPD que son los de licitud, lealtad y transparencia; limitación de la finalidad; minimización de los datos; exactitud; limitación del plazo de conservación; integridad y confidencialidad; y responsabilidad proactiva, siendo de la mayor importancia que el responsable del tratamiento acredite el cumplimiento de estos (art. 5.2 RGPD). Es menos evidente la necesidad de tener en cuenta el artículo 22 RGPD habida cuenta que las máquinas o bien no aportan una solución automática o si lo hacen no produce efectos significativos o bien el interesado ha podido dar su consentimiento. En todos los casos estamos en presencia de sistemas de IA de no alto riesgo por lo que su régimen de gobernanza no estaría sujeta más al marco voluntario de los códigos de conducta que puedan ser creados en el ámbito empresarial.

Cabe por el contrario afirmar que los chatbots o asistentes conversacionales, para ser respetuosos con el marco de protección de datos, no deberían guardar información personal más allá del tiempo necesario para tramitar la reclamación del sujeto (principios de minimización y limitación temporal del tratamiento). No parece que la actividad exija una conservación prolongada en el tiempo ni admita usos secundarios posteriores compatibles; pero si se considerase que sí deberán advertirse de manera clara, destacada y sencilla a los usuarios y justificarse adecuadamente a efectos de cumplir con el principio de responsabilidad proactiva. Como los datos van a ser tratados de forma automatizada se debería informar al sujeto de que se está tratando con un sistema automático (principio de transparencia).

La creciente capacidad computacional de los sistemas de IA hace que cada vez sea más común la posibilidad de elaborar perfiles. Algunos sistemas de IA del mercado ofrecen información sobre las trayectorias de jueces y magistrados, pudiendo incluso indicar su inclinación hacia la resolución de un caso en un determinado sentido. La elaboración de cuantificaciones sobre datos relativos a las decisiones de jueces y magistrados concretos afecta a sus derechos a la intimidad y vida privada y a la protección de datos de carácter personal.³⁹ Esta consideración ha sido llevada al

³⁹ Artículos 7 y 8, respectivamente, de la Carta de los Derechos Fundamentales de la Unión Europea; artículo 8 del Convenio Europeo de Derechos Humanos; artículo 18.1 de la Constitución española, que se refiere específicamente al derecho al honor, a la intimidad personal y familiar y a la propia imagen; apartado V de la Carta de Derechos Digitales de 2021, Carta Ética Europea

F. Esteban de la Rosa, “Construyendo el marco legal europeo para el diseño y uso de herramientas de inteligencia artificial en el entorno de la justicia civil-administrativa”, en el libro F. Herrera Triguero, A. Peralta Gutiérrez y L.S. Torres López, Universidad de Granada, pp. 423-451.

extremo en Francia donde ha sido tipificado penalmente el uso de máquinas de IA para crear perfiles y predecir el comportamiento de los jueces.⁴⁰

Debe tenerse en cuenta que los perfiles personales son también datos personales, en la medida en que son “información sobre una persona física identificada o identificable” (art. 4.1 RGPD). El concepto de dato personal es considerablemente amplio, y abarca cualquier información, ya sea objetiva o subjetiva, verdadera o errónea, mientras que pueda referirse a una persona identificada o identificable. En el amplio abanico que establece esta definición deben entenderse comprendidos los juicios o evaluaciones sobre una persona, sean éstos verdaderos o erróneos, o fruto de un prejuicio personal o de un análisis masivo de datos.⁴¹ A los perfiles personales les resulta plenamente aplicable el RGPD, con todos sus requisitos y consecuencias.⁴² Por esta razón es conveniente excluir de la materia del análisis predictivo la trayectoria de jueces y magistrados concretos, tanto en lo que se refiere a datos en bruto como a finalidades del tratamiento de análisis, ya que, además de afectar al derecho fundamental a la protección de datos personales, pueden crear presiones y afectar a la independencia judicial. Esta información puede ser sustituida por otra equivalente y menos invasiva como una distribución de resultados (favorables o desfavorables, por ejemplo, a ciertas pretensiones) del órgano judicial (juzgados X a Z; Sección X de la Audiencia Provincial Y).

2. La IA en la administración de justicia

del Uso de IA en sistemas judiciales CEPEJ y Propuesta de Reglamento sobre Inteligencia Artificial.

⁴⁰ Así se dispone en el artículo 33 de la Ley nº 2019-222 de 23 de marzo de 2019 de programación 2018-2022 y de reforma de la justicia, recogido en la Sección 3, capítulo segundo del Título II, sección que aparece bajo el epígrafe Conciliar la publicidad de las decisiones de justicia y el derecho al respecto a la vida privada. De manera concreta, este artículo modifica una serie de textos legales, entre ellos, el código de justicia administrativa, introduciendo en su artículo L-10 la siguiente previsión: “Los datos de identidad de los magistrados y miembros del Registro no pueden ser reutilizados con el propósito o el efecto de evaluar, analizar, comparar o predecir sus prácticas profesionales reales o presuntas. La violación de esta prohibición se castiga con las sanciones previstas en los artículos 226-18, 226-24 y 226-31 del Código Penal, sin perjuicio de las medidas y sanciones previstas en la Ley Nº 78-17, de 6 de enero de 1978, en relación con el procesamiento de datos, archivos y libertades” Información disponible en https://www.legifrance.gouv.fr/jorf/article_jo/JORFARTI000038261761?r=SqlReSoaHM

⁴¹ Véase el Grupo de Trabajo del Artículo 29, Dictamen 4/2007 sobre «el concepto de datos personales», WP136, de 20 de junio de 2007, p. 6. Disponible en: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_es.pdf. También sobre el concepto de dato personal y su amplitud, C. M. Romeo Casabona, «Datos personales (Comentario al artículo 4.1 RGPD)», en A. Troncoso Reigada, *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, pp. 573-589.

⁴² Acerca de las consecuencias de aplicar el RGPD a los perfiles personales su importancia económica, J. A. Castillo Parrilla, “Derecho al patrimonio digital. Bienes digitales y datos como bienes”, en A. Troncoso Reigada, *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo II, pp. 4346-4368.

F. Esteban de la Rosa, “Construyendo el marco legal europeo para el diseño y uso de herramientas de inteligencia artificial en el entorno de la justicia civil-administrativa”, en el libro F. Herrera Triguero, A. Peralta Gutiérrez y L.S. Torres López, Universidad de Granada, pp. 423-451.

Según un estudio reciente de la Comisión Europea⁴³ las autoridades judiciales de los Estados miembros están utilizando cada vez más aplicaciones basadas en la IA. En España el pasado 16 de diciembre de 2021 el CGPJ daba la noticia de la puesta a disposición, a favor de los miembros de la Carrera Judicial, de un conjunto de aplicaciones informáticas basadas en la IA y el Machine Learning que permiten la seudonimización automática de un documento y el acceso automático a la legislación y jurisprudencia citadas en las resoluciones judiciales.⁴⁴ A esta clase de herramientas cabe sumar otras, de interés sobre todo para las personas sobre quienes recae la responsabilidad de tramitar un proceso judicial, como los programas de conversión y transcripción de audio a texto por reconocimiento de voz o los de traducción automática.

De un número importante de las funciones concretas a que pueden servir las herramientas de IA en el entorno de la justicia da cuenta el informe que acompaña a la Carta Ética Europea sobre el uso de la Inteligencia Artificial en sistemas judiciales y su entorno.⁴⁵ Entre las finalidades descritas por el informe se encuentran el logro de una mejora de los motores de búsqueda de jurisprudencia mediante el uso de Machine Learning, no limitado a las palabras clave; el logro de una mejora en el acceso a la legislación mediante el uso de chatbots; la generación automática de documentos, la mejora de la eficiencia en la decisión sobre la inversión en recursos humanos y materiales en la administración de justicia; el ofrecimiento de una solución sin intervención humana (decisión automática) permitiendo a las partes introducir los datos de un caso mediante una interfaz online (ODR); el apoyo en el diseño de baremos de utilidad para ciertos litigios en el ámbito civil; las herramientas de predicción que ofrecen una estimación sobre las posibilidades de éxito de un litigio; o la creación de perfiles relativos a la actividad de los jueces de la que hemos hablado en el apartado anterior.⁴⁶

Junto a estas funciones en algunos tribunales la utilización de la IA está posibilitando que los tribunales presten servicios a los usuarios de la justicia que van más allá de las tradicionales de juzgar y hacer y ejecutar lo juzgado.⁴⁷ Las experiencias más innovadoras en los sistemas judiciales dan cuenta de la utilización de herramientas de IA de forma combinada con herramientas TIC en los llamados tribunales online que permiten el inicio y seguimiento en línea de un proceso.⁴⁸ La creación de esta clase de

⁴³ Tal y como ilustran el documento de trabajo de los servicios de la Comisión que acompaña a la Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las regiones, La digitalización de la justicia en la UE: Un abanico de oportunidades y el *Study on the use of innovative technologies in the justice field* (Estudio sobre el uso de tecnologías innovadoras en el ámbito de la justicia).

<https://op.europa.eu/en/publication-detail/-/publication/4fb8e194-f634-11ea-991b-01aa75ed71a1/language-en>

⁴⁴ La noticia puede verse en el siguiente enlace: <https://tinyurl.com/yc7vkmxx> (fecha de consulta: 7/07/2022).

⁴⁵ *Cit.*

⁴⁶ Véase la European Ethical Charter on the use of Artificial Intelligence in judicial systems and their environment, pp. 16-17.

⁴⁷ Véase F. Esteban de la Rosa (director), *Guía para el diálogo sobre el diseño y uso eficiente, de calidad y ético de herramientas tecnológicas en la justicia civil*, Fundación Cotec para la Innovación, 2022.

⁴⁸ Vid. R. Susskind, *Tribunales online y la justicia del futuro*, Wolters Kluwer, 2019, p. 27.

F. Esteban de la Rosa, “Construyendo el marco legal europeo para el diseño y uso de herramientas de inteligencia artificial en el entorno de la justicia civil-administrativa”, en el libro F. Herrera Triguero, A. Peralta Gutiérrez y L.S. Torres López, Universidad de Granada, pp. 423-451.

tribunales responde a las necesidades de los litigantes sin representación. Hay que tener en cuenta que en la mayoría de los países anglosajones la intervención del abogado no es preceptiva y poner a disposición de los usuarios esta clase de máquinas puede ahorrarle el coste de la contratación de un abogado. Los tribunales online pretenden proporcionar a los usuarios de la justicia un servicio integral de información legal y de resolución de litigios con un diseño centrado en las necesidades de los usuarios (User Centric Judicial System) y han proliferado en numerosas jurisdicciones tanto en Europa (Reino Unido) como en Estados Unidos (Connecticut, Florida, Hawaii, Iowa, Kansas, Michigan, Nevada, New Hampshire, New Mexico, Nueva York, Utah o Texas).

Cada vez más jurisdicciones están dando el paso adelante de proponer un uso combinado de herramientas TIC y de IA diseñadas para proporcionar información de manera que las partes puedan obtener un diagnóstico o una predicción del resultado del proceso para afrontar en mejor posición una negociación directa realizada en línea. Cabe mencionar el Dutch platform Rechtwijzer en Países Bajos, el CRT en la Columbia Británica de Canadá, los tres tribunales de Internet creados en China y la experiencia en Singapur.⁴⁹ Esta información se ofrece con la finalidad de que las partes puedan llegar a un acuerdo con la información ofrecida por la máquina. Respecto del CRT los informes anuales de actividad revelan la relación entre las consultas de diagnóstico de casos y el número de demandas presentadas.⁵⁰ Con esta orientación se alinea el proyecto en Singapur.⁵¹ Lo que caracteriza a estos tribunales online es que ofrecen a los usuarios de la justicia herramientas de información así como de ayuda para la toma de decisiones, herramientas de negociación automatizada o programas informáticos para calcular la compensación utilizando la teoría de los juegos o con base en programas de utilizan IA, al igual que lo hacen los profesionales. También pueden ofrecer programas de ayuda para la redacción de acuerdos.

⁴⁹ Véase F. Esteban de la Rosa y J. Zeleznikow, “Making Intelligent Online Dispute Resolution Tools available to Self-Represented Litigants in the Public Justice System: Towards an Ethical use of the AI technology in the administration of Justice”, en *Eighteenth International Conferente for Artificial Intelligence and Law* (ICAIL’21), June 21-25, 2021. <https://doi.org/10.1145/3462757.3466077>

⁵⁰ Los informes anuales de actividad del CRT se encuentran disponibles en este enlace: <https://civilresolutionbc.ca/about-the-crt/presentations/>

⁵¹ Fernando Esteban de la Rosa and John Zeleznikow. 2021. Making Intelligent Online Dispute Resolution Tools available to Self-Represented Litigants in the Public Justice System: Towards and Ethical use of the AI technology in the administration of Justice. In *Eighteenth International Conference for Artificial Intelligence and Law* (ICAIL’21), June 21–25, 2021, São Paulo, Brazil. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3462757.3466077>

F. Esteban de la Rosa, "Construyendo el marco legal europeo para el diseño y uso de herramientas de inteligencia artificial en el entorno de la justicia civil-administrativa", en el libro F. Herrera Triguero, A. Peralta Gutiérrez y L.S. Torres López, Universidad de Granada, pp. 423-451.

Experimentos sobre un sistema predictivo se han llevado a cabo en Francia,⁵² Países Bajos⁵³ e incluso en el Tribunal Europeo de los Derechos Humanos.⁵⁴ El ofrecimiento por el sistema público de justicia de herramientas de predicción o, mejor dicho, de indicación de probabilidades sobre la posibilidad de ganar un litigio presenta luces y sombras.⁵⁵ En un trabajo publicado en 1949, Loevinger sostenía que el estadio actual (positivista) de la ciencia y sus métodos empíricos y matemáticos, que permiten asignar valores numéricos a los hechos observados y otros datos y realizar determinados cálculos con ellos (especialmente a través de la rama de las matemáticas conocida como estadística) llevaría, si se aplicara al Derecho, al paso a una nueva era en la ciencia jurídica. Se trataba del paso de la Jurisprudencia (Jurisprudence) a la Jurimetría (Jurimetrics). En defensa de la Jurimetría como ciencia Loevinger sostenía que ayuda a garantizar principios básicos de la democracia, como el derecho de la ciudadanía a ser informada acerca de los asuntos que requieren su control y que gestionan las autoridades que ella elige. Tal derecho no puede sino descansar en una información fiable que solo puede ser la que se basa, a su vez, en una investigación científica libre y competitiva.⁵⁶ La puesta a disposición de los usuarios de la justicia de esta información podría ser considerado como una contribución a la satisfacción de los principios propios de una organización política democrática, especialmente el principio de igualdad vinculado al de seguridad jurídica (artículo 14 de la CE en relación con el 9.3).

No obstante, ofrecer esta información también tiene riesgos. La información suministrada, ofrecida por el sistema público de justicia, puede afectar de forma negativa al acceso a la justicia en la medida en que pueda conducir a una renuncia a la vía jurisdiccional, la cual podría ser además contraria a, o elusiva de, la normativa jurídica sobre resolución de litigios. Otro riesgo adicional es el de datificación. Se trata de un riesgo que ha sido descrito por Galindo Ayuda, siguiendo el estudio de Re & Solow-Niederman sobre la aplicación de IA al Derecho.⁵⁷ De manera concreta, la datificación

⁵² Los Tribunales de Apelación de Douai y Rennes realizaron un ensayo de tres meses en 2017 con un programa informático calificado de "predictivo" por un grupo de jueces. Véase CEPEJ, Apéndice I, para un estudio en profundidad sobre el uso de la IA en los sistemas judiciales, especialmente las aplicaciones de IA que procesan decisiones y datos judiciales.

⁵³ La Autoridad Judicial de los Países Bajos (Rechtspraak) llevó a cabo un experimento en septiembre de 2017 en el que se presentaron dos casos reales para que el sistema los evaluara y pusiera a prueba la capacidad de su programa de inteligencia artificial llamado LexIQ para proponer soluciones a esos casos judiciales reales. Las decisiones tomadas por LexIQ se contrastaron con las tomadas por jueces humanos. Véase <https://towardsdatascience.com/legal-certainty-and-the-possibility-of-computer-decision-making-in-the-courtroom-ac4b1a6c42d1> (fecha de consulta: 11 de febrero de 2021).

⁵⁴ Se anuncia que el modelo construido puede predecir las decisiones del tribunal con una gran precisión (79% de media). Véase Nikolaos Aletras, Dimitrios Tsarapatsanis, Daniel Preotjiuc-Pietro y Vasileios Lampos, "Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective", PeerJ Comput. Sci. 2:e93; DOI 10.7717/peerj-cs.93.

⁵⁵ C. Liccopes y L. Dumolin (2019), "Le travail des juges et les algorithmes de traitement de la jurisprudence. Premières analyses d'une experimentation de « justice predictive » en France", Droit et Société, 2019/3 N° 103, pp. 535-554.

⁵⁶ L. Loevinger, "Jurimetrics. The Next Step Forward", op. cit., p. 493.

⁵⁷ R.M. RE y A. Solow-Niederman, (2019). "Developing Artificially Intelligent Justice", Stanford Technology Law Review, Stanford, v. 22, n.1, p. 242-289, Septiembre; pp. 252-254 para la cita.

F. Esteban de la Rosa, “Construyendo el marco legal europeo para el diseño y uso de herramientas de inteligencia artificial en el entorno de la justicia civil-administrativa”, en el libro F. Herrera Triguero, A. Peralta Gutiérrez y L.S. Torres López, Universidad de Granada, pp. 423-451.

implicaría un predominio en la resolución de litigios de las reglas de cálculo matemático sobre las reglas jurídicas, cuya legitimidad descansa en su origen democrático; pero este es un tipo de legitimidad de la que carecen las reglas matemáticas, aunque tengan otras propiedades notables.

La perspectiva del uso de las herramientas de IA resulta distinta cuando su objetivo es ayudar al juez ofreciéndole un borrador o proyecto de sentencia o, en su caso, dictar una decisión que pueda llegar a sustituir al juez. Un proyecto en la primera dirección ha sido presentado recientemente en Alemania en mayo de 2022, en concreto en el distrito de Frankfurt. El proyecto Frauke comporta la utilización de IA para ayudar a los jueces a resolver litigios en un ámbito muy concreto, a saber, el de las reclamaciones derivadas por cancelaciones o retrasos derivados del transporte aéreo de pasajeros.⁵⁸ El dato que se ha tenido en cuenta para impulsar este proyecto ha sido el elevadísimo número de procedimientos (entre 10.000 y 15.000) que han de ser decididos por el Tribunal de Distrito de Frankfurt cada año, lo que genera un gran volumen de trabajo asociado exclusivamente a esta clase de procedimientos. La denominación Frauke proviene de las siglas “Frankfurt Judgment Configurator Electronic”. En el lanzamiento del proyecto piloto participan el tribunal de distrito de Frankfurt, la compañía IBM y la Oficina Central de protección de datos del Estado de Hesse. La decisión de la máquina no pretende sustituir a la del juez. La IA es utilizada en este caso para poner en manos del juez propuestas de decisión que no son vinculantes para él. Para elaborar la decisión la máquina analiza, entre otros aspectos, las tarjetas de embarque, los tiempos de vuelo, los datos meteorológicos así como las decisiones anteriores sobre casos del mismo tipo dictadas por el tribunal de distrito. El empleo de la IA en este caso se ve favorecido por el hecho de que muchas reclamaciones comparten el hecho generador de la indemnización, al verse afectados todos los pasajeros del vuelo que ha sido cancelado, y el hecho de que las tasas de compensación estén reguladas a un tanto alzado por el Reglamento de la UE sobre los derechos de los pasajeros aéreos.

Por tanto, la máquina queda configurada como un asistente del juez por lo que con claridad debe ser entendida como un sistema de IA de alto riesgo. En nuestra opinión el empleo de máquinas para ofrecer probabilidades de éxito de la demanda a los usuarios de la justicia debería también ser consideradas como herramientas de alto riesgo a pesar de que, en puridad, no estamos en presencia de una asistencia al juez a la hora de interpretar o aplicar la ley a unos hechos concretos. De acuerdo con la letra de la propuesta de Reglamento esta clase de máquinas no entrarían entre los sistemas de alto riesgo.

IV. El concepto de intervención humana significativa

cita tomada de F. Galindo (2019). “Inteligencia Artificial y acceso a documentación jurídica: sobre el uso de las TICs en la práctica jurídica”, Revista Democracia Digital e Governo Eletrônico, Florianópolis, v. 1, n. 18, p. 144-166. Citado por F. Galindo Ayuda (2020), “Avances en sistemas jurídicos de recuperación de documentos”. Sciere, 26: 1(en-jun 2020), pp. 63-74; pp. 69 y 70 en relación a este asunto.

⁵⁸ La noticia puede verse en <https://www.hessenschau.de/panorama/amtsgericht-frankfurt-kuenstliche-intelligenz-hilft-bei-massen-urteilen,amtsgericht-roboter-100.html>.

F. Esteban de la Rosa, "Construyendo el marco legal europeo para el diseño y uso de herramientas de inteligencia artificial en el entorno de la justicia civil-administrativa", en el libro F. Herrera Triguero, A. Peralta Gutiérrez y L.S. Torres López, Universidad de Granada, pp. 423-451.

El juez-robot proyectado en Estonia plantea el debate sobre la admisibilidad de una máquina que pueda dictar decisiones con la fuerza y autoridad de una sentencia pero sin contar con la motivación que caracteriza a estas decisiones. Desde la doctrina se han ofrecido argumentos a favor de esta posibilidad.⁵⁹ El hecho de que pueda haber predicciones que, una vez conocido el resultado de la decisión judicial, lleguen a arrojar un porcentaje de acierto superior al 95% y de que existan países donde se acumulan retrasos judiciales de más de 100 millones de casos, como en Brasil, están llevando a alimentar un argumento de utilidad dirigido a permitir la conversión de las predicciones en determinaciones sin necesidad de que vengan avaladas por una motivación. Este último tipo de algoritmos es el que, en la actualidad, genera mayores problemas para el Derecho y el que ha despertado el interés de una parte de la doctrina jurídica por los riesgos que entraña para los valores y principios sobre los que se asientan los Estados democráticos y los derechos fundamentales de la ciudadanía.

El principio de control humano de la decisión de IA goza de arraigo en Europa.⁶⁰ El principio quinto de la Carta Ética del Consejo de Europea lleva *por título* "bajo control del usuario" con el propósito de excluir un enfoque prescriptivo y garantizar que los usuarios sean actores informados. El respeto a este principio es especialmente importante en los sistemas de inteligencia artificial de alto riesgo.⁶¹ Como contrapunto a esta directriz cabe destacar la innovación que parecen introducir las Directrices del Comité de ministros del Consejo de Europa de 16 de junio de 2021. El apartado 20 de las mencionadas Directrices deja la cuestión en manos de los Estados del Consejo de Europa. A su parecer, si el Derecho nacional permite las decisiones automáticas tales decisiones deberían quedar sometidas a revisión ante un juez.

Para los países miembros de la UE la obligación de garantizar la intervención humana, lo que conlleva la exclusión de las decisiones automatizadas, se encuentra en el artículo

⁵⁹ Véase R. Susskind, *Tribunales online y la justicia del futuro*, Walter kluwers, 2020, p. 329.

⁶⁰ Véase el apartado 4.3 del Dictamen del Comité Económico y Social Europeo sobre «La revolución digital teniendo en cuenta las necesidades y los derechos de los ciudadanos» Doc 2019/C 190/03; la Resolución del Parlamento Europeo 2019-2024 TEXTOS APROBADOS P9_TA (2020)0275 Marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas Resolución del Parlamento Europeo, de 20 de octubre de 2020; apartado 1.4 del Dictamen del Comité Económico y Social Europeo. TEN/679. La revolución digital teniendo en cuenta las necesidades y los derechos de los ciudadanos: El CESE tiene una opinión clara sobre la cuestión relativa al grado en que es éticamente aceptable delegar las decisiones (con implicaciones morales) a sistemas basados en la inteligencia artificial (IA): los sistemas automatizados, independientemente de su complejidad, tienen que funcionar de acuerdo con el principio de control humano. Solo los seres humanos toman la decisión final y asumen la responsabilidad por ella. Entre las exigencias de la propuesta de Reglamento en materia de Inteligencia Artificial se incluye como requisito obligatorio la supervisión o vigilancia humana, que figura como uno de los siete requisitos para el desarrollo de una IA fiable, en consonancia con las recomendaciones sobre aplicaciones de alto riesgo en el Libro Blanco sobre IA. Véase G. Lazcoz Moratinos, 2020, "Análisis de la propuesta de Reglamento sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas", *IUS ET SCIENTIA*, 6(2), pp. 32-35.

⁶¹ Véase A. Obregón Fernández y G. Lazcoz Moratinos "La supervisión humana de los sistemas de inteligencia artificial de alto riesgo. Aportaciones desde el Derecho Internacional Humanitario y el Derecho de la Unión Europea", *Revista electrónica de estudios internacionales*, nº 42, 2021, p. 3.

F. Esteban de la Rosa, “Construyendo el marco legal europeo para el diseño y uso de herramientas de inteligencia artificial en el entorno de la justicia civil-administrativa”, en el libro F. Herrera Triguero, A. Peralta Gutiérrez y L.S. Torres López, Universidad de Granada, pp. 423-451.

22 RGPD. De acuerdo con este precepto, “todo interesado tendrá derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado (...) que produzca efectos jurídicos en él o le afecte significativamente de modo similar” (art. 22.1 RGPD). Esta prohibición no resulta aplicable cuando (1) la decisión, aunque plenamente automatizada, no comporta efectos jurídicos o análogos de importancia o (2) cuando no es plenamente automatizada en la medida en que se incluye una decisión humana en la última fase de la toma de decisiones, aunque comporte efectos jurídicos o análogos de importancia.

Para evaluar las herramientas de IA teniendo en cuenta la perspectiva de su acomodación al cumplimiento de la prohibición que contiene el artículo 22 RGPD se tienen en cuenta los siguientes criterios:

- si los resultados de la actividad o herramienta producen efectos jurídicos o significativos en el titular de los datos;
- si, existiendo estos efectos, se trata de una sugerencia (sujeta a posterior revisión humana) o una decisión automatizada autónoma;
- si, tratándose de una sugerencia, la persona que revisa o tiene la última palabra, puede apartarse libremente del criterio de la máquina sin sufrir penalización o una mayor fiscalización por ello;
- si, en caso de cumplirse los presupuestos anteriores, pueden implementarse medidas o correctivos que permitan mantener la herramienta cumpliendo con los principios de protección de datos personales.

De acuerdo con estos criterios es posible hacer una evaluación de las diferentes herramientas de IA.

Por ejemplo, la facilitación de la tarea de presentar una demanda mediante el uso de un Sistema Experto que reclama la información necesaria e impide la presentación si faltan datos se puede considerar como una herramienta que facilita meras sugerencias que las partes eligen. Las herramientas de generación automática de documentos no conllevan tampoco un efecto jurídico significativo y atribuible al tratamiento automatizado. Las herramientas de negociación asistida o automatizada se pueden considerar como de mera asistencia a las partes, o a quien debe facilitar su acuerdo por lo que escapan al impacto del artículo 22 RGPD.

Mayores problemas presenta la asistencia para la argumentación del fallo del juez. En estos casos es necesario que se acredite la necesaria independencia de facto del juez a la hora de seguir o separarse de las argumentaciones generadas con IA. En cuanto a las herramientas de asistencia al juez en la valoración de la prueba también es necesario que se acredite la necesaria independencia de facto del juez a la hora de seguir o separarse de las argumentaciones generadas con IA.

En cuanto a las plataformas que resuelven litigios de forma automática hay que decir que no cumplen con la normativa de protección de datos. Una resolución comporta evidentes y directos efectos jurídicos significativos, por lo que el proceso no puede automatizarse plenamente. La plataforma Cognicor Technologies es capaz de producir

F. Esteban de la Rosa, “Construyendo el marco legal europeo para el diseño y uso de herramientas de inteligencia artificial en el entorno de la justicia civil-administrativa”, en el libro F. Herrera Triguero, A. Peralta Gutiérrez y L.S. Torres López, Universidad de Granada, pp. 423-451.

una resolución automática sujeta a posible revisión por un agente previa solicitud de las partes. Esta intervención humana en “fase de revisión” salva la automatización plena. No obstante, sería deseable que una persona pudiese supervisar el resultado antes de dárselo a las partes con la suficiente libertad de apartarse del criterio automatizado.

La asistencia al juez en la valoración de la prueba es una actividad que se puede considerar como peligrosa no sólo en términos del art. 22 RGPD sino por favorecer o consolidar criterios discriminatorios, tal como señaló el Relator especial de la ONU Philip Alston en su informe anual de 2019 en relación con el algoritmo SyRI del gobierno holandés para la preselección de población más proclive a cometer fraude a la seguridad social.

Respecto de la asistencia en fase de ejecución, a través de un algoritmo para compilar inembargabilidades tal y como propone Jordi Nieva Fenoll,⁶² se trata de una actividad que produce evidentes efectos jurídicos que afectan significativamente al sujeto, por lo que su funcionamiento no debería estar plenamente automatizado de acuerdo con el art. 22.1 RGPD. Su funcionamiento debería incorporar la presencia humana en la toma de la última decisión, con independencia suficiente respecto del criterio del algoritmo.

Es preciso advertir que la garantía de una intervención humana a la que obliga el artículo 22 RGPD no resulta aplicable si la decisión, aun siendo plenamente automatizada y comportando efectos significativos, (1) es necesaria para la celebración o la ejecución de un contrato entre el interesado y el responsable del tratamiento, (2) está autorizada por el Derecho de la UE o de alguno de los Estados Miembros, o (3) se basa en el consentimiento explícito del interesado (art. 22.2 RGPD), siempre que en cualquiera de los casos la normativa o el responsable del tratamiento en su caso contemplen medidas adecuadas de salvaguarda de los derechos, libertades e intereses legítimos del interesado.

La autorización por una norma de la UE o de los Estados miembros es, por tanto, título para la admisibilidad de las decisiones automatizadas. En los Estados miembros de la Unión Europea se detectan posturas diferentes.

- Como se ha dicho, Estonia ha anunciado el proyecto para la creación de un juez robot. No obstante, esta opción ha sido propuesta en Estonia para un tribunal que se encargaría de resolver litigios de escasa cuantía de hasta 7.000 Euros,⁶³ del que sin embargo no se tiene por ahora ulterior noticia.
- En Francia, aunque no en el ámbito judicial sino en el de la resolución alternativa de litigios, se ha establecido por ley de 2019 que las empresas que proporcionan servicios en línea de mediación o de arbitraje tienen prohibido que la decisión final sea dejada por completo en manos de los algoritmos de las máquinas, debiendo existir siempre una fase última que garantice la presencia humana.⁶⁴

⁶² J. Nieva Fenoll, *Inteligencia artificial y proceso judicial*. Madrid: Marcial Pons, 2018.

⁶³ See Eric Niiler, (2019) *Can AI be a Fair Judge in Court? Estonia Thinks So*. <https://www.wired.com/story/can-ai-be-fair-judge-court-estonia-thinks-so/> (fecha de consulta: 5 de febrero de 2021).

⁶⁴ Véase el art. 4.3 [LOI n° 2019-222 du 23 mars 2019 de programmation 2018-2022 et de réforme pour la justice](#).

F. Esteban de la Rosa, “Construyendo el marco legal europeo para el diseño y uso de herramientas de inteligencia artificial en el entorno de la justicia civil-administrativa”, en el libro F. Herrera Triguero, A. Peralta Gutiérrez y L.S. Torres López, Universidad de Granada, pp. 423-451.

En España, el actual artículo 42 de la LUTICAJ contempla la “actuación judicial automatizada”. A pesar del tenor de la disposición, que se refiere a actos de trámite, actos de comunicación y actos resolutorios de procedimientos, cabe inferir que la disposición no se refiere a sentencias, no sólo porque por su importancia debería ser objeto de mención explícita sino porque la producción de sentencias automatizadas podría contravenir los artículos 117.3 y 120.3 en relación con los artículos 18.4 y 24 de la Constitución⁶⁵. Con carácter general la norma podría referirse a supuestos de actos de trámite como diligencias de ordenación o de constancia (art. 456.2 LOPJ), actos de comunicación como diligencias de comunicación (art. 456.2 LOPJ) o notificaciones (art. 150 LEC) y actos resolutorios de procedimientos como los autos de sobreseimiento (arts. 414.3, 415, 421, 423, 566, 688 LEC, entre otros).

Sobre este tipo de decisiones automatizadas resulta de interés el criterio propuesto por el Anteproyecto de Ley de Medidas de Eficiencia Digital del Sistema Público de Justicia, que distingue entre actuaciones automatizadas proactivas y actuaciones asistidas. El Anteproyecto contempla que los sistemas informáticos utilizados en la Administración de Justicia posibiliten la automatización de las actuaciones de trámite o resolutorias simples, que no requieren interpretación jurídica. Las actuaciones proactivas son definidas como actuaciones automatizadas, auto-iniciadas por los sistemas de información sin intervención humana, que aprovechan la información incorporada en un expediente o procedimiento de una Administración pública con un fin determinado, para generar avisos o efectos directos a otros fines distintos, en el mismo o en otros expedientes, de la misma o de otra Administración pública, en todo caso conformes con la ley.

Entre tales actos se citan expresamente: a) el numerado o paginado de los expedientes; b) la remisión de asuntos al archivo cuando se den las condiciones procesales para ello; c) la generación de copias y certificados; d) la generación de libros; e) la comprobación de representaciones; f) la declaración de firmeza, de acuerdo con la ley procesal. El legislador ha previsto ciertas cautelas respecto de estas actuaciones automatizadas que resultan coherentes con las exigencias que provienen del artículo 22 RGPD, que se refiere expresamente a la adopción de “medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado”. Entre las medidas previstas se encuentran las siguientes: a) que todas las actuaciones automatizadas y proactivas se puedan identificar como tales, trazar y justificar; b) que sea posible efectuar las mismas actuaciones en forma no automatizada; c) que sea posible deshabilitar, revertir o dejar sin efecto las actuaciones automatizadas ya producidas.

⁶⁵ La potestad jurisdiccional corresponde en exclusiva a Juzgados y Tribunales (art. 117.3 CE) y las sentencias deben ser siempre motivadas (art. 120.3 CE). El art. 18.4 establece que la ley limitará el uso de la informática con el objetivo, entre otros, de garantizar el pleno ejercicio de los derechos. Este precepto aconseja una interpretación de los dos anteriores según la cual las sentencias (exponente por excelencia del ejercicio de potestad jurisdiccional) deben ser emitidas y motivadas por las personas titulares de Juzgados y Tribunales. Excluir la intervención humana en la producción de sentencias chocaría frontalmente con los tres preceptos mencionados y con el derecho a la tutela judicial efectiva del art. 24 CE.

F. Esteban de la Rosa, “Construyendo el marco legal europeo para el diseño y uso de herramientas de inteligencia artificial en el entorno de la justicia civil-administrativa”, en el libro F. Herrera Triguero, A. Peralta Gutiérrez y L.S. Torres López, Universidad de Granada, pp. 423-451.

Dado que el artículo 22 RGPD se refiere a decisiones basadas “únicamente” en el tratamiento automatizado, escapan a la prohibición de este precepto aquellas actuaciones en las que la intervención humana queda garantizada en las actuaciones judiciales automatizadas. Por esta razón quedan fuera de la prohibición las denominadas por el ALMEDSPJ actuaciones asistidas, que son definidas como aquellas para las que el sistema de información de la Administración de Justicia genera un borrador total o parcial de documento complejo con base en datos, que puede ser producido por algoritmos y puede constituir fundamento de apoyo de una resolución judicial o procesal. El Anteproyecto da así entrada en nuestro sistema judicial a una fundamentación jurídica que pueda ser construida por algoritmos, sometida al único requisito de que pueda ser revisada por el juez.

Un ejemplo de este tipo de automatización es ofrecido por los nuevos Tribunales de Internet creados en China. El sistema permite la generación automática de un documento de sentencia. El documento generado tiene caracteres verdes y negros. Los caracteres negros son la parte fija del borrador. Los caracteres verdes contienen la información personalizada de las partes, el número del caso, los hechos del caso y las disposiciones legales aplicables, todo lo cual es recibido desde la plataforma. Solo se tarda diez segundos en formar el borrador del documento tras la vista y el juez entonces solo tiene que revisarlo y hacer las modificaciones necesarias para finalizar el documento que contendrá la sentencia.⁶⁶

La intensidad y libertad de la presencia humana conectan con la diferencia entre una decisión y una sugerencia automatizadas. El criterio distintivo entre una y otra es la noción de “intervención humana significativa” o de la capacidad de la persona para influir realmente en la decisión que finalmente llegue al sujeto, porque tiene autoridad y competencia para influir y no se verá expuesta a represalia alguna (por ejemplo, mayor grado de responsabilidad en caso de fallos) por apartarse del criterio automatizado.⁶⁷

El artículo 22 RGPD da así respuesta a muchas de las cuestiones que plantea la admisibilidad y utilización de decisiones automatizadas en el ámbito de la justicia civil-administrativa. Se trata de un régimen que debe ser considerado como uniforme para el conjunto de los países miembros de la Unión Europea.

Una vez que se parte de la admisibilidad de la decisión automatizada en el entorno de la administración de justicia, resulta necesario trazar con la mayor claridad posible los límites para esta posibilidad.

Debe garantizarse la intervención humana en el ámbito de las decisiones que tienen lugar en la Administración de justicia con carácter general, lo cual no impide la

⁶⁶ Xuhui Fang, “Recent ODR Developments in China”, *International Journal on Online Dispute Resolution*, 2017 (4), 2, pp. 32-37.

⁶⁷ Sobre este particular, véase las WP29 Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679, 6 de Febrero de 2018, p. 21. Accesible en: <https://ec.europa.eu/newsroom/article29/items/612053>; y, aplicado a un caso concreto, G. Lazcoz Moratinos, Automated Decision-Making Under Amsterdam’s District Court Judgements: Drivers v. Uber and Ola en Time to reshape the digital society. 40th anniversary of the CRIDS, Éditions Larcier, 263-264.

F. Esteban de la Rosa, “Construyendo el marco legal europeo para el diseño y uso de herramientas de inteligencia artificial en el entorno de la justicia civil-administrativa”, en el libro F. Herrera Triguero, A. Peralta Gutiérrez y L.S. Torres López, Universidad de Granada, pp. 423-451.

automatización progresiva del sector, siempre que se respete el núcleo esencial del derecho a la tutela judicial efectiva.

Es recomendable el fomento de la automatización de actuaciones procesales, respetando las siguientes pautas:

- La intervención humana debe ser más intensa y libre cuanto mayor sea la trascendencia del acto procesal y, consiguientemente, sus efectos en el ciudadano.
- Los actos de mero trámite, comunicación y, en general, no definitivos pueden ser sometidos a una mera supervisión humana pasiva general.
- Los actos de calificación o resolutorios y, en general, aquellos que en las normas de procedimiento no necesiten de un proceso deliberativo y argumentativo, pueden ser sometidos a una mera supervisión humana individualizada.
- Los actos resolutorios que necesiten de un proceso deliberativo y argumentativo en el marco del derecho a la tutela judicial efectiva deben contar con una “intervención humana significativa” individualizada, lo que no excluye que se vean asistidos por programas automatizados como, por ejemplo, aquellos que proveen de elaboración automatizada de argumentaciones frecuentes.
- Los Jueces y Tribunales y, en general, titulares de órganos encargados de dictar resoluciones (incluido el ámbito administrativo), deben gozar de libertad para apartarse de las sugerencias automatizadas, en cumplimiento de los criterios del WP29 en esta materia, que resultan plenamente aplicables al sector de la administración de justicia.

Un uso distinto de las herramientas de IA se produce cuando la decisión de la máquina se configura para sustituir a la del juez. Por ahora solo se conoce un proyecto de este tipo, el anunciado por el Gobierno de Estonia relativo a la creación de un juez robot con la finalidad de decidir en casos de escasa cuantía de hasta 7.000 Euros.⁶⁸ A día de hoy, sin embargo, no se tiene más noticia de esta iniciativa.

V. Consideraciones finales

La orientación que se propone para el marco legal que completará la regulación de la gobernanza de la IA respecto del uso de herramientas de IA en el ámbito de los servicios legales y en el entorno de la administración de justicia parecen marcar una línea de solución razonable que abre el juego a la utilización de la IA en la mayoría de los ámbitos jurídicos a través de un régimen dual que separa entre sistemas de IA de alto riesgo y sistemas de IA que no son de alto riesgo dependiendo del contexto de uso.

El uso de la IA en el ámbito jurídico posee un campo importante de desarrollo en materia de anonimización de datos relativos a sentencias, transcripción de declaraciones o utilización de traducciones, así como en tareas rutinarias como la programación de audiencias o la generación automática de documentos. Muchas de estas actividades

⁶⁸ Véase Eric Niiler, (2019) Can AI be a Fair Judge in Court? Estonia Thinks So. <https://www.wired.com/story/can-ai-be-fair-judge-court-estonia-thinks-so/> (fecha de consulta: 7 de julio de 2022).

F. Esteban de la Rosa, “Construyendo el marco legal europeo para el diseño y uso de herramientas de inteligencia artificial en el entorno de la justicia civil-administrativa”, en el libro F. Herrera Triguero, A. Peralta Gutiérrez y L.S. Torres López, Universidad de Granada, pp. 423-451.

pueden suponer una gran contribución a la eficiencia de la Administración de Justicia y entrañan un riesgo bajo o nulo para los derechos y libertades. El marco normativo contribuye a dar respaldo al uso por parte del sistema judicial de herramientas de IA que pueden ser muy útiles y que no deben pasar de forma obligatoria por un proceso de supervisión antes de ser utilizadas. Para esta clase de sistemas, al igual que para los sistemas utilizados al margen del sistema público de justicia por profesionales, seguirán valiendo, en su mayor parte, las normas generales del régimen de protección de datos sin perjuicio de la adopción de códigos de conducta dispuestos para imitar la supervisión y control de los sistemas de IA de alto riesgo.

Al mismo tiempo la propuesta normativa europea lanza una llamada de atención respecto de la incorporación a los sistemas de justicia de herramientas para ser usadas para asesorar al juez en la interpretación o aplicación de la ley. Estas herramientas sí que estarán sometidas al régimen de supervisión más estricto para poder ser utilizadas.

El uso de IA en materia de decisiones automatizadas debe tener muy presentes los límites marcados por el artículo 22 del RGPD, que fundamentalmente exigen que se garantice una intervención humana libre que supervise las sugerencias automatizadas proporcionadas por la IA cuando éstas puedan producir efectos jurídicos o significativos en la persona. Por otra parte, debe evitarse el riesgo de estigmatización que puede producir una excesiva confianza en decisiones o preselecciones automatizadas, por mucho que éstas sean posteriormente revisadas por un humano. Con todo, el éxito de esta clase de asistencia a los jueces, a modo de inteligencia aumentada, podría quedar asegurado en sectores específicos de litigios con pocas variables y con pruebas limitadas, tal y como sucede con el caso de uso del proyecto que se ha presentado en el tribunal de distrito de Frankfurt en mayo de 2022.

Junto a esta clase de contexto de uso, en una visión innovadora del uso de sistemas de IA los sistemas judiciales están comenzando a ofrecer máquinas de IA a los usuarios de la justicia. En nuestra opinión, dado que el diagnóstico y evaluación de la situación litigiosa puede interferir con el acceso a la tutela judicial efectiva, y a la vista de la autoridad de quien proviene el diagnóstico, sería deseable que estas máquinas fueran consideradas como sistemas de IA de alto riesgo por el futuro marco europeo para la gobernanza de la IA.