Revista Española de Derecho Europeo
Núm. 76 | Octubre – Diciembre 2020
pp. 9-44
Madrid, 2020
DOI:10.37417/num_76_2020_534

ISSN: 2695-7191
Recibido: 20/12/2020 | Aceptado: 15/01/2021

EL LIBRO BLANCO SOBRE INTELIGENCIA ARTIFICIAL DE LA COMISIÓN EUROPEA: REFLEXIONES DESDE LAS GARANTÍAS ESENCIALES DEL PROCESO PENAL COMO "SECTOR DE RIESGO"

THE EUROPEAN COMMISSION'S WHITE PAPER ON ARTIFICIAL INTELLIGENCE: REFLECTIONS ON THE ESSENTIAL GUARANTEES OF CRIMINAL PROCEEDINGS AS A "RISK SECTOR"

Montserrat de Hoyos Sancho*

RESUMEN: Se realiza en este trabajo un análisis crítico de los aspectos del Libro Blanco sobre Inteligencia Artificial de la Comisión Europea relacionados con los derechos y garantías esenciales en un proceso penal. Expuestas las características clave del que será el nuevo marco regulador y los potenciales riesgos a que se tendrá que hacer frente en este ámbito, se abordan algunas herramientas basa-

* Catedrática de Derecho Procesal y exdirectora del Instituto de Estudios Europeos de la Universidad de Valladolid. Correo-e: montserrat.dehoyos@uva.es. ORCID: 0000-0002-0972-2259.

Este trabajo se enmarca en los siguientes Proyectos y Grupos de Investigación: Plan Nacional I+D+i –Excelencia–Ministerio de Economía y Competitividad: "Garantías procesales de investigados y acusados: necesidad de armonización y fortalecimiento en el ámbito de la Unión Europea" –DER 2016-78096-P–; Junta de Castilla y León: "Sociedades seguras y garantías procesales: el necesario equilibrio" –VA-135-G18–; Generalitat Valenciana: "Claves de la justicia civil y penal en la sociedad del miedo" –Prometeo 2018/2011-; Grupo de Investigación Reconocido, Universidad de Valladolid: "Garantías procesales y Unión Europea"; FEDER-Junta de Andalucía: "Derechos y garantías de las personas vulnerables en el Estado del Bienestar" –UMA18-JA175– y «El uso de las TICs en la cooperación jurídica penal internacional: construyendo la sociedad digital andaluza del futuro" -P18-RT-1059.

das en Inteligencia Artificial, en materia probatoria y para el análisis predictivo de comportamientos futuros. Se concluye exponiendo los que se consideran requisitos esenciales para que los sistemas IA puedan ser una valiosa herramienta de ayuda en el ejercicio de la función jurisdiccional, sin merma de las garantías procesales esenciales, en el contexto del "espacio de libertad, seguridad y justicia" de la Unión Europea.

PALABRAS CLAVE: Unión Europea; inteligencia artificial; garantías procesales; prueba penal; análisis predictivos; armonización legislativa.

ABSTRACT: This paper makes a critical analysis of the aspects of the White Paper on Artificial Intelligence launched by the European Commission that are in relation with the essential rights and safeguards within criminal proceedings. After exposing the key features of the new legal framework and the potential risks that will have to be faced in this matter, several tools based on artificial intelligence, on issues of proof and on predictive analysis of future behaviour are addressed. It concludes by considering and exposing which essential requirements are deemed necessary for the AI systems to be a valuable tool to assist with the exercise of the jurisdictional function, without undermining the essential procedural safeguards, in the context of the "Area of Freedom, Security and Justice".

KEYWORDS: European Union; artificial intelligence; procedural safeguards; criminal evidence; predictive analysis; legislative harmonisation.

INTRODUCCIÓN

El 19 de febrero de 2020 se publicó en Bruselas el *Libro Blanco sobre la inteligencia artificial – un enfoque europeo orientado a la excelencia y la confianza*, elaborado por la Comisión Europea.

Como todo Libro Blanco, tiene por finalidad poner de manifiesto las necesidades normativas y, más en general, de intervención de las instituciones europeas en un determinado sector; al mismo tiempo, hacer propuestas de base para tratar de solventar problemas y mejorar ese ámbito de manera armonizada en el contexto de la Unión Europea. Todo ello se pretende trasladando una consulta al conjunto de los expertos y colectivos implicados, de forma que se active institucionalmente el *feed-back* que permita conocer su

parecer y sus propuestas, para de esa manera poder diseñar una acción política y legislativa en materia de inteligencia artificial –IA en lo sucesivo-, con el mayor acierto posible.

En cualquier caso, no es este, ni mucho menos, el primer compendio institucional que se publica sobre la materia en el contexto europeo. Entre otros, debemos destacar la *Carta ética europea sobre el uso de la inteligencia artificial en los sistemas judiciales y su entorno*, aprobada por la Comisión europea para la eficacia de la justicia —CEPEJ[1]— con fecha de 4 de diciembre de 2018. Este documento es un antecedente particularmente relevante, porque parte de su contenido encuentra reflejo directo en el *Libro Blanco* que ahora nos ocupa.

La referida *Carta ética* es un documento de *soft law*, pero ya se concluye en este, entre otros extremos, la importancia de observar en este ámbito las siguientes garantías que resumimos a continuación: cuando los instrumentos IA se empleen como apoyo a los procesos, se debe asegurar que no se vulnera el derecho de acceso al juez y el derecho a un proceso equitativo –igualdad de armas y respeto al contradictorio-; que se respeta el principio de no discriminación; que los datos manejados por los sistemas son seguros y fiables, obtenidos en procedimientos cuya trazabilidad puede conocerse y cuya integridad puede asegurarse; que se garantiza la transparencia e imparcialidad técnica del sistema –el proceso algorítmico ha de ser accesible-, lo cual pasa por crear una autoridad pública independiente que pueda valorar y certificar las herramientas IA *a priori,* y después monitorizar su funcionamiento; finalmente, el sistema IA debe permanecer en todo momento bajo el control de los usuarios del mismo, los cuales no estarán necesariamente vinculados a las soluciones sugeridas por tal sistema IA.

En el Apéndice I a dicha "*Carta*" se insiste en el necesario respeto a los principios de igualdad de armas y de presunción de inocencia, y se destaca la importancia de que la persona interesada –su defensa- tenga la posibilidad de contestar la validez científica del algoritmo y el peso atribuido a los respectivos datos de que se nutre; en todo caso ha de respetarse el "derecho de acceso al juez" y han de protegerse los datos personales[2].

Por lo demás, a estas alturas del desarrollo y aplicaciones de la Inteligencia Artificial pocas dudas caben sobre las importantes consecuencias que su empleo tiene y tendrá sobre muchos sectores económicos y sociales en sentido amplio. Desde luego, la administración de justicia no está, ni estará, al margen de tal influencia, tratándose además de un ámbito que el propio *Libro*

---

[1] Organismo creado en 2002 e integrado en el Consejo de Europa. *Vid.* más ampliamente: https://www.coe.int/en/web/cepej/home/.

[2] Un comentario sobre los efectos de esta *Carta Ética* desde la perspectiva procesal puede encontrarse en Gialuz (2019, esp. pp. 12 y ss.). *Vid.* también, anteriormente, Quattrocolo (2018). En España, valora el instrumento Martín Diz (2019a, esp. pp. 822 y ss.).

Blanco califica como "sector de elevado riesgo", por las razones que después expondremos.

Es preciso por tanto que también desde la perspectiva del Derecho Procesal se lleve a cabo una lectura y análisis crítico de este *Libro Blanco*, en cuanto que algunos de sus contenidos afectan de manera directa e inmediata a derechos y garantías procesales, esenciales en todo el contexto de la Unión Europea, los cuales desde luego deben ser preservados con ahínco, por mucho que el empleo de la IA pueda influir y, en cierta medida, cambiar algunos aspectos de la forma de proporcionar tutela judicial efectiva a los ciudadanos.

Comenzando por la propia definición de IA[3], el *Libro Blanco* nos ofrece la que ha de servir de sirve de base para su estudio, de cara a las futuras iniciativas políticas, y por tanto la que asumiremos también en este nuestro análisis:

> "Los sistemas de inteligencia artificial son programas informáticos —y posiblemente también equipos informáticos— diseñados por seres humanos que, dado un objetivo complejo, actúan en la dimensión física o digital mediante la percepción de su entorno mediante la adquisición de datos, la interpretación de los datos estructurados o no estructurados, el razonamiento sobre el conocimiento o el tratamiento de la información, fruto de estos datos y la decisión de las mejores acciones que se llevarán a cabo para alcanzar el objetivo fijado".

Los principales elementos que integran un sistema IA son los "datos" y los "algoritmos[4]"; a su vez, la IA puede incorporarse en distintos tipos de soportes o equipos informáticos.

En relación con las llamadas "técnicas de aprendizaje automático", el *Libro Blanco* explica que constituyen un subapartado de la IA. Es preciso tener en cuenta que los algoritmos pueden ser entrenados para inferir concretos modelos a partir de un conjunto de datos, con el objetivo de determinar las acciones que se requieren para alcanzar un objetivo prefijado. Así, los algoritmos podrán configurarse para seguir aprendiendo mientras se utilizan y a medida que se van nutriendo progresivamente de nuevos datos, lo que se conoce como "aprendizaje automático" —*Machine Learning*—[5].

---

[3] Probablemente sería posible encontrar tantas definiciones de IA como analistas se han ocupado del tema desde las distintas perspectivas. Por tal motivo, preferimos acoger directamente la que se maneja en el propio *Libro Blanco*, que además es clara y concisa. Trae su origen de la definición previamente elaborada por el "Grupo de expertos de alto nivel" de la Comisión Europea, con base en la *Comunicación sobre la inteligencia artificial para Europa,* redactada por la propia Comisión (2018). Así se indica expresamente en la p. 20 del *Libro Blanco*.

[4] En la citada Carta Ética Europea se definen los algoritmos, precisamente a estos efectos que nos ocupan —*vid.* p. 69— como "*Finite sequence of formal rules (logical operations and instructions) making it possible to obtain a result from the initial input of information. This sequence may be part of an automated execution process and draw on models designed through machine learning*".

[5] Estos sistemas son capaces de identificar patrones no predeterminados y de generar nuevas relaciones entre tales patrones y los nuevos datos de que se nutren, de forma que pueden hacer surgir sucesivas predicciones o recomendaciones, en prin-

En todo caso, si bien es cierto que los productos basados en sistemas IA pueden llegar a funcionar de manera autónoma a partir de su percepción del entorno y sin seguir un conjunto predefinido de instrucciones tasadas, su comportamiento y objetivos sí estarán definidos, programados y acotados previamente por las personas, pues son estas quienes diseñarán los distintos sistemas de IA tratando de optimizar la consecución de tales resultados pretendidos por cada uno de ellos[6].

Por lo que respecta a la definición de aplicaciones de IA de "riesgo elevado", como son las que se utilizarían en la lucha contra la delincuencia o en la administración de justicia en general, precisamente por los derechos y libertades fundamentales que pueden verse afectados, se insiste en el *Libro Blanco* en que "debe ser clara y fácil de entender y de aplicar para todas las partes interesadas", por la relevancia y consecuencias del tipo de decisiones que se adoptan en ese contexto.

Y precisamente en ese ámbito tan amplio que es la administración de justicia, resulta pertinente realizar una distinción relevante para nuestro análisis: entre la IA disponible para los abogados[7], por un lado, de aquella otra que podría ser utilizada por los que ejercen función jurisdiccional, jueces y magistrados. En este estudio nos centraremos en esta última, en el contexto geográfico y normativo de la Unión Europea.

Al margen quedarán también en este momento las referencias a los mecanismos de mera "automatización" o "digitalización" de la administración de justicia, como pueden ser los sistemas automáticos o telemáticos que se utilizan simplemente para facilitar y agilizar la tramitación de expedientes, o para las notificaciones procesales, ya que por sí mismos, tal automatiza-

---

cipio no previstas en las especificaciones de la programación inicial del algoritmo. *Vid*. más ampliamente, entre otros muchos trabajos sobre la materia, Mitchel (2020), pp. 43 y ss. Sobre su uso en los sistemas jurisdiccionales de Inglaterra y Gales, *vid*. *Algorithms in the Criminal Justice System,* Informe elaborado por *The Law Society* (2019, pp. 10 y 11), sobre tipos y utilidades del *Machine learning* y *Deep learning*. En España, Barona Vilar (2019, pp. 44 y ss.) y Martín Diz (2019b, pp. 533 y ss.).

[6] Los sistemas basados en la IA pueden consistir simplemente en un programa informático –*v. gr*.: asistentes de voz, programas de análisis de imágenes, motores de búsqueda, sistemas de reconocimiento facial y de voz…-, pero la IA también puede estar incorporada en dispositivos de *hardware* –*v.gr*.: robots avanzados, automóviles autónomos, drones o aplicaciones del internet de las cosas-, como explica el citado documento Comisión (2018, p. 1).

[7] Por ejemplo, la llamada "Jurimetría del Caso, del Magistrado o del Abogado", que proporciona a los letrados información relativa al tratamiento que los órganos jurisdiccionales hacen de un tipo delictivo concreto, y su correlación con otros delitos, lo que facilita la argumentación procesal en la causa, pues además ofrece un estudio de la trayectoria y líneas jurisprudenciales de los distintos jueces y magistrados, o incluso del argumentario de la contraparte en ese tipo procesos, entre otras muchas utilidades. Las principales editoriales y bases de datos jurídicas de carácter privado ya ofrecen este tipo de herramientas a los profesionales del derecho.

ción de trámites, la celebración de juicios telemáticos u *online*[8], o el empleo de escritos y documentos en soporte electrónico, no implica el uso de IA en sentido estricto[9].

En este trabajo pretendemos extraer del *Libro Blanco* de la Comisión Europea aquellos aspectos que pueden ser más relevantes para el ejercicio de la función jurisdiccional, abordar algunas de las posibilidades presentes y futuras de los sistemas IA en ese contexto, reseñando al mismo tiempo las limitaciones que para tal uso conlleva la imprescindible garantía de los derechos y libertades fundamentales del ciudadano en relación con el *debido proceso*. En particular, dada la amplitud del tema y la necesariamente limitada extensión de este trabajo, centraremos nuestro análisis en las posibles fuentes y medios de prueba que se asientan en la aplicación de sistemas IA, en el uso de algoritmos de ayuda a la valoración judicial de la prueba, así como en los llamados "algoritmos predictivos" que pueden auxiliar al juez o tribunal en el desempeño de sus funciones, concretamente en el específico contexto del proceso penal.

1. PREMISAS Y FINALIDADES DEL LIBRO BLANCO SOBRE INTELIGENCIA ARTIFICIAL DE LA COMISIÓN EUROPEA. CARACTERÍSTICAS CLAVE DEL NUEVO MARCO REGULADOR

Ya en sus primeras líneas el *Libro Blanco* pone sobre el tapete los principales riesgos potenciales que hay que enfrentar cuando se hace uso de los sistemas de IA: la opacidad en la toma de decisiones, las discriminaciones de género o de otro tipo, la intromisión en la privacidad y su uso con fines delictivos.

Por tanto, se hace preciso arbitrar una acción política y normativa que minimice en todo lo posible estos riesgos, de tal forma que se pueda generar un *ecosistema de confianza* con un enfoque antropocéntrico[10], como objetivo político en sí mismo que ha de ofrecer suficiente seguridad para aprovechar al máximo las oportunidades de negocio, de progreso científico y, en general, de mejora social que puede ofrecer la IA.

---

[8] De imprescindible consulta en este punto los trabajos de Susskind (2020, pp. 175 y ss.); en pp. 182 y ss. explica el autor los que, a su juicio serían los casos más adecuados para el enjuiciamiento *online*. Además de la cuantía del asunto –para causas civiles-, que no sería determinante, habría que considerar: complejidad de la ley aplicable, patrón de hechos, volumen de documentación que integra la causa, sensibilidad de los datos en conflicto, dificultad de los problemas legales objeto del litigio, credibilidad de los testigos y eficacia actual de los procesos de ese tipo.

[9] Nieva Fenoll (2018) califica estos sistemas de "IA *débil*" (pp. 24 y ss).

[10] Puede consultarse además el documento *Generar confianza en la inteligencia artificial centrada en el ser humano* (Comisión, 2019).

Se muestra necesario entonces actuar de manera conjunta en el ámbito UE, para poder así mantener un nivel adecuado de competitividad en el exigente contexto mundial, evitando una fragmentación del mercado único —destaca la Comisión—, a la vez que se garantiza correctamente el respeto a los derechos y libertades de los ciudadanos y de las personas jurídicas. La IA se ha de asentar en nuestros valores y derechos fundamentales[11], como son la dignidad humana y la protección de la privacidad.

Por todo ello, expresa también la Comisión su deseo y la importancia de que se formulen alternativas políticas que permitan alcanzar esos objetivos, así que se invita a los Estados miembros, a otras instituciones europeas, a la industria, a los interlocutores sociales, a las organizaciones de la sociedad civil, a los investigadores y a los ciudadanos en general interesados en materia, a que presenten sus respectivas opiniones sobre las opciones que se plantean en dicho *Libro Blanco*, que servirán para que la Comisión tome sus decisiones en este ámbito.

Conviene destacar al mismo tiempo la trascendencia que los datos tienen para el correcto funcionamiento de la IA; es por ello que el plan que se articule en torno a la IA debe estar estrechamente ligado a lo que la propia Comisión denomina la *Estrategia Europea de Datos*, que incluye una gestión responsable de los mismos, cumpliendo en todo caso los conocidos como "principios FAIR"[12]. Y desde luego es preciso que toda esta acción conjunta sobre IA y datos, pueda contar con el soporte *hardware* adecuado; por tanto, se deberán realizar las inversiones necesarias en infraestructura y tecnología informática[13], a lo que deberá unirse un esfuerzo en formación de operadores y usuarios.

En la generación y mantenimiento del referido *ecosistema de confianza* pretendido por la Comisión, el establecimiento de un marco regulador juega sin duda un papel determinante[14]; en particular, en aquellos sectores en los que los derechos de los ciudadanos se pueden ver afectados de manera más directa, como "por ejemplo, en el caso de las aplicaciones de IA empleadas por los cuerpos y fuerzas de seguridad y el poder judicial", destaca el *Libro Blanco*[15]. A la vez se pone de relieve que, si bien los desarrolladores e implementadores de IA ya están sujetos a toda la legislación europea

---

[11] *Vid. Libro Blanco*, p. 2.

[12] Acrónimo de "Fáciles de encontrar, Accesibles, Interoperables y Reutilizables". *Vid.* el informe final y el Plan de Acción del Grupo de Expertos en datos FAIR de la Comisión (2018).

[13] Aspecto que también se destaca en el propio *Libro Blanco*, *vid.* esp. p. 10.

[14] En los estudios realizados sobre el sector empresarial en 2019 se observó que en los regímenes jurídicos o reguladores de numerosos sectores económicos existían carencias en relación con los requisitos de transparencia y supervisión humana de los sistemas IA. *Vid. Libro Blanco*, p. 12, con referencia al *Informe del Grupo de expertos de alto nivel*, creado por la Comisión, que en abril 2019 publicó las que se debían considerar *Directrices* éticas *para una IA fiable*.

[15] *Vid.* p. 12.

en materia de derechos fundamentales —protección de datos, privacidad, no discriminación…—, "algunas características específicas de la IA, como la opacidad, pueden hacer que la aplicación y ejecución de la legislación sea más compleja".

Por tanto, es preciso revisar la legislación actual, comprobar que reúne las condiciones para hacer frente a los riesgos que plantea el uso de la IA, también en el contexto de la administración de justicia, y en su caso, modificarla o completarla en los extremos que fuera necesario.

No obstante, ese "marco regulador sólido que garantice una IA fiable" al que se refiere la Comisión[16], deberá permitir que se deje un margen para el desarrollo a corto-medio plazo, pues la IA evoluciona a una velocidad vertiginosa, y la legislación UE no puede ser un corsé que constriña su evolución[17], siempre que esta sea acorde con los derechos, libertades y garantías esenciales de los ciudadanos y de las personas jurídicas.

El referido marco regulador armonizador en el contexto UE debe ofrecer, desde luego, seguridad jurídica, así como un suficiente grado de certeza a los Estados y a los operadores. En todo caso, también ha de tratar de minimizar los riesgos sobre la vigencia de los derechos fundamentales de la ciudadanía; en este punto, según hemos indicado, la Comisión destaca los siguientes: protección de datos, privacidad, no discriminación y seguridad[18], aunque ya el Consejo de Europa[19] puso de relieve otros derechos fundamentales que también podían verse afectados por el uso de algoritmos y otras técnicas de procesamiento automatizado de datos, como serían los siguientes: *fair trial* y *due process*, libertad de expresión, libertad de reunión y asociación, disponibilidad de recursos efectivos contra la violación de un derecho, derechos sociales y acceso a servicios públicos, e incluso el derecho a elecciones libres.

Todos estos riesgos, o incluso algún otro más indirecto, pueden ser el resultado de defectos en el diseño de los sistemas de IA, vgr.: errores técnicos, falta de supervisión humana…, y/o por el uso de datos erróneos, insuficientes

---

[16] *Vid. Libro Blanco,* p. 13.

[17] Sobre la necesidad de introducir nuevas formas regulatorias que permitan una revisión regular de la materia, un buen ritmo de adaptación normativa a las circunstancias que se vayan planteando, llegando a sugerir incluso la pertinencia de una "regulación líquida o biodegradable" en determinados aspectos del uso de la IA, *vid.* Cotino Hueso (2019, pp. 21 y ss.).

[18] Dejamos al margen de nuestro estudio todas las cuestiones que tienen que ver con la responsabilidad derivada del uso de IA, materia desde luego de enorme trascendencia práctica. La materia se aborda en el *Informe sobre el marco de seguridad y responsabilidad civil de la Inteligencia Artificial, el Internet de las cosas y la robótica,* adjunto al propio Libro Blanco sobre IA. Sobre esta cuestión, en nuestro país, entre los más recientes, *vid.* el trabajo de Núñez Zorrilla (2019), y con referencia al contexto europeo, de la misma autora (2018, pp. 9 y ss.).

[19] En el documento *Algorithms and Human Rights. Study on the Human Rights Dimensions of Automated Data Processing Techniques and Possible Regulatory Implication* [Committee of Experts on Internet Intermediaries (MSI-NET), 2018].

o sesgados, introducidos en el sistema sin correcciones previas, *v.gr.:* referidos solo a hombres, sin considerar las especificidades de las mujeres u otros colectivos no binarios. Y desde luego, a nuestro juicio, la falta de transparencia sobre la configuración de los algoritmos empleados y sobre su funcionamiento en la toma de decisiones, o acerca del origen y tratamiento de los distintos tipos de datos de los que estos se nutren[20], son algunos de los principales peligros que acechan al uso de la IA en el ámbito de la administración de justicia, por su posible repercusión sobre el derecho de defensa e igualdad de partes, en definitiva, sobre el debido proceso o proceso con todas las garantías.

Como es sabido, las posibilidades de alegar, de probar o de impugnar, y sus respectivos resultados sobre la causa, dependen directamente de que se puedan conocer con suficiente detalle los datos y razonamientos conforme a los cuales resolverá el órgano decisor, y/o sobre los que basa su actuación la contraparte o posición acusadora en el proceso penal.

Por lo demás, si se llegaran a aplicar sistemas de IA en la adopción de decisiones jurisdiccionales sin la suficiente supervisión/control judicial de las resoluciones finales, también se vería afectado el derecho a la tutela *judicial* efectiva de ciudadanos y personas jurídicas.

Dejamos en todo caso a un lado los supuestos en que se pudiera hacer un uso fraudulento o delictivo de esos algoritmos y/o datos que, desde luego, también es un riesgo posible.

La Comisión Europea concluye que el futuro marco regulador y la utilización que se haga de estos instrumentos de IA en el contexto UE, concretamente en este sector de "riesgo elevado" que nos ocupa –la administración de justicia-, han de reunir una serie de "características clave" y prever controles eficaces que permitan minimizar los riesgos apuntados, siempre de forma específica y proporcionada. En todo caso, es muy importante que ofrezcan seguridad jurídica a ciudadanos y empresas.

Tales características esenciales son resumidamente las siguientes[21]:

Datos de entrenamiento. El funcionamiento de la IA, los resultados que con esta se puedan alcanzar y las decisiones que a través de ella de adopten, dependen directamente de los datos con los que se "entrena" el sistema, en

---

[20] La propia Comisión se refiere en el *Libro Blanco* al riesgo de que determinados algoritmos empleados en IA para predecir la reincidencia delictiva puedan partir de prejuicios raciales, de género, o de extranjería. Cierto es que cuando la decisión la debe tomar un humano, también existe el riesgo de subjetividad o discriminación, incluso inconsciente, pero cuando es un sistema de IA el que trabaja con sesgos discriminatorios fundados en prejuicios, las consecuencias pueden ser mucho más graves, pues afectarán a más personas, y además es muy difícil que se pueda materializar el control social de tales sesgos o prejuicios en el funcionamiento de la IA. *Vid. Libro Blanco*, p. 14.

[21] Se recogen en el *Libro Blanco*, pp. 22 y ss.

definitiva, de los que se nutre. Por lo tanto, deben adoptarse las medidas que garanticen que tal compilación de datos respeta los valores y normas de la UE en relación con la seguridad y protección de derechos fundamentales. Por ejemplo, debe poder asegurarse que los equipos IA se "entrenan" con conjuntos de datos suficientemente amplios, que no van a generar resultados que conlleven discriminaciones ilícitas, y que la privacidad de los datos está suficientemente protegida.

Conservación de registros y datos. Deben poder conservarse los registros sobre la metodología de programación de los algoritmos, así como conocer el origen de los datos con los que se "entrenan" los sistemas IA; es decir, saber de qué forma se seleccionaron y contar con un registro de los mismos[22]. Esto permitirá hacer un seguimiento y supervisar las decisiones que adoptan tales sistemas, lo que será especialmente útil en circunstancias eventualmente problemáticas o si se formulan reclamaciones.

Suministro de información. Los sistemas IA deben facilitar información adecuada y de manera proactiva acerca de cómo deben ser utilizados aquellos que se califican "de alto riesgo". En particular, tal información se referirá a las capacidades y limitaciones del sistema, a los objetivos a los que se destina, a las condiciones en que se espera funcione, así como al nivel de exactitud previsible en la consecución de sus objetivos. Todos esos datos serán muy relevantes, no solo para los implementadores del sistema, sino también para las autoridades competentes y/o las partes afectadas por su funcionamiento y decisiones.

Y muy importante también: la Comisión destaca que más allá de lo antedicho, debe informarse claramente a los ciudadanos de cuándo están interactuando con un sistema de IA y no con un ser humano, salvo cuando "sea inmediatamente evidente para los ciudadanos". Tal información ha de ser concisa y fácilmente comprensible, adaptada al contexto específico.

Solidez y exactitud. Todos los sistemas IA, y en particular las aplicaciones de "riesgo elevado", han de tener unas condiciones técnicas de solidez y precisión que las haga altamente fiables. En su desarrollo y configuración se habrán valorado de forma previa y adecuada todos los riesgos conocidos que su funcionamiento puede conllevar; además, se habrán adoptado todas las medidas razonables y posibles para reducir al mínimo el peligro de que su utilización, en todas las fases de su ciclo de vida prevista, pueda producir daños.

Será preciso arbitrar los mecanismos que permitan detectar errores o incoherencias en su funcionamiento, y tratar de que los sistemas IA sean resi-

[22] El registro, documentación y, en su caso, el propio conjunto de datos utilizados, se podrán/deberán conservar durante un tiempo en todo caso limitado, y deberán poder suministrarse, previa solicitud, en caso de ser requeridos por las autoridades competentes, adoptando las medidas que fueran necesarias para proteger la información confidencial.

lientes en caso de ataques o intentos de manipulación de los propios algoritmos o de los datos de que se nutre, articulando además las medidas precisas para evitar todos estos problemas.

Supervisión humana. Esta vigilancia por parte de personas ha de evitar que la IA provoque efectos no deseados o perversos, especialmente cuando opera en sectores de alto riesgo.

En particular, los resultados de la aplicación de un sistema IA no serán efectivos hasta que un humano los revise y valide[23]. Aunque tales resultados pudieran ser inmediatos, ha de garantizarse la intervención y el examen humano posterior.

Además, deberá ser posible realizar un seguimiento del sistema IA mientras funciona, intervenir en tiempo real y, en su caso, proceder a su desactivación —*v. gr.:* si no se dan las condiciones de seguridad necesarias—.

Por otro lado, determinados sistemas de IA pueden evolucionar y aprender de la experiencia, lo que requeriría vigilancia periódica, evaluaciones reiteradas, control de cumplimiento de las garantías y, si fuera preciso, la corrección de los errores que se detecten a lo largo del ciclo de vida de dichos sistemas[24].

Finalmente, en el caso de aplicaciones IA de "riesgo elevado", deberá garantizarse una "acción judicial efectiva" para las partes que puedan haber sufrido repercusiones negativas derivadas de tales sistemas IA.

El *Libro Blanco* de la Comisión le dedica también un apartado concreto a la cuestión de la identificación biométrica remota[25]; esta puede conllevar distintos riesgos específicos para los derechos fundamentales: vida privada, protección de datos personales, no discriminación…, en particular cuando los sistemas de reconocimiento facial se instalan en lugares públicos.

La normativa UE actualmente vigente impide, como regla general[26], el tratamiento de datos biométricos a fin de identificar de manera unívoca a una

---

[23] En el *Libro Blanco*, p. 25, se pone el ejemplo de la decisión que deniega una prestación de seguridad social, que solo podrá ser finalmente adoptada por un humano.

[24] *Vid*. *Libro Blanco*, p. 28.

[25] Sobre este particular –*Facial Recognition Technology*-, *vid*. Informe de la Agencia Europea para los Derechos Fundamentales (2020), así como las recomendaciones contenidas en el *Study on the use of innovative technologies in the justice field,* Informe Final presentado por la Comisión Europea (2020, pp. 40 y ss.).

[26] *Vid*. Reglamento 2016/679, sobre protección de datos de personas físicas –RGPD, esp. art. 9: prohibición de tratamiento de datos personales de carácter biométrico y sus posibles excepciones-; en semejantes términos, el art. 10 del Reglamento 2018/1725, sobre protección de las personas físicas en relación con el tratamiento de datos personales por instituciones, órganos u organismos UE y libre circulación de tales datos. *Vid.* también el art. 10 de la Directiva 2016/680, sobre protección de las personas físicas en lo relativo al tratamiento de datos personales en ámbito penal: se permitirá el tratamiento de los datos biométricos que permitan identificar de manera unívoca a una persona si lo autoriza el Derecho de la UE o del Estado miembro, si

persona física. Solo será posible de manera excepcional en supuestos en que esté justificado por concurrir un interés público significativo, con respeto al principio de proporcionalidad, a los derechos de protección de datos y demás garantías adecuadas al caso contenidas en la Carta Derechos Fundamentales UE, que deberán ser comunes "para evitar la fragmentación del mercado interior", afirma la Comisión en el *Libro Blanco*.

Entendemos no obstante que, más que para esa cuestión —garantizar la unidad del mercado interior—, este extremo —la garantía en todo caso de los derechos contenidos en la CDFUE— resulta de gran importancia para que estén debidamente armonizadas las condiciones de su uso en todo el "espacio de libertad, seguridad y justicia UE", ya que desde luego, la aprobación y vigencia de legislaciones nacionales posiblemente dispares en materia de tratamiento de datos relativos a la identificación biométrica remota de personas, puede afectar muy negativamente al reconocimiento mutuo de resoluciones judiciales penales; en particular, a la eficacia de la cooperación judicial y policial transfronteriza en la investigación y enjuiciamiento de hechos delictivos[27].

Expuestas las premisas y principales características del marco regulador que se propone en el *Libro Blanco*, nos ocuparemos a continuación de alguna de las utilidades que, a nuestro juicio, puede tener la IA en el contexto de la administración de justicia, concretamente en el proceso penal, respetando las pautas fijadas por la Comisión en el mismo, aunque en este momento lo haya hecho de forma genérica. Tendremos también en cuenta, lógicamente, las advertencias formuladas por el Consejo de Europa en el ya referido documento de marzo de 2018, sobre *Algoritmos y Derechos humanos*.

2. ALGUNAS HERRAMIENTAS BASADAS EN INTELIGENCIA ARTIFICIAL QUE PUEDEN EMPLEARSE EN EL CONTEXTO DE LA ADMINISTRACIÓN DE JUSTICIA

Según se ha indicado, tanto la actuación de las Fuerzas y Cuerpos de Seguridad en el campo de la prevención y lucha contra la delincuencia, como la función jurisdiccional, o incluso el ámbito de decisión de los que actúan en el contexto del poder judicial en sentido amplio, son considerados por la propia Comisión Europea como "sectores de riesgo elevado" en lo que se refiere al uso de los instrumentos de IA[28]. No obstante, estos sectores no deberían permanecer por mucho tiempo al margen de un uso regulado de la IA, así que

---

es necesario para proteger intereses vitales del interesado o de otra persona física, o bien si el tratamiento se refiere a datos que el interesado ha hecho públicos de forma manifiesta.

[27] Sobre esta cuestión, *vid*. más ampliamente De Hoyos Sancho (2019a y 2019b).

[28] *Vid. Libro Blanco*, pp. 21 y 30.

será preciso fijarnos en alguna de las principales herramientas actualmente disponibles basadas en esta tecnología que podrán utilizarse en el específico marco del proceso penal, y valorar entonces si pueden cumplir con los parámetros básicos que se enuncian en el *Libro Blanco* sobre IA.

De partida debemos tener en cuenta que el ámbito jurídico en general, y el jurisdiccional en particular, se caracteriza por ser refractario a los cambios; y lo son más aún a la incorporación de las innovaciones tecnológicas, ya sea por desconfianza hacia lo que no se conoce bien, ya por la dificultad de frenar la inercia en los métodos de trabajo, por el insuficiente soporte normativo necesario para operar en este sector, por el escaso grado de adaptabilidad de las personas que actúan en este contexto, por la obsolescencia y/o escasez de sus equipamientos técnicos, o por una suma de todas estas y otras razones.

No hay más que ver lo complicado que está siendo implantar en nuestro país el llamado "expediente digital", que ya funciona hace años en otros sectores de la Administración, tanto o más complejos, sin grandes dificultades, *v. gr.:* Agencia Tributaria; o incluso recordar la "revolución" que supuso en su día que todos los implicados confiaran en la "firma electrónica avanzada" como medio seguro en el tráfico jurídico[29].

No obstante lo dicho, más pronto que tarde acabaremos viendo un uso creciente de sistemas IA en la administración de justicia, como ya sucede en otros países, y tendremos que asegurarnos de que esos sistemas cumplen con los requisitos esenciales ya expuestos: respeto al debido proceso o proceso con todas las garantías, a los más concretos derechos de defensa e igualdad de partes, a la presunción de inocencia, a las garantías de privacidad, de no discriminación, de transparencia en la toma de decisiones, de seguridad de los sistemas, de control humano en la toma final de decisiones, etc. A todos estos requisitos deberá sumarse la suficiente previsión normativa que fuera necesaria en cada caso.

Por lo demás, tengamos presente que hoy por hoy los sistemas IA han de entenderse como "ayuda" al ejercicio de la función jurisdiccional por el órgano competente o, en su caso, a las labores de los letrados de la acusación y/o la defensa. Las decisiones intermedias y finales que hayan de adoptarse sobre la causa corresponderán necesariamente y en exclusiva al titular/es del juzgado o tribunal[30].

---

[29] *Vid.* en su día, De Hoyos Sancho (2003, pp. 4 y ss.).

[30] *Vid.* art. 22 del citado Reglamento 2016/679 –RGPD-, y también la mencionada *Carta ética* del CEPEJ: toda persona tiene derecho a no estar sometida a una decisión que produzca efectos jurídicos o tenga consecuencias significativas sobre ella, fundada exclusivamente en un tratamiento automatizado de datos destinados a valorar ciertos aspectos de su personalidad. En Italia, el Decreto legislativo de 18 de mayo de 2018, núm. 51, en su art. 8, establece la prohibición expresa de decisiones basadas únicamente en un tratamiento automatizado, incluida la "*profilazione*/perfilado", que produzcan efectos negativos sobre el interesado, salvo que lo autorice el Derecho de la UE o específicas disposiciones legales, y siempre y cuando estas establezcan garantías

2.1. Los sistemas IA como fuente de prueba y como ayuda en la valoración judicial de la prueba

Cuando hablamos de sistemas IA que pueden ser, primero fuente de prueba y, en su caso, posteriormente aportados al proceso como medios de prueba, estamos pensando en las posibilidades que para formar la convicción del juzgador ofrecen herramientas tan dispares como las que se engloban bajo los conceptos de domótica, de asistencia a la conducción, los sistemas de compra conectados a la información que recogen los *smart phones*, los relojes inteligentes con sensores biológicos que registran multitud de datos, prevén y sugieren pautas de conducta o, en general, el llamado "internet de las cosas".

Todos estos y otros muchos dispositivos "inteligentes" recogen y procesan abundantísima información sobre un gran número de individuos, y también sobre el usuario concreto del sistema, para poder construir perfiles segmentando por comportamiento y finalmente por individuos, de tal manera que, después de hacer un seguimiento acerca de cómo cada uno de ellos interactúa con esos dispositivos, es capaz de determinar, y por tanto de predecir, pautas de conducta o necesidades de una persona concreta; por ejemplo, sobre cuestiones tan concretas como las siguientes: a qué hora está en casa los días laborables, porque lo advierte el geolocalizador de sus dispositivos móviles y además porque se produce un cambio en la temperatura del interior del domicilio, que detecta el sistema de domótica; qué días tiene comensales invitados, ya que su frigorífico "inteligente" está mucho más lleno de lo habitual; el asistente a la conducción de su vehículo conoce los trayectos habituales para ir de casa al trabajo y también aquellos en los que el tráfico es más fluido, por lo que si por algún motivo decide desviarse de estos, le sugerirá el nuevo camino de vuelta a casa; su reloj "inteligente" conoce perfectamente sus principales constantes vitales en las distintas horas del día, por lo que detectará cualquier cambio en las mismas, como la presión sanguínea y el pulso extraordinariamente altos en una franja horaria que nunca se dedica a la actividad física, o nivel de ruido muy elevado en un momento que habitualmente no es de vigilia.

En definitiva, como puede suponerse a la vista de estos y otros muchísimos ejemplos que ya son realidad actualmente, todos estos sistemas IA pueden proporcionar información muy valiosa para una investigación y, en su caso, son susceptibles de llegar a acceder como medio de prueba al enjuiciamiento penal, si bien no dejan de ser datos que resultan automáticamente de la aplicación de algoritmos que gobiernan el *software* de todos y cada uno de esos sistemas.

---

adecuadas para los derechos y las libertades del interesado. En todo caso, se garantiza el derecho de obtener la intervención humana por parte del titular de tratamiento. *Vid.* Maldonato (2019, p. 403).

Como bien destaca la doctrina que se ha ocupado recientemente del tema[31], la captación y el tratamiento de datos personales generados automáticamente, la utilización en el proceso penal de estos elementos cognoscitivos de gran impacto, plantea problemas de envergadura equivalente a su creciente importancia práctica, como lo es, entre otros[32], el hecho de que la "eficiencia tecnológica" acabe siendo un criterio autosuficiente sobre la fiabilidad de la prueba[33], reemplazando así el juicio humano y dejando prácticamente sin efecto la presunción de inocencia. Además, también la igualdad de armas entre las partes puede verse afectada por el uso procesal de datos que han sido generados y tratados automáticamente, a través de algoritmos más o menos complejos creados o no específicamente para ser empleados en el marco de un proceso penal.

Concretamente en relación con la paridad de armas que ha de regir el proceso penal —acusación y defensa deben poder alegar y probar, conociendo previamente los elementos esenciales de la causa—, si no fuera posible acceder y conocer el "código fuente" del algoritmo que gobierna el sistema IA, generalmente protegido por el derecho de propiedad intelectual y creado para fines ajenos al enjuiciamiento penal, sería casi imposible cuestionar o impugnar los resultados/datos que proporciona el sistema y que se podrían utilizar como prueba en una causa penal.

Se produciría así lo que Quattrocolo (2019, p. 12) califica en este punto de "asimetría o desequilibrio cognoscitivo", ya que generalmente una parte —la pública, el Ministerio Fiscal—, tendrá acceso a la tecnología más moderna y dispondrá de medios económicos que de forma habitual no estarán al alcance del particular investigado/acusado, quien por tanto no tendrá opciones reales de rebatir o impugnar los resultados que ofrezca la "prueba algorítmica". La inaccesibilidad del código fuente o la imposibilidad de conocer características esenciales del *software* protegido, impedirán a la defensa cuestionar la exactitud y fiabilidad de la prueba incriminatoria.

---

[31] *Vid.* entre otros el trabajo de Quattrocolo (2019) y la abundante literatura anglosajona de referencia que allí se cita.

[32] Quattrocolo (2019) analiza toda esta problemática desde la perspectiva convencional del art. 8 CEDH, precepto inspirado en el concepto ya clásico de *privacy* –derecho a la vida familiar y privada-, que como es sabido puede ceder puntualmente y de manera proporcionada en pro de otros derechos fundamentales, como son la seguridad nacional o la prevención y represión del delito, siempre que haya previsión normativa y se considere "necesario en una sociedad democrática". (*Vid.* más ampliamente pp. 5 y ss).

[33] A pesar de que la aplicación de los algoritmos que se utilizan en los sistemas IA puede conducir a resultados erróneos, y no por incompetencia o malicia humana, o por un sesgo en los datos de que se nutre, sino porque en ocasiones el "aprendizaje automático", que de alguna manera se escapa del control del creador del sistema, puede derivar en conclusiones erróneas, ni siquiera previsibles por los técnicos que lo configuraron. *Vid.* más ampliamente Kearns y Roth (2020, pp. 101 y ss. y p. 262).

En definitiva, si no hay suficiente transparencia –acceso al código fuente, *inputs* y *outputs* del *software*- no podrá asegurarse la necesaria y suficiente paridad de armas entre acusación y defensa, el justo equilibrio procesal entre ambas posiciones[34]. Incluso suponiendo que se tuviera acceso a tal información, sería preciso además que las partes pudieran disponer de peritos en la materia que certificaran –o no- la fiabilidad del sistema IA y de sus resultados en ese concreto supuesto[35].

Por lo que respecta al uso de sistemas IA en la fase de valoración de los medios de prueba, entendemos que estos no podrán reemplazar al juez-persona en la valoración libre, conjunta y racional de toda la prueba lícita practicada en la causa[36], pues es este —el juez o tribunal competente para el enjuiciamiento— el que en definitiva tiene que formar su convicción más allá de toda duda razonable.

Si bien es claro que los actuales y futuros instrumentos de inteligencia artificial pueden ser de gran ayuda en esas tareas muchas veces muy complejas, no es menos cierto que tal función de valoración de la prueba practicada es estrictamente jurisdiccional, y por tanto indelegable.

En todo caso, ya los analistas de esta materia vienen haciendo referencia a algunas posibles utilidades de la IA en este punto[37]: ayuda a la determinación de la credibilidad que ha de otorgarse a las declaraciones de testigos o de las propias partes[38]; ayuda al esclarecimiento de la posible autoría o de la voluntariedad del consentimiento reflejado en la redacción de un documento, con base en el lenguaje usado o en un estilo de escritura; ayuda en la valoración de un dictamen pericial, detectando fallos o incoherencias en el mismo; reconstrucción virtual de hechos delictivos complejos, etcétera.

---

[34] También Nieva Fenoll (2018, pp. 139 y ss.) pone de relieve la importancia que la "desclasificación de los algoritmos" tiene como garantía del sistema y del derecho de defensa: no es posible elaborar una mínima estrategia de defensiva si no se puede conocer cómo decide "la máquina".

[35] Va todavía más allá Quattrocolo (2019, p. 16), pues advierte de que podría producirse una compleja y confusa "batalla entre expertos" –peritos que concluyen la fiabilidad del sistema IA, frente a otros que la niegan-, lo que obligaría al juez a erigirse en árbitro de una discusión sobre materias que toda probabilidad le resultan absolutamente ajenas e incomprensibles.

[36] Siempre interesante la lectura de los trabajos de Ferrer Beltrán; en este punto, en particular (2007).

[37] *Vid.* con mucho más detalle las explicaciones de Nieva Fenoll (2018, pp. 79 y ss.). Por su parte, Bueno de Mata (2020, esp. pp. 22 y ss.), se refiere al uso de los sistemas IA como una suerte de "prueba de indicios virtual", pues se trataría de una prueba indirecta, al no recaer sobre los hechos constitutivos del delito.

[38] Nos recuerda Nieva Fenoll (2018, p. 87), que ya la Policía española introdujo en 2017 el uso de una aplicación de inteligencia artificial, conocida como "VeriPol", que sirve para detectar palabras reveladoras del posible engaño en denuncias por el robo de teléfonos móviles, con base en datos y análisis de estadísticas previas sobre denuncias falsas.

Por lo demás, y según también se argumentó *supra*, para que estas herramientas puedan usarse como apoyo en la tarea de valoración de la prueba, en cuanto pueden influir en el resultado de la misma, y para que no se vean vulnerados los derechos de defensa y contradicción, las partes deberían poder conocer previamente los elementos y características esenciales de ese sistema IA. Es decir, debería cumplirse con el requisito de transparencia que claramente demanda el *Libro Blanco* que nos ocupa y, desde luego, la decisión final sobre la valoración de la prueba no puede atribuirse de forma exclusiva a la herramienta de inteligencia artificial.

2.2. Los algoritmos de análisis predictivo

Las herramientas de IA que sirven para realizar análisis predictivos se basan en la utilización de un gran número de datos, de carácter personal y de otros tipos —*Big Data*—, los cuales, convenientemente procesados a través de algoritmos *ad hoc*, proporcionan unos resultados que pueden servir para predecir o vaticinar el posible comportamiento futuro de una persona en distintos contextos. Así, pueden ayudar a determinar un peligro de reincidencia delictiva o de revictimización, el grado de riesgo de incumplimiento de obligaciones procesales, de las condiciones que pudieran imponerse con carácter cautelar en una causa, o en la fase de ejecución de sentencias, entre otras utilidades. Generalizadamente se conocen como *risk assessment tools 39*.

La gran mayoría de los trabajos recientes sobre la materia, cuando abordan esta concreta cuestión de los algoritmos predictivos —o justicia predictiva[40]—, hacen referencia al trascendente asunto *Eric Loomis*, y a la correlativa sentencia dictada en 2016 por la Corte de Wisconsin[41]. Este es un claro ejemplo de aplicación práctica por los tribunales norteamericanos de tal tipo de herramientas IA que se utilizan para predecir comportamientos futuros[42].

Además de servir para poner de manifiesto sus principales utilidades, el caso *Loomis* ha resultado también valioso para evidenciar los riesgos que

---

[39] Entre los trabajos más recientes puede destacarse: McKay (2020, pp. 22-39).

[40] En nuestro país, de interesante lectura las reflexiones formuladas recientemente por De la Oliva Santos (2019, pp. 30 y ss.).

[41] Pueden consultarse al respecto: De Miguel Beriain (2018); Maldonato (2019), Occhiuzzi (2019, pp. 391 y ss.), Gialuz (2019, pp. 6 y ss.); Signorato (2020, pp. 611 y ss.), Burchard (2019, pp. 1924 y ss.).

[42] Como puede suponerse, el empleo de herramientas de evaluación del riesgo no se limita al ámbito estadounidense; también pueden referirse experiencias de este tipo en Europa. Seguramente la más relevante sea la inglesa; desde 2017, la policía de Durham, en colaboración con la Universidad de Cambridge, ha puesto en marcha un sistema denominado con el acrónimo *HART —Harm Assessment Risk Tool—*, una herramienta de análisis predictivo utilizada destacadamente para decidir sobre la *diversion*, esto es, acerca la posible aplicación de un programa de rehabilitación a un detenido, como alternativa al ejercicio de la acción penal. *Vid*. más ampliamente Gialuz (2019, pp. 10 y ss.).

conlleva una utilización de los resultados de la aplicación de tales algoritmos de pronóstico sin las garantías suficientes, en particular en el orden jurisdiccional penal.

Resumiremos a continuación los principales elementos del *leading case State v. Loomis* (2016) y de la referida resolución de la Corte de Wisconsin dictada haciendo uso del algoritmo predictivo contenido en el *software* conocido como COMPAS[43]. Pondremos de relieve las principales objeciones planteadas a su utilización y, finalmente, destacaremos los aspectos que, entendemos, sí pueden ser aprovechables de estas herramientas de pronóstico, teniendo en cuenta desde luego las pautas que ya han quedado establecidas en el *Libro Blanco* sobre IA que nos ocupa, así como alguna reciente sentencia de tribunales nacionales europeos objetando el uso que se ha hecho de tales instrumentos por la Administración pública.

El programa COMPAS fue concebido para poder determinar el grado de peligrosidad de una determinada persona, el riesgo de su reincidencia delictiva, y se utiliza en algunos estados de Estados Unidos para ayudar a los jueces en la determinación de la pena, considerando una serie de datos personales y factores sociales que, según entienden los expertos que han participado en la configuración del programa, son determinantes del grado de probabilidad de que el sujeto vuelva a delinquir.

El Sr. Loomis fue enjuiciado por los delitos de receptación de vehículo y resistencia a la autoridad. En la fase de investigación preliminar se redactó un informe por las autoridades competentes que, además de suministrar al tribunal información personal sobre el imputado, incluía también una valoración del riesgo de reincidencia —*risk assessment*— resultado de aplicar COMPAS, programa que operaba con numerosos datos introducidos previamente en el sistema y con el resultado de un cuestionario de más de cien preguntas[44] que se formularon concretamente al Sr. Loomis. El software COMPAS ofreció una puntuación –*score*- sobre el riesgo de reincidencia del Sr. Loomis, y el tribunal de Wisconsin le condenó a seis años de prisión, precisamente a la vista de los resultados de la aplicación del algoritmo.

Debe llamarse la atención sobre una cuestión técnica relevante: el programa COMPAS está amparado por el secreto comercial —*trade secret*— de la empresa que lo ha diseñado, por lo que los tribunales que lo utilizan solo pueden tener acceso al concreto resultado de su aplicación en un determinado supuesto —dato numérico sobre el riesgo de reincidencia en ese caso—,

[43] Acrónimo que corresponde a *Correctional Offender Management Profilling for Alternative Sanction*.

[44] Entre otros extremos, se preguntará al detenido: cuántos de sus amigos han sido arrestados alguna vez, cuántas veces se ha mudado de casa en el último año, con qué frecuencia apenas tiene dinero, o se siente aburrido. En el siguiente enlace puede consultarse el propio cuestionario -137 preguntas- que utiliza COMPAS, y que se le entrega al sospechoso en el momento de la detención: https://www.documentcloud.org/documents/2702103-Sample-Risk-Assessment-COMPAS-CORE.

pero no les es posible conocer, ni a estos ni a las defensas, otra información acerca de los mecanismos de funcionamiento del *software* –de qué concretos datos se nutre[45], qué específicas variables pondera el algoritmo, en qué proporción, etcétera–.

Tras su condena el Sr. Loomis interpuso un recurso alegando la violación de su derecho al "proceso justo", pues no había tenido una sentencia "individualizada" basada en informaciones precisas sobre su persona, ya que COMPAS suministra datos relevantes por grupos o tipologías de individuos, y haciendo uso de algoritmos que no podían ser verificados de ninguna manera por la defensa.

En 2016 la Corte Suprema de Wisconsin dio respuesta al recurso presentado por el Sr. Loomis y declaró —por unanimidad— la legitimidad del uso del algoritmo predictivo por el tribunal para valorar el riesgo de reincidencia, si bien especificó que tal instrumento predictivo no podía ser el único elemento en que se fundamentara la sentencia condenatoria; es decir, COMPAS podrá "asistir" al órgano jurisdiccional en su decisión del caso —*sentencing*—, pero no sustituirlo. Finalmente, la Corte Suprema de los Estados Unidos confirmó esta decisión, rechazando el *writ of certiorari* que se presentó frente a la anterior resolución[46].

El sucintamente expuesto *asunto Loomis* puso enseguida sobre el tapete una serie de problemas generales que plantea la utilización de este tipo de algoritmos: la imposibilidad de conocer de qué concretos datos se nutre el sistema, de saber cómo funciona precisamente el algoritmo aplicado, en qué medida pondera este los diversos parámetros de referencia, dónde y porqué se ha colocado el umbral de riesgo bajo/medio/alto en un determinado punto de corte —*cut off*—, o si se respetan los principios de igualdad y no discriminación. Además, el hecho de que ese *software* esté protegido por el secreto comercial de la empresa que lo crea —y que lo vende para su uso por la Administración— hace que sea opaco para los operadores jurídicos, una *black box* que realiza un cálculo y ofrece un concreto resultado numérico, una puntuación que determina el contenido de una sentencia, la adopción de una medida cautelar, o la concesión de un permiso penitenciario, pero cuyo funcionamiento y banco de datos de los que se nutre no pueden ser conocidos, ni por el tribunal, ni por la defensa del investigado/acusado.

Una de las principales críticas formuladas al uso de COMPAS tiene que ver con los indicadores de riesgo que resultan de aplicar este *software* a sujetos

---

[45] Únicamente se sabe que los algoritmos que determinan el riesgo de reincidencia, en cualquier delito y en un futuro más o menos próximo, se basan en el análisis de datos sobre comportamientos de individuos previamente detenidos o condenados. Muchos de esos factores no se refieren al concreto delito objeto de la investigación, ni tampoco al posible historial criminal del sujeto en cuestión.

[46] *Vid.* las valoraciones de Gialuz (2019, p. 9), con referencia a la sentencia de la Corte Suprema de Estados Unidos *Loomis v. Wisconsin* (2017).

investigados o acusados que son afroamericanos[47]. Si uno de los datos a tener en cuenta es el número de arrestos previos de personas con esta característica personal, resulta que, por el mero hecho de ser afroamericano, el nivel de riesgo que indicará el algoritmo se multiplicará notablemente –resultará *high risk*-, ya que en Estados Unidos la cifra de detenciones de personas afroamericanas, y de otras minorías, es muy elevada. Sin embargo, el hecho de que se produzca ese mayor número de contactos entre la policía y los sujetos pertenecientes a minorías étnicas, no significa necesariamente una mayor propensión delictiva de estos, sino más bien que existe un control policial mucho más estricto sobre los sujetos pertenecientes a esos grupos raciales minoritarios en Estados Unidos.

Utilizar el dato del porcentaje de detenciones a sujetos de una determinada etnia como medida del riesgo de reincidencia, hace que sea mucho más difícil, *v. gr.*, la defensa de un afroamericano, pues directamente será "puntuado" como individuo de alto riesgo a estos efectos. Se trata por tanto de una desviación —*bias*— del sistema, que necesariamente habría de tenerse en cuenta.

En consecuencia, a pesar de la aparente neutralidad y objetividad de este tipo de algoritmos predictivos[48], tal y como está concebido al menos el *software* COMPAS, se puede concluir que su utilización vulnera el derecho de defensa, la igualdad de partes y la necesaria transparencia en los sistemas utilizados en la adopción de decisiones judiciales; por mucho que, como indicó la Corte Suprema de Wisconsin, no pueda ser el único elemento en que se fundamente la sentencia condenatoria[49].

Si nos colocamos ya en el concreto ámbito UE, deberíamos tener en cuenta además que, si se pretende la utilización de *risk assessment tools* en los procesos penales, sería de aplicación la referida Directiva (UE) 2016/680[50], que

---

[47] *Vid.* más ampliamente las explicaciones de Maldonato (2019, pp. 406 y ss.) con referencia a la doctrina norteamericana que ha analizado la cuestión.

[48] De imprescindible consulta a este respecto, el trabajo monográfico de Kearns y Roth (2020, pp. 101 y ss.), en particular acerca de los conceptos de imparcialidad, exactitud, equidad y paridad estadística del algoritmo. Sobre la "aparente neutralidad" de la IA, *vid.* también Ubertis (2020, pp. 2 y ss.).

[49] Interesante la consulta del documento publicado por el *Pretrial Justice Institute* (2020), en el que la citada institución concluye que "*We now see that pretrial risk assessment tools, designed to predict an individual's appearance in court without a new arrest, can no longer be a part of our solution for building equitable pretrial justice systems. Regardless of their science, brand, or age, these tools are derived from data reflecting structural racism and institutional inequity that impact our court and law enforcement policies and practices. Use of that data then deepens the inequity*". "*We have consistently opposed the use of pretrial risk assessment tools to make detention decisions. We now expand that to oppose their use to determine restrictions placed on a person's pretrial liberty (reporting visits, electronic monitoring, curfews, drug testing, etc.)*".

[50] Relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales por parte de las autoridades competentes para fines de prevención, investigación, detección o enjuiciamiento de infracciones penales o de ejecución de

específicamente en su art. 11, apdo. 1º, establece que: "Los Estados miembros dispondrán la *prohibición de las decisiones basadas únicamente en un tratamiento automatizado*, incluida la elaboración de perfiles, que produzcan efectos jurídicos negativos para el interesado o le afecten significativamente, salvo que estén autorizadas por el Derecho de la Unión o del Estado miembro a la que esté sujeto el responsable del tratamiento y que establezca medidas adecuadas para salvaguardar los derechos y libertades del interesado, al menos el derecho a obtener la intervención humana por parte del responsable del tratamiento[51]".

En fin, como pone de relieve Gialuz (2019, p. 17), junto con la obligatoria e imprescindible valoración *humana* de todos los datos que obran en la causa, el *output* resultante de la aplicación de sistemas IA solo podrá ser considerado como un "mero indicio", y para que pueda ser tenido en cuenta en la decisión judicial, debería siempre estar corroborado por otros elementos efectivamente probatorios.

Esta interpretación se confirma por la excepción a la regla general —prohibición de decisiones basadas únicamente en un tratamiento automatizado de datos—, que se encuentra también en la citada Directiva: estas solo se admitirán si está prevista una tutela suficiente de los derechos con intervención humana.

Además, si estuviera en juego la libertad personal del sujeto imputado, debería tenerse en cuenta lo previsto en los arts. 5 CEDH y 6 CDFUE: el interesado tiene en todo caso derecho a que sobre su *status* se pronuncie un juez en persona, quien deberá valorar también otros elementos probatorios más allá del *output* resultante del *risk assessment tool*.

Por tanto, aunque excepcionalmente el derecho de la Unión y/o de los Estados miembros estableciera la posibilidad de delegar en un *software* o sistema IA la adopción de una decisión en este ámbito, siempre debería ser posible recurrir la decisión ante un juez/tribunal persona física[52].

---

sanciones penales, y a la libre circulación de dichos datos y por la que se deroga la Decisión Marco 2008/977/JAI del Consejo. *Vid.* art.1, párr.1: "La presente Directiva establece las normas relativas a la protección de las personas físicas en lo que respecta al tratamiento de los datos personales por parte de las autoridades competentes, con fines de prevención, investigación, detección o enjuiciamiento de infracciones penales o de ejecución de sanciones penales, incluidas la protección y la prevención frente a las amenazas contra la seguridad pública".

[51] Muy interesante en este punto el Informe publicado por Fair Trials (2020), *Regulating AI for Use in Criminal Justice Systems in the EU*, en el que se llama la atención sobre la posible amplitud de las excepciones a la prohibición de decisiones automatizadas, pues basta con que lo autorice el Derecho de la Unión o de un Estado miembro; no queda claro con qué salvaguardas esto sería posible, ni qué ha de entenderse por "intervención humana". *Vid.* pp. 1-6 del referido Informe.

[52] En este sentido, con referencia a esta insustituible función del "*giudice in carne*", *vid.* Gialuz (2019, p. 18).

A mayor abundamiento, el apdo. 2º de este mismo art. 11 de la Directiva (UE) 2016/680, en relación con el art. 10, indica que tales decisiones basadas en un tratamiento automatizado de datos que puedan producir efectos jurídicos negativos o afecten significativamente a los individuos, no podrán *como regla general* basarse en categorías de datos personales "que revelen el origen étnico o racial, las opiniones políticas, las convicciones religiosas o filosóficas, o la afiliación sindical, así como el tratamiento de datos genéticos, datos biométricos dirigidos a identificar de manera unívoca a una persona física, datos relativos a la salud o a la vida sexual o las orientaciones sexuales de una persona física", *salvo que* se hayan tomado las medidas adecuadas para salvaguardar los derechos, libertades e intereses legítimos del interesado. Es más, el apdo. 3º de este mismo art. 11 prohíbe expresamente la elaboración de perfiles que dé lugar a una discriminación de las personas físicas con base en las categorías de datos antes enunciadas.

Por otro lado, más allá de la mayor o menor fiabilidad de estas herramientas predictivas, su uso suscita a nuestro juicio otros problemas de fondo no menores, como por ejemplo que el concreto montante de una condena pueda depender, no solo del hecho delictivo cometido y de la culpabilidad del autor, o de conductas ilícitas del condenado en el pasado —reincidencia—, sino de una previsión acerca de potenciales comportamientos futuros de esa persona. Estaríamos transitando de un Derecho penal del hecho, a un inaceptable Derecho penal del perfil del autor[53].

En relación con el uso judicial de estas *risk assessment tools*, es destacable también una sentencia dictada ya en el contexto europeo, concretamente por el Tribunal de Distrito de la Haya[54] el 5 de febrero de 2020[55], en la que se declara que el sistema algorítmico utilizado por los ayuntamientos de los municipios holandeses para valorar y prevenir el riesgo de defraudación tributaria o a la Seguridad Social por parte de sus ciudadanos, vulnera especialmente el derecho a la vida privada protegido en el art. 8 CEDH[56], pues no cumple con los imprescindibles requisitos de proporcionalidad y transparencia.

---

[53] El análisis de esta cuestión excede con creces de los objetivos de este trabajo. Sobre el particular, *vid.* las reflexiones de Feijoo Sánchez (2007, esp. p. 13): no es admisible una determinación de la pena basada en pronósticos; es imposible enunciar un juicio de pronóstico mínimamente seguro respecto a la conducta futura del sujeto.

[54] Resolviendo la demanda interpuesta por una serie de agrupaciones entre las que se contaban la "Plataforma para la Protección de los Derechos Civiles", el "Comité Jurídico Holandés para los Derechos Humanos" y el "Consejo Nacional de Consumidores".

[55] Esta sentencia ha sido reseñada en nuestro país por Cotino Hueso (2020), y anteriormente por Fernández (2020). *Vid.* también las valoraciones de Battaglini (2020).

[56] "1. Toda persona tiene derecho al respeto de su vida privada y familiar, de su domicilio y de su correspondencia. 2. No podrá haber injerencia de la autoridad pública en el ejercicio de este derecho sino en tanto en cuanto esta injerencia esté prevista por la ley y constituya una medida que, en una sociedad democrática, sea necesaria para la seguridad nacional, la seguridad pública, el bienestar económico del país, la defensa del orden y la prevención de las infracciones penales, la protección de la salud o de la moral, o la protección de los derechos y las libertades de los demás".

El sistema conocido como *SyRI*[57] —Sistema de Indicación del Riesgo— asigna a las personas físicas y jurídicas un determinado nivel de riesgo de defraudación o de uso ilegal de fondos públicos, en función de un gran número de datos y parámetros, personales, laborales, económicos, fiscales, antecedentes de incumplimientos, etc., que se procesan por medio del algoritmo en cuestión, el cual finalmente identifica factores de riesgo de un determinado sujeto o persona jurídica.

La norma que se encontraba en la base del funcionamiento y uso gubernamental del sistema *SyRI* fue impugnada por varias asociaciones holandesas de defensa de derechos civiles, por vulneración del art. 8, párrafo 2º CEDH: la injerencia del referido sistema en el derecho al respeto a la vida privada no era necesaria, ni proporcional para la defensa del interés social, es decir, al propósito previsto en la ley.

El referido Tribunal de Distrito de la Haya concluyó que la legislación holandesa no cumplía con el "justo equilibrio" que según el CEDH ha de existir entre el interés social en evitar el fraude o el uso indebido de fondos públicos de una forma más eficaz[58], y el respeto a la vida privada de los ciudadanos. Además, se tuvieron en consideración también los principios fundamentales de la protección de datos en el derecho UE, en particular los de transparencia y verificabilidad del sistema de determinación del riesgo, y se concluyó que el gobierno holandés no había hecho público el tipo de algoritmos utilizados para determinar ese nivel de riesgo, ni el método de análisis predictivo empleado; tampoco las personas cuyos datos se procesaban sabían que esa información personal se estaba utilizando por el referido sistema *SyRI*.

En fin, el Tribunal concluyó que la legislación que sustentaba el uso de *SyRI* no proporcionaba la información suficiente para conocer qué datos objetivos justificaban la conclusión de que una persona física o jurídica tenía un "riesgo alto" en los términos expuestos. Faltaba por tanto la necesaria transparencia; no era posible verificar el diseño del mecanismo de decisión empleado por el algoritmo en cuestión, por lo que existía un peligro de que se produjeran sesgos o desviaciones en el sistema. Además, se consideró que no existía una intervención humana significativa que evitase los peligros de desviación o error en el mismo.

En cualquier caso, y aunque el Tribunal holandés abordó la cuestión desde la perspectiva del derecho a la privacidad del art. 8 CEDH, este no deja de llamar la atención sobre otro aspecto a nuestro juicio de gran relevancia:

---

[57] *Systeem Risicoindicatie*, conocido por el acrónimo SyRI, sistema regulado por una Ley de 2013 y desarrollado por un Decreto gubernativo del año 2014.

[58] Finalidades que se consideran legítimas. El Tribunal entiende que, en efecto, se deben utilizar esas nuevas tecnologías para prevenir y combatir el fraude, que es una necesidad social evidente, pues afecta a la integridad del sistema económico y a la confianza en las instituciones financieras del país. El uso de estos sistemas que permiten definir "perfiles de riesgo" no es *per se* contrario al art. 8.2 CEDH, según el Tribunal de Distrito de La Haya.

"es difícil comprender cómo un interesado podría defenderse contra el hecho de que se haya presentado un informe de riesgos sobre él o ella" –apdo. 6.90 de la sentencia-. Como destaca Cotino (2020, p. 5), con referencia también al caso *Loomis,* además de a la propia sentencia holandesa, la afectación de las garantías del debido proceso son determinantes en la utilización de herramientas de este tipo, por la opacidad del sistema, porque el modelo y los indicadores de riesgo utilizados son secretos, y por tanto no pueden ser conocidos por los interesados.

Por lo demás, si bien estas objeciones y conclusiones que acabamos de resumir se sitúan en el contexto del uso de sistemas IA en el Derecho administrativo sancionador, entendemos que la gran mayoría de ellas son perfectamente extrapolables y aplicables también al ámbito de la toma de decisiones de naturaleza jurisdiccional cuando se pretende hacer uso de tales algoritmos predictivos.

CONCLUSIONES

A nuestro juicio, los sistemas IA basados en el correcto procesamiento de *Big Data* podrán ser una gran ayuda en el ejercicio de la función jurisdiccional[59]. Ya se utilizan desde hace tiempo en otros ámbitos de la vida pública y privada, donde generalmente resultan valiosos.

A la luz del contenido del *Libro Blanco* sobre la materia aprobado por la Comisión Europea y del resto de argumentos puestos de relieve en los epígrafes precedentes, consideramos que tales instrumentos podrían ser empleados por el aparato judicial, en particular por lo que respecta al orden penal, si se pueden garantizar al menos los siguientes requisitos y presupuestos:

1) Previsión normativa del ámbito de aplicación y forma de utilización de estos sistemas IA. El posible uso de este tipo de herramientas de apoyo a la función jurisdiccional debería estar de alguna manera expresamente previsto en los correspondientes preceptos legales relativos, *v. gr.,* a la determinación judicial del riesgo de reiteración delictiva, de fuga u ocultamiento, circunstancias que como es sabido son decisivas para la adopción de medidas cautelares personales[60] *-vid*. 503 LECrim-, a la fijación del montante de cauciones y

[59] En definitiva, como explican Kearns y Roth (2020, p. 261), también la toma de decisiones por personas son algorítmicas, pues se basan en la lógica y en la experiencia previa. Si pudiéramos describir esos procesos con suficiente exactitud, lo haríamos por medio de algoritmos. Por tanto, no se trata de decidir si podemos usar o no algoritmos en las decisiones judiciales, sino de establecer y utilizar algoritmos definidos correctamente y con suficiente precisión.

[60] Es complicado admitir juicios predictivos de peligrosidad de un sujeto cuando todavía tiene la presunción de inocencia a su favor; no obstante, puesto que las medidas cautelares son necesarias y es aceptado que uno de sus fundamentos es evitar el riesgo de reiteración delictiva, no tiene mucho sentido limitar los instrumentos cognitivos del juez. Por lo demás, como destaca Gialuz (2019), si admitimos las "máxi-

fianzas —*vid*. art. 505.4 LECrim—, a los supuestos, condiciones y circunstancias en que se puede acordar la suspensión de la ejecución de la pena privativa de libertad —*vid*. art. 80 CP—, o a la concesión de permisos penitenciarios con base en un "informe de pronóstico" y un "juicio de probabilidad sobre el comportamiento futuro" —*vid*. art. 67 LOGP[61]—.

Y lo mismo cabría decir de los instrumentos de ayuda a la valoración judicial de determinados medios de prueba. Esos sistemas IA podrían ser propuestos por las partes, como complemento al medio de prueba que ellas mismas solicitasen practicar, o incluso ser acordados como diligencias probatorias por el Tribunal juzgador en aplicación de lo ya dispuesto en el apdo. 2º del art. 729 LECrim, precepto que se refiere a aquellas diligencias de prueba que, no habiendo sido propuestas por ninguna de las partes, considere necesarias el Tribunal para la comprobación de cualquiera de los hechos que hayan sido objeto de los escritos de calificación —*v. gr.*: una reconstrucción de hechos a través de sistemas IA—, o incluso en virtud del apdo. 3º del mismo artículo, que se refiere a la posibilidad de que el Tribunal admita diligencias de prueba de cualquiera clase que en el acto ofrecieran las partes para acreditar alguna circunstancia que pudiera influir en el valor probatorio de la declaración de un testigo –*v. gr.*: sistemas IA que ayuden a determinar su credibilidad-.

Entendemos que, además de la referencia legal expresa a la posibilidad de usar estas herramientas IA en el ejercicio de la función jurisdiccional, debería existir una regulación suficiente en otros cuerpos normativos, mejor fuera de la LECrim, que permitiera su modificación y actualización siempre que fuera necesario[62], igual que sucede con los que en general se califican de

---

mas de experiencia" que se utilizan en el razonamiento judicial para efectuar juicios predictivos en relación con medidas cautelares, que no dejan de estar basados en la generalización de experiencias con *otros* sujetos, ¿por qué no se va a poder usar las herramientas IA con las debidas garantías?, se cuestiona el autor. Estas almacenan más datos aún que la experiencia propia del juez que tiene que decidir; si el algoritmo es transparente y no discriminatorio, puede ser una herramienta de valiosa ayuda al juez en la toma de decisiones. No se deben excluir instrumentos potencialmente útiles para obtener una decisión más acertada y menos arbitraria, aunque la última palabra siempre será del juez. *Vid.* más ampliamente, pp. 19 y ss.

[61] Gascón Inchausti (2019) ha puesto de relieve la utilidad de estos sistemas IA para ayudar a concretar el riesgo de reincidencia, que es un factor determinante del sentido de las decisiones que se puedan adoptar en relación con estas cuestiones. *Vid.* pp. 191 y ss., esp. p. 202.

[62] Se refieren algunos analistas en este punto, pese a la reticencia que suscita entre otros, a la pertinencia de modelos de "regulación líquida"; es decir, de un nuevo tipo de legislación que pueda revisarse regularmente —"reglas biodegradables"—, una normativa que se adapte y siga el paso a estos desarrollos tecnológicos tan cambiantes. Las fórmulas regulatorias sobre estos extremos han de ser necesariamente más dinámicas, con remisiones a órganos capacitados técnicamente y con suficiente legitimación. Solo así el Derecho puede desplegar una eficacia razonable ante los nuevos retos. *Vid.* más ampliamente Cotino Hueso (2019, esp. pp. 21 y ss.).

"Aspectos accesorios de las actuaciones judiciales", que caen bajo el marco de la potestad reglamentaria del Pleno del Consejo General del Poder Judicial —CGPJ—. Tengamos en cuenta en este punto el Acuerdo de 15 de septiembre de 2005 del CGPJ, que aprueba el *Reglamento 1/2005, de aspectos accesorios de las actuaciones judiciales*, en el que ya se contiene, entre otros extremos, el procedimiento de aprobación de los programas, aplicaciones y sistemas informáticos de la Administración de Justicia, o la gestión de ficheros de datos automatizados de carácter personal que se encuentran bajo la responsabilidad de los órganos jurisdiccionales y del CGPJ.

2) Control previo de la legalidad/admisibilidad del concreto algoritmo y del tratamiento de los datos que procesa el sistema IA. Supervisión de su aplicación y funcionamiento. Ya se trate de un algoritmo elaborado por una entidad privada, o bien por una de carácter público —esto último preferible[63], por el tipo de datos personales que se van a manejar, así como para garantizar la estabilidad en el servicio de seguimiento y actualización—, de forma previa a su utilización por los órganos jurisdiccionales de nuestro país, sería preciso que todo el sistema IA, incluida la selección de datos de los que se va a nutrir[64], fuera aprobado y supervisado por un organismo público, que bien podría ser una Comisión *65* enmarcada en el Consejo General del Poder Judicial.

Además, en el caso de que durante su utilización los juzgados y tribunales pusieran de manifiesto algún sesgo o desviación en los resultados que ofreciera el sistema o un anquilosamiento del mismo, deberían comunicarlo a tal Comisión a fin de que se realizaran los ajustes, actualizaciones y correcciones que fueran necesarias.

Naturalmente, operarían también en este contexto de elaboración y utilización de los sistemas IA todas las medidas legal y reglamentariamente previstas en nuestro país para la protección de los datos de carácter personal que se pudieran emplear en el funcionamiento de estos sistemas IA, además del resto de derechos y libertades fundamentales que garantiza nuestro or-

---

[63] Sobre la pertinencia de que sea una "agencia pública" quien controle la fiabilidad de estos instrumentos IA en el proceso, *vid.* también Gialuz (2019, p. 21).

[64] Como destaca Cotino Hueso (2019, pp. 12 y 13), las afirmaciones sobre la neutralidad, objetividad y precisión de los datos son engañosas, pues siempre hay un proceso de "limpieza de datos" inherentemente subjetivo y, desde luego, discriminaciones indirectas no intencionadas y/o errores en el uso de macrodatos que pueden afectar a la igualdad y otros principios constitucionales. Sobre los conceptos de imparcialidad del algoritmo y paridad estadística, *vid.* más ampliamente Kearns y Roth (2020, pp. 101 y ss.).

[65] *Vid.* arts. 98 y ss. del citado Reglamento 1/2005, del Pleno del Consejo, sobre aspectos accesorios de las actuaciones judiciales. En particular, las referencias en que allí se contienen a la Comisión de Informática Judicial del CGPJ, quien podría asumir también este tipo de funciones de propuesta de aprobación de sistemas IA por el Pleno del CGPJ, en el marco de sus obligaciones de control de programas y aplicaciones informáticas que se puedan usar en la Administración de Justicia.

denamiento, destacadamente las distintas prohibiciones de discriminación legalmente garantizadas.

Por otro lado, este tipo de herramientas IA también pueden ser empleadas por la Policía, o incluso por la Fiscalía cuando dirige sus propias investigaciones, y desde luego resultarán muy útiles para determinar sectores de riesgo sobre los que conviene actuar en cada momento[66]. Pensemos sin ir más lejos en la ciberdelincuencia; la inteligencia artificial puede ser de gran ayuda, o incluso a veces ser la única herramienta realmente eficaz para detectar la peligrosidad de determinadas conductas en la red, lo que a su vez permitiría dirigir los recursos policiales y de investigación —siempre escasos— con la mayor efectividad posible[67].

3) Necesaria intervención humana —del juez— en la adopción de la decisión final —resolución jurisdiccional en sentido estricto—. Posibilidad de "acción judicial efectiva" contra la decisión judicial basada en sistemas IA. Según se viene argumentando en las líneas precedentes, entendemos que las herramientas IA han de servir como instrumentos de apoyo/asesoramiento en la toma de decisiones judiciales, como pueden serlo el informe de un forense, el de la Junta de Tratamiento de un centro penitenciario, o una pericia informática complementaria —valgan los símiles—. En todo caso, la decisión final no puede basarse exclusivamente en lo que resulte de la aplicación del algoritmo; del mismo modo que no sería aceptable que el juez delegase en un perito la resolución sobre la pretensión punitiva, por muy complejo que pueda ser el objeto de la pericia.

Igual que el juez puede desmarcarse en su decisión final de la valoración de los hechos que propone un determinado informe pericial, podrá también discrepar en sus resoluciones del resultado sugerido por la aplicación del algoritmo.

---

[66] *Vid. Libro Blanco*, p. 2: "Las herramientas de inteligencia artificial pueden ofrecer una oportunidad para proteger mejor a los ciudadanos de la UE de la delincuencia y los actos de terrorismo. Este tipo de herramientas podrían, por ejemplo, ayudar a detectar propaganda terrorista en línea, descubrir transacciones sospechosas en la venta de productos peligrosos, detectar objetos peligrosos ocultos o productos y sustancias ilícitos, ofrecer asistencia a los ciudadanos en situaciones de emergencia y servir de orientación al personal de primera intervención".

[67] *Vid.* Parolli y Sellaroli (2019, p. 70), quienes llaman la atención sobre algunos posibles efectos indeseables. Cierto es –afirman- que la IA puede servir para racionalizar el uso de los insuficientes recursos personales y materiales de los que se dispone para la investigación de los delitos en general, pero también pueden llegar a producirse ciertas desviaciones en los datos sobre peligrosidad de determinados sectores. Así, si en un determinado momento se aprecia que se cometen más delitos en algunas concretas zonas de la ciudad o por los que allí viven, esto provocará un aumento de la vigilancia policial, y por tanto al descubrimiento de más delitos, lo que originará un incremento mayor aún de índice de peligrosidad del lugar y de los que lo habitan. Estas desviaciones han de ser consideradas también.

Será preciso además hacer frente en este punto a un riesgo cierto, que no es otro que el peligro de que el juez tienda a acomodar sus decisiones, sin apenas más valoraciones adicionales, a lo que resulte de aplicar el sistema IA pertinente al asunto que debe resolver[68].

Entendemos que una resolución judicial fundada exclusivamente, de forma automática, en el resultado de la aplicación de un algoritmo —al que incluso se remitiera de forma expresa—, no podría considerarse una decisión respetuosa con el derecho a la motivación de las resoluciones judiciales[69]. De otro lado, si el juez estimara que debe apartarse en su decisión final del resultado "recomendado" por la aplicación del algoritmo, también deberá motivarlo en los fundamentos de su resolución.

La suficientemente razonada motivación de la resolución judicial jugará entonces un papel determinante en este punto, pues permitirá a las partes conocer el fundamento fáctico y jurídico de tal decisión y, en su caso, ejercitar de manera efectiva el derecho a recurrir las resoluciones judiciales perjudiciales que sean impugnables. Se vería entonces satisfecho el derecho que se recoge en el propio *Libro Blanco* a la "acción judicial efectiva" contra los resultados de la aplicación de sistemas IA; "acción" que se materializaría a través de los posibles recursos contra las decisiones judiciales adoptadas utilizando sistemas IA.

Esta cuestión enlaza desde luego también con la necesidad de que las partes puedan conocer qué concretos factores pondera el algoritmo, cómo fun-

---

[68] Que ya destacara, entre otros Gascón Inchausti (2019, p. 204): debido a la dificultad intrínseca de la toma de decisiones en ciertos escenarios complejos, es comprensible la tendencia humana a tratar de delegar esas decisiones o parte de ellas en un tercero —perito— o bien en una "máquina", que gozaría de una cierta "apariencia de mejor condición", al menos por su apariencia de mayor objetividad, lo que puede conducir a que esos sistemas tengan una repercusión sobre el sentido de la decisión que, si bien no puede decirse que sea "automático", sería desde luego muy determinante. Afirma también Gascón (2019, p. 205), que en tales casos se produciría un claro peligro para la efectividad del derecho de defensa, e incluso el riesgo de cierta inversión de la carga probatoria, pues no será sencillo cuestionar en un caso concreto el fundamento científico, metodológico o empírico del sistema de inteligencia artificial y la fiabilidad de sus resultados. Insiste igualmente Gialuz (2019, pp. 19 y ss.) en que el juez deberá evitar lo que se denomina *"automation complacency"* o *"automation bias"*, es decir, la tendencia humana a ignorar o a no buscar información adicional que pueda contradecir la solución generada por el ordenador, que es aceptada como "la correcta". *Vid.* también Ubertis (2020, p. 12), y las advertencias que el autor realiza sobre el riesgo del "mito tecnológico" de la IA, que podría inducir al juez a no desmarcarse del resultado ofrecido por la máquina.

[69] Hay quien ya alude a la necesidad de reconocer expresamente la existencia de un nuevo derecho fundamental, denominado "derecho a las decisiones no basadas exclusivamente en tratamientos automatizados", que derivaría de la dignidad humana, como derecho de un hombre a ser juzgado por otro hombre. *Vid.* más ampliamente, Signorato (2020, p. 613).

ciona el sistema IA, de qué manera se obtiene un determinado resultado y, naturalmente, de qué datos se puede nutrir este.

4) Transparencia y trazabilidad de los sistemas IA. Seguridad y verificabilidad de los datos empleados, sin sesgos ni discriminaciones. Estos requisitos para el uso de los algoritmos que están en la base de los sistemas IA son abordados con detalle en el *Libro Blanco* que nos ocupa, lo que es buena muestra de su importancia, más si cabe en el sector de "riesgo elevado" que es la administración de justicia.

Ya hemos puesto de relieve anteriormente cómo las principales objeciones formuladas al programa predictivo COMPAS tienen que ver precisamente con estos requisitos que se entiende han de reunir los sistemas IA que se vayan a utilizar en este ámbito. También hemos mencionado la normativa UE que debería ser tenida en cuenta en relación con el tratamiento de datos personales.

Por tanto, no se debería autorizar el empleo de herramientas que puedan ser determinantes del sentido de una resolución judicial, si las partes no pueden conocer los elementos que integrarán los algoritmos y el propio funcionamiento del sistema de IA. No puede aceptarse en este ámbito el uso de las llamadas *"black box"*[70], por lo que no será posible ampararse en la protección del derecho al secreto empresarial del creador intelectual de la herramienta IA para impedir el conocimiento del algoritmo, del peso ponderado que se otorga a cada factor variable, o del tipo de datos de que este se nutre[71].

Coincidimos en este punto con Nieva Fenoll (2018, p. 143) cuando concluye que los tiempos modernos exigen un nuevo concepto de publicidad procesal, adaptando esta al empleo de las nuevas herramientas tecnológicas; solo de esta forma se podrán seguir garantizando los esenciales principios de defensa y contradicción.

---

[70] Como destaca Fair Trials (2020), la transparencia es un aspecto fundamental del proceso contradictorio y del "debido proceso", que se ha visto reforzada con la Directiva 2012/13/EU sobre el derecho a la información, la cual exige a los Estados UE que faciliten el acceso a las defensas a todo el material probatorio en posesión de las autoridades competentes. Y los posibles intereses comerciales de las empresas que diseñan y venden los sistemas IA no son justificación suficiente para la *non-disclosure*. *Vid.* más ampliamente el citado Informe de *Fais Trials*, p. 29.

[71] Gascón Inchausti (2019, pp. 203 y ss.), destaca que estos problemas relativos a la necesaria transparencia y trazabilidad del algoritmo pueden ser efectivamente paliados por el legislador, quien podría forzar el acceso a esa información o supeditar un acceso proporcionado para permitir su empleo por las autoridades públicas. Refiere el citado autor ejemplos nacionales que funcionan desde hace años, como los sistemas de determinación del riesgo de reiteración de delitos de violencia de género –conocido como *VioGen*-, que se basan en una serie de variables ponderadas y de protocolos para la valoración policial del nivel de riesgo para la víctima, sistemas que no son secretos y que se utilizan para acordar medidas cautelares para la seguridad de aquella, instrumentos aprobados por Instrucciones de la Secretaría de Estado de Seguridad.

Además, el hecho de que se pueda recurrir una decisión basada en el resultado de la utilización de un sistema IA, cuyo funcionamiento se conoce, puede posibilitar la propia evolución/corrección del algoritmo y de los datos de que se nutre, en caso de que se evidencie que hay sesgos indeseables, carencias o excesos en los factores a ponderar, o si se produce un anquilosamiento de los mismos con el paso de los años.

5) Necesidad de una regulación armonizada sobre el uso de algoritmos por las autoridades jurisdiccionales en el ámbito del "espacio de libertad, seguridad y justicia" de la Unión Europea. Aunque el *Libro Blanco* que nos ocupa se refiere destacadamente a la importancia de arbitrar una acción política y normativa conjunta en el contexto UE, que pueda generar un *ecosistema de confianza con un enfoque antropocéntrico* en materia de inteligencia artificial, *para poder así mantener un nivel adecuado de competitividad en el exigente contexto mundial, evitando una fragmentación del mercado único*[72], entendemos que, a pesar de que no se haga referencia a otras finalidades, tal acción normativa conjunta sobre IA resulta a nuestro juicio decisiva también para poder alcanzar satisfactoriamente otros objetivos que, desde luego, son fundamentales para la consolidación y el buen funcionamiento del "espacio de libertad, seguridad y justicia" de la Unión.

No olvidemos que en materia de cooperación judicial transfronteriza rige el principio de reconocimiento mutuo de resoluciones judiciales[73], y que presupuesto fundamental para su eficacia es, además de la confianza mutua entre las autoridades implicadas, la previa armonización/aproximación de las regulaciones nacionales sobre la materia objeto de la petición de cooperación[74].

Así, por mencionar varios ejemplos que pueden resultar ilustrativos, si una autoridad judicial competente solicita el traslado de una persona condenada a otro Estado de la Unión en el que se estima que esta puede tener más posibilidades de reinserción social –reconocimiento mutuo de sentencia a pena privativa de libertad[75]-, o si se requiere que tenga eficacia transfronteriza una orden europea de protección de víctimas[76], o si se pretende que

---

[72] *Vid. Libro Blanco*, esp. p. 3.

[73] *Vid.* arts. 81 y ss. del Tratado de Funcionamiento de la Unión Europea.

[74] Desde la perspectiva procesal, véase Arangüena, De Hoyos y Rodríguez-Medel (dirs.) (2015), Arangüena y De Hoyos (2018), Arangüena, De Hoyos y Hernández (2020), y De Hoyos Sancho (2019c).

[75] *Vid.* Decisión Marco 2008/909, del Consejo, de 27 de noviembre de 2008, y la transposición en España en el Título III de la Ley 23/2014, de 20 de noviembre, de reconocimiento mutuo de resoluciones penales en la Unión Europea, arts. 63 y ss. Más ampliamente, De Hoyos Sancho (2015a, pp. 107 y ss.).

[76] *Vid.* Directiva 2011/99/UE y Reglamento UE 606/2013, respectivamente sobre reconocimiento de órdenes de protección dictadas en procesos penales y en materia civil, así como el Título VI de la citada Ley de reconocimiento mutuo de resoluciones penales, arts. 130 y ss. Un estudio de dicho instrumento puede encontrarse en De Hoyos Sancho (2015b, pp. 271 y ss.).

tenga validez una prueba penal recabada en otro Estado de la Unión —orden europea de investigación—[77], será preciso que, si en la recogida de datos personales y en la toma de decisiones judiciales ha intervenido un sistema IA, cosa que podrá ocurrir no tardando mucho[78], este cumpla con los requisitos de transparencia, seguridad, trazabilidad, no discriminación, acción judicial efectiva…, y todos los demás a que venimos haciendo referencia a lo largo de las líneas precedentes. En otro caso, la eficacia de la cooperación judicial transfronteriza basada en el reconocimiento mutuo de resoluciones se podría ver frustrada por la vulneración de derechos y libertades fundamentales, precisamente a través del empleo de sistemas IA que han resultado determinantes del sentido de las decisiones adoptadas, pero que no reunían las garantías imprescindibles.

FUENTES CITADAS

A. Bibliografía

Arangüena, C., De Hoyos, M., Rodríguez-Medel, C. (dirs.) (2015). *Reconocimiento mutuo de resoluciones penales en la Unión* Europea. Aranzadi.

Arangüena, C. y De Hoyos, M. (dirs.) (2018). *Garantías procesales de investigados y acusados. Situación actual en el ámbito de la Unión Europea.* Tirant lo Blanch.

Arangüena, C., De Hoyos, M. y Hernández, A. (eds.) (2020). *Procedural Safeguards for Suspects and Accused Persons in Criminal Proceedings.* Springer.

Armenta Deu, T. (2019). Orden europea de investigación y exclusión probatoria. En González Cano, I. (dir.), *Orden europea de investigación y prueba transfronteriza en la Unión Europea* (pp. 767-796). Tirant lo Blanch.

Barona Vilar, S. (2019): Inteligencia artificial o la algoritmización de la vida y de la justicia: ¿solución o problema? *Revista Boliviana de Derecho,* 28, 18-49.

Battaglini, M. (2020). Sentencia histórica del Tribunal de la Haya anulando la elaboración de perfiles para el fraude de la Seguridad Social (SyRI). https://www.linkedin.com/pulse/sentencia-histórica-del-tribunal-de-la-haya-anulando-manuela/.

Bueno de Mata, F. (2020). Macrodatos, inteligencia artificial y proceso: luces y sombras. *Revista General de Derecho Procesal,* 51.

Burchard, Ch. (2019). L'intelligenza artificiale come fine del Diritto penale? Sulla trasformazione algorítmica della societá. *Rivista italiana di diritto e procedura penale*, 4 (62), 1909-1942.

---

[77] González Cano (2019), y en particular los trabajos, en relación con la admisión de la prueba transfronteriza, Armenta Deu (2019), y Laro González (2019). También, De Hoyos Sancho (2019b).

[78] *V. gr.:* utilización de un sistema IA que ayude a determinar las posibilidades de reinserción social de un condenado en un determinado país, en función de parámetros que se ponderasen a través de un algoritmo, como pueden ser, entre otros, vínculos familiares, sociales, lingüísticos, culturales, situación profesional, económica, sanitaria, etc., etc.; o la aplicación de sistemas IA para determinar el riesgo de revictimización en supuestos de violencia de género; o la aceptación de fuentes de prueba obtenidas en otros Estados y que son resultado de la utilización de algoritmos y *big data*, a los que hemos hecho referencia *supra*.

Cotino Hueso, L. (2019). Riesgos e impactos del *big data*, la inteligencia artificial y la robótica. Enfoques, modelos y principios de la respuesta del derecho. *Revista General de Derecho Administrativo*, 50.

Cotino Hueso, L. (2020). SyRI, ¿a quién sanciono? Garantías frente al uso de inteligencia artificial y decisiones automatizadas en el sector público y la sentencia holandesa de febrero de 2020. *La Ley Privacidad*, 4.

De Hoyos Sancho, M. (2003). Firma digital y comunicaciones procesales. *Actualidad Jurídica Aranzadi*. 571, 4-5.

— (2015a). El reconocimiento mutuo de resoluciones por las que se impone una penal o medida privativa de libertad. En Arangüena Fanego, C.; de Hoyos Sancho, M. y Rodríguez-Medel Nieto, C. *Reconocimiento mutuo de resoluciones penales en la Unión Europea* (pp. 107-128). Aranzadi.

— (2015b). La orden europea de protección de víctimas: análisis normativo. En Arangüena Fanego, C.; de Hoyos Sancho, M. y Rodríguez-Medel Nieto, C. *Reconocimiento mutuo de resoluciones penales en la Unión Europea* (pp. 271-302). Aranzadi.

— (2019a). La orden europea de investigación: reflexiones sobre su potencial efectividad a la vista de los motivos de denegación del reconocimiento y ejecución en España. *Revista General de Derecho Procesal*, 47.

— (2019b). Orden europea de investigación: avanzando hacia la integración en materia procesal penal. En Barona Vilar, S. (Ed.), *Claves de la Justicia Penal: feminización, inteligencia artificial, supranacionalidad y seguridad* (pp. 343-370). Tirant lo Blanch.

— (2019c). El principio de subsidiariedad y la autonomía procesal de los Estados de la Unión Europea. *Jueces para la Democracia*, 96, 36-47.

De la Oliva Santos, A. (2019). "Justicia predictiva", interpretación matemática de las normas, sentencias robóticas y la vieja historia del "Justizklavier". *El Cronista del Estado Social y Democrático de Derecho*, 80, pp. 30-37.

De Miguel Beriain, I. (2018). Does the use of risk assessments in sentences respect the right to due process? A critical analysis of the Wisconsin v. Loomis ruling. *Law, Probability and Risk*, 1 (17), 45-53.

Fair Trials (2020). *Regulating AI for Use in Criminal Justice Systems in the EU,* informe publicado en https://www.fairtrials.org.

Feijóo Sánchez, B. (2007). Individualización de la pena y teoría de la pena proporcional al hecho, *InDret*, 1.

Fernández, C. B. (2020, 13 feb.). Primera sentencia europea que declara ilegal un algoritmo de evaluación de características personales de los ciudadanos. *Diario La Ley*.

Ferrer Beltrán, J. (2007). *La valoración racional de la prueba.* Marcial Pons.

Gascón Inchausti, F. (2019). Desafíos para el proceso penal en la era digital: externalización, sumisión pericial e inteligencia artificial. En Conde Fuentes, J.; Serrano Hoyo, G.; Arrabal Platero, P. y García Molina, P. (dirs.), *La justicia digital en España y en la Unión Europea (pp. 191-206)*. *Atelier.*

Gialuz, M. (2019, 29 may). Quando la giustizia penale incontra l'intelligenza artificiale: luci e ombre dei *Risk Assessment Tools* tra Stati Uniti ed Europa, *Diritto Penale Contemporaneo*, en https://archiviodpc.dirittopenaleuomo.org.

González Cano, M. I. (Dir.) (2019). *Orden europea de investigación y prueba transfronteriza en la Unión Europea.* Tirant lo Blanch.

Kearns, M. y Roth, A. (2020). *El algoritmo ético. La ciencia del diseño de algoritmos socialmente responsables.* La Ley-Wolters Kluwer.

Laro González, M. E. (2019). La orden europea de investigación. Especial consideración de los problemas que plantea el doble control de admisibilidad en la obtención de pruebas. En González Cano, I. (dir.), *Orden europea de investigación y prueba transfronteriza en la Unión Europea* (pp. 797-814). Tirant lo Blanch.

Maldonato, L. (2019). Algoritmi predittivi e discrezionalitá del giuidice: una nuova sfida per la giustizia penale, *Diritto Penale Contemporaneo*, 2, 401-414.

Martín Diz, F. (2019a). Inteligencia artificial y proceso: garantías frente a eficiencia en el entorno de los derechos procesales fundamentales. En Jiménez Conde, F. y Bellido Penadés, R.; Llopis Nadal, P. y de Luis García, E. (coords.). *Justicia:* ¿garantías versus eficiencia*?* (pp. 815-827) Tirant lo Blanch.

— (2019b). Aplicaciones de Inteligencia Artificial en procesos penales por delitos relacionados con la corrupción. En Rodríguez García, N.; Carrizo González-Castell, A.; Rodríguez López, F.; Sánchez Bernal, J. y Carrillo del Teso, A.E. (coords.), *Corrupción: 'compliance', represión y recuperación de activos* (pp. 533-568). Tirant lo Blanch.

McKay, C. (2020). Predicting risk in criminal procedure: actuarial tools, algorithms, AI and judicial decisión-making. *Current Issues in Criminal Justice,* 1 (32), 22-39.

Mitchel, J. *et al.* (2020). Machine learning for determining accurate outcomes in criminal trials. *Law, Probability and Risk*, 1 (19), 43-65.

Nieva Fenoll, J. (2018). *Inteligencia artificial y proceso judicial,* Marcial Pons, 2018.

Núñez Zorrilla, M. C. (2018). Los nuevos retos de la UE en la regulación de la responsabilidad civil por los daños causados por la inteligencia artificial. *Revista Española de Derecho Europeo*, 66, 9-53.

—(2019). *Inteligencia artificial y responsabilidad civil derivada de daños ocasionados por robots autónomos con inteligencia artificial.* Reus.

Occhiuzzi, B. (2019). Algoritmi predittivi: alcune premesse metodologiche, *Diritto Penale Contemporaneo*, 2, 391-400.

Parolli, C. y Sellaroli, V. (2019). Sistema penale e intelligenza artificiale: molte speranze e qualche equivoco. *Diritto Penale Contemporaneo,* 6, 47-71.

Pretrial Justice Institute (2020, 7 de feb.). *Updated Position on Pretrial Risk Assessment Tools*, https://www.pretrial.org/wp-content/uploads/Risk-Statement-PJI-2020.pdf.

Quattrocolo, S. (2018, 22 de mar). Intelligenza artificiale e giustizia: nella cornice della Carta ética europea, gli spunti per un'urgente discussione tra scienze penali e informatiche, en www.lalegislazionepenale.it.

Quattrocolo, S. (2019). Equitá del processo penale e *automated evidence* alla luce della Convenzione europea dei diritti dell'uomo. *Revista ítalo-española de Derecho Procesal,* vol. 2, 1 y ss.

Signorato, S. (2020). Giustizia penale e intelligenza artificiale. Considerazioni in tema di algoritmo predittivo. *Rivista di Diritto Processuale*, 2 (75), 605-616.

Susskind, R. (2019). *On line Courts and the Future of the Justice,* Oxford University Press.

Susskind, R. (2020). *Tribunales on-line y la Justicia del futuro,* La Ley-Wolters Kluwer.

The Law Society (2019). *Algorithms in the Criminal Justice System*, en https://www.lawsociety.org.uk/topics/research/algorithm-use-in-the-criminal-justice-system-report.

Ubertis, G. (2020). *Intelligenza artificiale, giustizia penale, controllo umano significativo, Sistema penale*, texto presentado en Milán el 15 de octubre de 2020, en http://www.ristretti.it/commenti/2020/novembre/pdf3/articolo_ubertis.pdf.

B. Legislación y jurisprudencia

Agencia de los Derechos Fundamentales de la Unión Europea. (2020). *Facial recognition technology. Fundamental rights considerations in the context of law enforcement* https://op.europa.eu/es/publication-detail/-/publication/95cbdd4c-3d8d-11ea-ba6e-01aa75ed71a1/language-es.

Carta de Derechos Fundamentales de la Unión Europea. *DOUE* C 202 (7 de jun. de 2016).

Carta ética europea sobre el uso de la inteligencia artificial en los sistemas judiciales y su entorno (2018). Aprobada por la European Commission for the Efficiency of Justice (CEPEJ). https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c.

Comisión Europea (2018). *Comunicación sobre la inteligencia artificial para Europa* [COM(2018) 237 final].

— (2019). Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones. Generar confianza en la inteligencia artificial centrada en el ser humano [COM(2019) 168 final].

— (2020). *Informe sobre el marco de seguridad y responsabilidad civil de la Inteligencia Artificial, el Internet de las cosas y la robótica,* adjunto al propio Libro Blanco sobre IA. [COM(2020) 64 final]. https://ec.europa.eu/transparency/regdoc/rep/1/2020/ES/COM-2020-64-F1-ES-MAIN-PART-1.PDF.

— (2020). *Libro Blanco sobre la inteligencia artificial – un enfoque europeo orientado a la excelencia y la confianza* [COM(2020) 65 final]. https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_es.pdf.

— (2020, sep). *Study on the use of innovative technologies in the justice field. Final Report*.

Committee of Experts on Internet Intermediaries (MSI-NET). (2018). *Algorithms and Human Rights. Study on the Human Rights Dimensions of Auto-mated Data Processing Techniques and Possible Regulatory Implication* (Council of Europe Study DGI(2017)12). https://rm.coe.int/algorithms-and-human-rights-en-rev/16807956b5/.

Convenio para la protección de los derechos humanos y de las libertades fundamentales, hecho en Roma el 4 de noviembre de 1950.

Decisión Marco 2008/909/JAI del Consejo, de 27 de noviembre de 2008, relativa a la aplicación del principio de reconocimiento mutuo de sentencias en materia penal por las que se imponen penas u otras medidas privativas de libertad a efectos de su ejecución en la Unión Europea. *DOUE* L 327 (5 de dic. de 2008).

Decisión Marco 2008/977/JAI del Consejo, de 27 de noviembre de 2008, relativa a la protección de datos personales tratados en el marco de la cooperación policial y judicial en materia penal. *DOUE* L 350 (30 de dic. de 2008).

Directiva (UE) 2016/680 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales por parte de las autoridades competentes para fines de prevención, investigación, detección o enjuiciamiento de infracciones penales o de ejecución de sanciones penales, y a la libre circulación de dichos datos y por la que se deroga la Decisión Marco 2008/977/JAI del Consejo. *DOUE* L 119 (4 de may. de 2016).

Directiva 2011/99/UE del Parlamento Europeo y del Consejo, de 13 de diciembre de 2011, sobre la orden europea de protección. *DOUE* L 338 (21 de dic. de 2011).

Directiva 2012/13/UE del Parlamento Europeo y del Consejo, de 22 de mayo de 2012, relativa al derecho a la información en los procesos penales. *DOUE* L 142 (1 de jun. de 2012).

Fair Trials (2020, nov.). *Regulating AI for Use in Criminal Justice Systems in the EU*. https://www.fairtrials.org.

Grupo de expertos de alto nivel sobre inteligencia artificial. (2019). *Directrices para una IA fiable.* https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top.

Grupo de Expertos en datos FAIR de la Comisión. (2018). *Turning FAIR into reality. Final report and action plan from the European Commission expert group on FAIR data* https://ec.europa.eu/info/sites/info/files/turning_fair_into_reality_1.pdf.

Loomis v. Wisconsin (2017) 137 S.Ct.2290.
Pretrial Justice Institute. (2020, 7 feb). *Updated Position on Pretrial Risk Assessment Tools*. https://www.pretrial.org/wp-content/uploads/Risk-Statement-PJI-2020.pdf.
Reglamento (CE) núm. 1/2005 del Consejo, de 22 de diciembre de 2004, relativo a la protección de los animales durante el transporte y las operaciones conexas y por el que se modifican las Directivas 64/432/CEE y 93/119/CE y el Reglamento (CE) núm. 1255/97. *DOUE* L 3 (5 de ene. De 2005).
Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos). *DOUE* L 119 (4 de may. De 2016).
Reglamento (UE) 2018/1725 del Parlamento Europeo y del Consejo, de 23 de octubre de 2018, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales por las instituciones, órganos y organismos de la Unión, y a la libre circulación de esos datos, y por el que se derogan el Reglamento (CE) núm. 45/2001 y la Decisión núm. 1247/2002/CE. *DOUE* L 295 (21 de nov. de 2018).
Reglamento (UE) núm. 606/2013 del Parlamento Europeo y del Consejo, de 12 de junio de 2013, relativo al reconocimiento mutuo de medidas de protección en materia civil. *DOUE* L 181 (29 de jun. de 2013).
State v. Loomis, 881 N.W.2d 749 (Wis. 2016).
The Law Society (2019, jun). *Algorithms in the Criminal Justice System*. file:///C:/Users/user/AppData/Local/Temp/algorithms-in-criminal-justice-system-report-2019.pdf
Tratado de Funcionamiento de la Unión Europea (versión consolidada). *DOUE* C 202 (7 de jun. de 2016).