LA INTELIGENCIA ARTIFICIAL APLICADA AL PROCESO PENAL DESDE LA PERSPECTIVA DE LA UE

Guillermo SCHUMANN BARRAGÁN
Profesor ayudante
Departamento de Derecho Procesal y Derecho Penal
Universidad Complutense de Madrid
gschuman@ucm.es

ORCID ID: https://orcid.org/0000-0003-1934-7808

Trabajo publicado en PEREIRA PUIGVERT, S., ORDÓÑEZ PONZ, F. (dirs.), *Investigación y proceso penal en el siglo XXI: nuevas tecnologías y protección de datos*, Aranzadi, Cizur Menor, 2021, pp. 517-539. ISBN: 978-84-1390-522-8.

SCHUMANN BARRAGÁN, G., "La inteligencia artificial aplicada al proceso penal desde la perspectiva de la UE" en PEREIRA PUIGVERT, S., ORDÓÑEZ PONZ, F. (dirs.), *Investigación y proceso penal en el siglo XXI: nuevas tecnologías y protección de datos*, Aranzadi, Cizur Menor, 2021, pp. 517-539.

LA INTELIGENCIA ARTIFICIAL APLICADA AL PROCESO PENAL DESDE LA PERSPECTIVA DE LA UE

Guillermo SCHUMANN BARRAGÁN
Profesor ayudante
Departamento de Derecho Procesal y Derecho Penal
Universidad Complutense de Madrid[1]

SUMARIO: I. INTRODUCCIÓN. — II. LA PERSPECTIVA DE LA UE SOBRE IA. 2.1. La Comisión Europea: el Libro Blanco sobre inteligencia artificial; 2.2. La Comisión Europea: la Propuesta de Reglamento sobre inteligencia artificial; 2.3. El Parlamento Europeo; 2.4. El Consejo de la UE; 2.5. La Agencia de los Derechos Fundamentales de la UE. — III. ALGUNAS REFLEXIONES ENTORNO A LA PERSPECTIVA DE LA UE SOBRE IA. 3.1. Un análisis técnico y multinivel de los derechos fundamentales en riesgo; 3.2. Los riesgos asociados al uso de sistemas de IA para el derecho a la tutela judicial efectiva y la presunción de inocencia; 3.3. La contradicción de la prueba (art. 47 CDFUE) y la opacidad del sistema de IA; 3.4. La externalización de la investigación penal: la eficacia de los derechos fundamentales entre particulares; 3.5. Transparencia y protección jurídica de la IA. — V. CONCLUSIÓN.[2]

I. INTRODUCCIÓN

La Inteligencia Artificial (en adelante, IA) es un conjunto de tecnologías que combinan información, algoritmos y capacidad informática.[3] Gráficamente puede decirse que los algoritmos son el motor de la IA y los datos su combustible. El uso de sistemas de IA por las Fuerzas y Cuerpos de Seguridad del Estado para investigar y perseguir el delito es desde hace tiempo ya una realidad en nuestro país —v.gr. VioGén o VeriPol[4] —. La IA puede influir en el proceso penal en áreas como el acceso a fuentes de prueba o la valoración de la prueba.[5]

Una Europa verde y digital es el *leitmotiv* que inspira la acción política de la Comisión Europea dirigida por Ursula VON DER LEYEN. El Pacto Verde Europeo y la transformación digital se sitúan en un lugar prioritario en la agenda política de la UE. En este contexto,

[1] ORCID ID: https://orcid.org/0000-0003-1934-7808

[2] Este trabajo se enmarca dentro del Proyecto de Investigación "Hacia un proceso civil convergente con Europa. Hitos presentes y retos futuros" (PGC2018-094693-B-I00), financiado por el Ministerio de Ciencia, Innovación y Universidades.

[3] Libro Blanco sobre la inteligencia artificial - un enfoque europeo orientado a la excelencia y la confianza [COM(2020) 65 final]. p. 2.

[4] VioGén es un sistema de IA utilizado por la Policía Nacional para identificar el riesgo de violencia de género. SÁNCHEZ LÓPEZ, B., "La diligencia policial de valoración del riesgo de violencia de género en el sistema Viogén", *FORO. Revista de Ciencias Jurídicas y Sociales, Nueva Época*, vol. 22, núm. 1. Disponible en: https://doi.org/10.5209/foro.66637 Por otro lado, VeriPol es un sistema de IA también utilizado por la policía que, a través del análisis del lenguaje, permite identificar denuncias falsas. MISURACA, G., AND VAN NOORDT, C., *Overview of the use and impact of AI in public services in the EU*, Publications Office of the European Union, Luxembourg, 2020, p. 47.

[5] HOYOS SANCHO, M., "El libro blanco sobre inteligencia artificial de la Comisión Europea: reflexiones desde las garantías esenciales del proceso penal como «sector de riesgo»", *Revista Española de Derecho Europeo*, núm. 76, 2020, pp. 22-25.

la regulación de la IA aplicada a distintos sectores de la vida social y, entre ellos, a la investigación y persecución del delito adquiere un protagonismo indiscutible a nivel europeo.

El proceso legislativo ordinario de la UE es un proceso complejo y compartido entre la Comisión —que tiene la iniciativa—, el Parlamento y el Consejo de la UE (art. 294 TFUE). Es usual que antes del inicio formal del proceso legislativo exista un diálogo abierto entre las Instituciones y entre estas y la sociedad civil para fijar las líneas generales de política legislativa que inspirará una futura regulación. Este diálogo suele tener su origen en la publicación de libros verdes y blancos por la Comisión o de recomendaciones de reglamentos o directivas por el Parlamento.

El propósito de estas líneas es examinar cuál es la posición de la UE en relación con aquellos usos de la IA con influencia en el proceso penal y hacer algunas reflexiones críticas al respecto. Para ello habrá que identificar los distintos enfoques de política legislativa de las Instituciones que intervienen en el proceso legislativo y que presumiblemente se materializarán en un futuro reglamento.

II. LA PERSPECTIVA DE LA UE SOBRE IA

En los últimos tiempos los diferentes actores institucionales en la UE han ido tomando posición en relación con la IA y han ofrecido indicaciones acerca de qué ha de regularse y cómo. Corresponde ahora analizar las tomas de posición de la Comisión, el Parlamento Europeo, el Consejo de la UE y la Agencia de Derechos Fundamentales.

Como se verá, una idea constante que inspira la posición de todas estas Instituciones es el «enfoque europeo»— *the European approach*— que debe adoptarse. En el escenario geopolítico actual, la UE aspira a tener una visión propia de la IA y a promoverla en el mundo. Es una posición que políticamente sitúa a Europa como garante y promotor internacional de un determinado modelo social: una transformación digital compatible con los derechos fundamentales y los valores sobre los que se construye la Unión.

2.1. La Comisión Europea: el Libro Blanco sobre inteligencia artificial

La Comisión publicó en febrero del 2020 el Libro Blanco «sobre inteligencia artificial — un enfoque europeo orientado a la excelencia y la confianza» [COM(2020) 65 final].[6] Los libros verdes y blancos de la Comisión son documentos que contienen propuestas o una primera aproximación de política legislativa sobre una determinada cuestión. Su

[6] El Libro Blanco está disponible en: https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/excellence-trust-artificial-intelligence_es También es de interés la página web de la Comisión dedicada a la IA: https://ec.europa.eu/digital-single-market/en/artificial-intelligence

intención es provocar un diálogo abierto con el resto de Instituciones, agentes nacionales y sociales.[7]

El Libro Blanco se construye sobre dos grandes pilares: innovación y confianza. En relación con este último, la Comisión considera que solo una regulación de la IA que genere confianza hará socialmente aceptable su uso en el sector privado y público. Se parte de la premisa de que la IA ya se aplica por parte de las administraciones públicas y de que debe potenciarse este uso, en lo que ahora nos interesa, como un instrumento para la persecución del delito.[8]

El Libro Blanco aborda una regulación de la IA desde una *aproximación de riesgos* —*Risikoadaptierte Regulierung*—.[9] Se cree que una regulación exhaustiva y detallada es un obstáculo para la innovación tecnológica. Por ello, se identifican los riesgos asociados al uso de la IA, qué usos concretos materializan estos riesgos y se propone una regulación proporcional y específica para ellos. Se considera que una aplicación de IA es de «riesgo elevado» en función del sector en el que se aplica —entre ellos, «determinados ámbitos del sector público [como] el poder judicial»[10] —. También lo es si de su uso se derivan «riesgos significativos» para los individuos, como la producción de efectos en su esfera jurídica o riesgos para su vida o integridad física.[11] En relación con estos usos de alto riesgo se propone una regulación que imponga obligaciones específicas a los desarrolladores y usuarios y un mecanismo de certificación o control previo que las fiscalice.[12]

En lo que ahora interesa, el principal riesgo que se identifica asociado al uso de la IA es el de vulneraciones de derechos fundamentales. Por un lado, el riesgo de que se vulnere el derecho a la igualdad y no discriminación. Su causa está en determinados *sesgos* —*bias*— *(i)* de la información con la que se entrena y/o posteriormente se nutre el algoritmo

[7] Después de la publicación del Libro Blanco se llevó a cabo una consulta pública que recibió 1215 observaciones de distintas instituciones públicas y privadas. Todas están accesibles en: https://ec.europa.eu/digital-single-market/en/news/white-paper-artificial-intelligence-public-consultation-towards-european-approach-excellence El Libro Blanco tiene como punto de partida los trabajos realizados por un Grupo de expertos de alto nivel sobre inteligencia artificial (*AI HLEG*) nombrado por la Comisión. Los trabajos realizados por el grupo de expertos están disponibles en: https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence También son de interés los estudios realizados por la *IA Watch*, un servicio de investigación sobre IA constituido por la Comisión. Sus trabajos de investigación está disponibles en: https://knowledge4policy.ec.europa.eu/ai-watch_en

[8] Libro Blanco sobre inteligencia artificial, p. 2.

[9] En octubre del 2019 la *Datenethikkommission* del Ministerio de Justicia federal alemán emitió una Opinión sobre una futura regulación sobre el uso de algoritmos y sistema de IA. Se proponía adoptar una regulación basada en los riesgos asociados al uso de la IA —*Risikoadaptierte Regulierung*—. Puede verse la influencia de esta Opinión y otros estudios en el Libro Blanco. La Opinión está disponible en: https://www.bmjv.de/DE/Themen/FokusThemen/Datenethikkommission/Datenethikkommission_node.html

[10] Libro Blanco sobre inteligencia artificial, nota al pie 50.

[11] Libro Blanco sobre inteligencia artificial, pp. 22-29.

[12] Libro Blanco sobre inteligencia artificial, pp. 21-22.

y/o *(ii)* del propio algoritmo que la procesa. El sesgo de los datos o del algoritmo puede potenciar o acentuar discriminaciones ya existentes.

> Un ejemplo recurrente de discriminación en la aplicación de IA es el relativo a los sistemas de evaluación de criminalidad o reincidencia delictiva —v.gr. el sistema COMPAS en EE. UU[13] o HART en Reino Unido[14]—. Es posible que el algoritmo del sistema asocie a una determinada condición personal un mayor o menor riesgo de incidencia delictiva —v.gr. género, edad, lugar de residencia, etnia, etc.—. También es posible que, aunque el algoritmo sea «imparcial», el sesgo venga de la propia información de la que se nutre—v.gr. que una mayor vigilancia policial de determinados grupos sobredimensione su criminalidad en las estadísticas policiales—. En ambos casos se corre el riesgo de que la aplicación de la IA potencie los sesgos humanos que consciente o inconscientemente ya existen.

También se identifica el riesgo de que se vulnere el derecho a la protección de datos al llevarse a cabo vigilancia biométrica masiva con sistemas de IA o al utilizarse estos para desanonimizar datos.[15]

Por último, se hace referencia a los riesgos que supone para el derecho a la tutela judicial efectiva la opacidad del algoritmo —*black-box effect*—.[16] En la medida en que el individuo no sepa que se está aplicando IA, no pueda conocer cómo funciona o esta no sea entendible, no podrá impugnar o defenderse frente a decisiones basadas en ella. Además, la opacidad del algoritmo también obstaculiza su control por parte de las autoridades públicas.

En el marco de los riesgos identificados y de los usos de IA de alto riesgo se propone, por un lado, adaptar la regulación existente, en concreto, aquellas normas de Derecho derivado que desarrollan los derechos reconocidos en la CDFUE: el Reglamento General de Protección de Datos[17], las «directivas sobre igualdad» —*the European non-discrimination law*—[18] y la Directiva (UE) 2016/680 sobre tratamiento de datos para fines de prevención, investigación, detección o enjuiciamiento de infracciones penales.[19]

---

[13] GASCÓN INCHAUSTI, F., "Desafíos para el proceso penal en la era digital: externalización, sumisión pericial e inteligencia artificial" en CONDE FUENTES, J., SERRANO HOYO, G. (dirs.), *La Justicia digital en España y la Unión Europea*, Atelier, Barcelona, 2019, pp. 201-204.

[14] ESTÉVEZ MENDOZA, L, "Algoritmos policiales basados en IA y derechos fundamentales a la luz de HART y VALCRI: garantías *versus* eficiencia" en JIMÉNEZ CONDE, F., BELLIDO PENADÉS, R., *Justicia: ¿garantía versus eficiencia?*, Tirant lo Blanch, Valencia, 2019, pp. 668-669.

[15] Es posible que a través del análisis y relación de distintas fuentes de información inicialmente anónima se desanonimice y se identifique a los individuos. La anonimización de la información permitiría «escapar» del Reglamento de protección de datos mientras que el análisis y tratamiento de distintas fuentes de información anónima permitiría desanonimizar la información a través de IA y perfilar a los individuos [vid. art. 72.1.p) LO 3/2018].

[16] Libro Blanco sobre inteligencia artificial, pp. 13-15.

[17] Reglamento (UE) 2016/679.

[18] Directivas 2000/43/CE, 2000/78/CE, 2004/113/CE y 2006/54/CE.

[19] 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

Además, se plantea adoptar una regulación específica de carácter horizontal cuyas líneas generales se desarrollan en el Libro Blanco y que se materializan en la Propuesta de Reglamento (*vid. infra*).

En definitiva, el Libro Blanco identifica determinados riesgos asociados a la IA y los usos que los materializan. Entre ellos, su uso como mecanismo para prevenir y perseguir el delito y su incidencia en determinados derechos fundamentales.[20] En relación con ellos se propone adaptar la regulación europea que desarrolla estos derechos y adoptar un nuevo reglamento que regule el desarrollo, uso y adaptación de los sistemas de IA.

2.2. La Comisión Europea: la Propuesta de Reglamento sobre inteligencia artificial

El 21 de abril del 2021 la Comisión presentó su Propuesta de Reglamento sobre inteligencia artificial. [21] Con ello se activa el procedimiento legislativo y se pone en marcha la primera y eventuales futuras lecturas entre el Parlamento y el Consejo (art. 294 TFUE).

La aproximación regulatoria horizontal del uso de la IA en distintos sectores de la vida social y la decisión de hacerlo a través de un reglamento ha supuesto que el título competencial finalmente activado por la Comisión sea el relativo al mercado (digital) interior (art. 114 TFUE) (considerando 2). Es una cuestión de interés que ayuda a interpretar la lógica y funcionamiento del sistema diseñado.

La Propuesta de Reglamento —de 85 artículos— incorpora la aproximación de riesgos ya adelantada en el Libro Blanco y clasifica los usos de inteligencia artificial como *(i)* prohibidos (art. 5), *(ii)* de alto riesgo (art. 6 y anexo III) y *(iii)* de riesgo limitado (art. 69).

En principio, la Propuesta prohíbe la vigilancia biométrica masiva en tiempo real [art. 5 (d)]. Excepcionalmente su uso está permitido para perseguir el delito siempre que sea para la identificación de víctimas, la prevención de peligros inminentes a personas o ataques terroristas y la identificación de determinados sospechosos o delincuentes [art. 5 (d) (i) (ii) y (iii)]. Además de esta restricción objetiva, en función de las circunstancias la medida tiene que ser proporcional y estar limitada temporal y geográficamente. Su uso exige la aprobación de una autoridad judicial o administrativa independiente (¡!).

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.

[20] En este sentido, vid. el estudio de la *IA WATCH* sobre el uso actual de la IA en el sector público en la UE. MISURACA, G., AND VAN NOORDT, C., *Overview of the use and impact of AI*, op. cit.

[21] Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts [2021/0106 (COD)].

La Propuesta de Reglamento no enumera directamente los usos de alto riesgo, sino que se remite al Anexo III que lo acompaña.[22] Se trata de una técnica legislativa utilizada por la Comisión que le permite a través de la figura de los actos delegados adaptar rápidamente la regulación a los cambios económicos, sociales o, en este caso, tecnológicos (art. 7). En el Anexo se señalan como usos de alto riesgo bajo la clasificación de *law enforcement*, entre otros, aquellos *(i)* para medir el riesgo de incidencia o reincidencia delictiva o el riesgo de las víctimas potenciales, *(ii)* para evaluar la fiabilidad de las pruebas en el curso de la investigación o enjuiciamiento de delitos penales, *(iii)* para elaborar perfiles — *profiling*— o *(iv)* para analizar la escena del crimen.[23]

En relación con estos usos de alto riesgo se establecen determinados requisitos técnicos (arts. 8-15). Entre ellos destaca la incorporación de deberes de documentación que pretenden potenciar la transparencia del algoritmo a través de manuales que permitan a los usuarios utilizarlo e interpretar su resultado de forma adecuada (arts. 13). Además, se fijan determinadas obligaciones para los desarrolladores, distribuidores, importadores y usuarios del sistema de IA que son operativas en sus distintas fases de desarrollo, comercialización y uso (arts. 16-29). Entre ellas, la de llevar a cabo una evaluación previa de conformidad del sistema de IA con el Reglamento (art. 19).

> La Propuesta de Reglamento no adopta un sistema de certificación previa como en cierta manera se defendía en el Libro Blanco y, como se verá, defiende el Parlamento. Por el contrario, crea un procedimiento de evaluación de conformidad —*conformity assessment*— (arts. 31- 32, 43 y 48) por medio del que se valora internamente o a través de terceros el cumplimiento de los requisitos. Una vez llevada a cabo se presenta el resultado positivo a una de

[22] Annexes to the Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts [COM(2021) 206 final]. Conforme al considerando 26 de la propuesta «[c]ertain AI systems intended for the administration of justice and democratic processes should be classified as high-risk, considering their potentially significant impact on democracy, rule of law, individual freedoms as well as the right to an effective remedy and to a fair trial. In particular, to address the risks of potential biases, errors and opacity, it is appropriate to qualify as high-risk AI systems intended to assist judicial authorities in researching and interpreting facts and the law and in applying the law to a concrete set of facts».

[23] En concreto, el punto 6 del Anexo III incorpora los siguientes usos de alto riesgo: (a) AI systems intended to be used by law enforcement authorities for making individual risk assessments of natural persons in order to assess the risk of a natural person for offending or reoffending or the risk for potential victims of criminal offences; (b) AI systems intended to be used by law enforcement authorities as polygraphs and similar tools or to detect the emotional state of a natural person; (c) AI systems intended to be used by law enforcement authorities to detect deep fakes as referred to in article 52(3); (d) AI systems intended to be used by law enforcement authorities for evaluation of the reliability of evidence in the course of investigation or prosecution of criminal offences; (e) AI systems intended to be used by law enforcement authorities for predicting the occurrence or reoccurrence of an actual or potential criminal offence based on profiling of natural persons as referred to in Article 3(4) of Directive (EU) 2016/680 or assessing personality traits and characteristics or past criminal behaviour of natural persons or groups; (f) AI systems intended to be used by law enforcement authorities for profiling of natural persons as referred to in Article 3(4) of Directive (EU) 2016/680 in the course of detection, investigation or prosecution of criminal offences; (g) AI systems intended to be used for crime analytics regarding natural persons, allowing law enforcement authorities to search complex related and unrelated large data sets available in different data sources or in different data formats in order to identify unknown patterns or discover hidden relationships in the data».

las denominadas autoridades de notificación — *notified body*— (art. 30). Esto permitirá que el sistema de IA incorpore un etiquetado europeo especial de conformidad — *CE marking of conformity*— (arts. 49). Esta cuestión, junto con la regulación de la vigilancia biométrica, previsiblemente serán los principales puntos de desencuentro sobre los que se centrará la posterior negociación entre el Parlamento y el Consejo en la primera o eventuales futuras lecturas.

Por último, la propuesta crea un entramado institucional a nivel europeo y nacional (art. 30-39, 56-59). En él destaca la creación del *European Artificial Intelligence Board* (art. 56-58) y de un registro público en el que se incluyan los usos de alto riesgo de IA operativos (art. 60).

En sintonía con la lógica competencial de preservar el mercado (digital) interior (art. 114 TFUE), la propuesta de reglamento incorpora en su art. 71 multas de hasta 30 millones de euros o 6% de la facturación anual mundial en caso de incumplimiento de las obligaciones que en él se establecen.[24]

En definitiva, la aplicación de la IA al proceso penal en la Propuesta de Reglamento se centra en la regulación de la vigilancia biométrica y en la consideración en el Anexo III de determinados sistemas para investigar y perseguir el delito como usos de alto riesgo. La aproximación regulatoria horizontal y el título competencial que se hace valer «diluye» una regulación específica y detallada de su aplicación en el proceso penal, aunque por otro lado potencia las multas frente a eventuales incumplimientos y, con ello, su propia eficacia.

2.3. El Parlamento Europeo

En octubre del 2020 el Parlamento Europeo aprobó la Resolución con recomendaciones a la Comisión sobre un «marco de los aspectos éticos de la inteligencia artificial» [2020/2012(INL)].[25] En esta el Parlamento propone adaptar la regulación comunitaria que ya existe y, además, incorpora una propuesta de reglamento «sobre principios éticos

[24] Conforme al art. 71 de la Propuesta de Reglamento: «4. The non- compliance of the AI system with any requirements or obligations under this Regulation, other than those laid down in Articles 5 and 10, shall be subject to administrative fines of up to 20 000 000 EUR or, if the offender is a company, up to 4 % of its total worldwide annual turnover for the preceding financial year, whichever is higher. 5. The supply of incorrect, incomplete or misleading information to notified bodies and national competent authorities in reply to a request shall be subject to administrative fines of up to 10 000 000 EUR or, if the offender is a company, up to 2 % of its total worldwide annual turnover for the preceding financial year, whichever is higher».

[25] La resolución está disponible en: https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275_ES.html Además, se aprobó una Resolución con recomendaciones para la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial [(2020/2014(INL))] así como en determinados derechos de propiedad intelectual de creaciones generadas directamente *por* sistemas de IA [(2020/2015(INI))]. Los documentos de las resoluciones y estudios previos del Parlamento están disponibles en: https://www.europarl.europa.eu/news/es/headlines/priorities/inteligencia-artificial-en-la-ue/20201015STO89417/regulacion-de-la-inteligencia-artificial-en-la-ue-la-propuesta-del-parlamento

para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas».[26]

La Propuesta de Reglamento del Parlamento parte de la misma aproximación de riesgos adoptada por la Comisión en su Libro Blanco y Propuesta. Se enuncian una serie de «principios éticos» para cualquier uso de IA (art. 5) y se fijan obligaciones aplicables solo para los usos de alto riesgo (art. 6 y siguientes). Se consideran usos de IA de alto riesgo aquellos que pueden causar lesiones, daños o vulnerar derechos fundamentales [art. 4.e)]. El riesgo debe evaluarse, entre otros criterios objetivos, en función del sector en el que se utilizan y su finalidad específica (art. 14). La Propuesta de Reglamento del Parlamento incorpora un anexo en el que también se señalan como sectores de riesgo alto el sector público y el sistema judicial.

Los usos de alto riesgo deben mantenerse bajo control y supervisión humanos (art. 7) y deben cumplir con determinados estándares de seguridad y transparencia —como informar a los usuarios cuando están interactuando con sistemas de IA [art. 8.1 f)] —. Se impone que los sistemas no incorporen sesgos (art. 9). Solo se permite un trato diferenciado por el sistema de IA cuando exista una «finalidad objetiva, razonable y legítima que sea proporcionada y necesaria» (art. 9.2). En relación con ello, se considera que «son objetivos legítimos que, en virtud del presente Reglamento, pueden justificar objetivamente cualquier diferencia de trato entre personas o grupos de personas la *protección de la seguridad* y la salud públicas, la *prevención de infracciones penales*» (considerando 25). También se regula el uso y recogida de datos biométricos, que solo podrá ser utilizado por autoridades públicas para fines de interés público en un lugar y tiempo determinado (art. 12).

Los usos de alto riesgo de IA se someterán a una certificación previa de la autoridad nacional de control (arts. 15 y 16) y estarán sometidos a evaluación y seguimiento (art. 15). Para ello se regula un marco institucional de supervisión entre autoridades nacionales (art. 18) y una coordinación europea por parte de la Comisión (art. 20).

El Parlamento hace hincapié en la necesidad de incorporar «vías de recurso efectivas» que permita una «revisión imparcial, efectuada por seres humanos, de todas las denuncias de vulneraciones de los derechos de los ciudadanos […] independientemente de si tienen su origen en agentes del sector público o del privado». Sin embargo, no existen en la Propuesta mecanismos procesales o medidas específicas para ello. La exigencia de transparencia del algoritmo se concreta en que el individuo sepa que está interactuando con un sistema de IA [art. 8.1.f)] y en que se documenten sus elementos, procesos y fases

[26] Como antecedente de la resolución es interés el estudio encargado por el Parlamento Europeo GONZÁLEZ FUSTER, G., *Artificial Intelligence and Law Enforcement Impact on Fundamental Rights*, European Parliament's Committee on Civil Liberties, Justice and Home Affairs. El estudio está disponible en: https://www.europarl.europa.eu/thinktank/en/document.html?reference=IPOL_STU(2020)656295

de desarrollo y uso para que puedan ser controlados por las autoridades nacionales (art. 8.2).

> Actualmente está en proceso de tramitación una Opinión «sobre la inteligencia artificial en el Derecho penal y su utilización por las autoridades policiales y judiciales en asuntos penales» [2020/2016(INI)].[27] En esta se identifica que la opacidad y la discriminación en los sistema de IA se potencia «en el ámbito policial y de la justicia penal, ya que pueden afectar a la presunción de inocencia, a los derechos fundamentales de libertad y seguridad de la persona y los derechos a una tutela judicial efectiva y a un juicio justo».[28]

2.4. El Consejo de la UE

En octubre del 2020 el Consejo de la Unión Europea publicó unas conclusiones sobre la aplicación de la CDFUE «en el contexto de la inteligencia artificial y el cambio digital».[29] Se trata de un documento de consenso pactado por las delegaciones nacionales de los EE. MM. con 30 conclusiones sobre la cuestión. En él se hace referencia al uso de la IA para facilitar la «labor de las fuerzas y cuerpos de seguridad» y la «búsqueda de pruebas fiables en causas penales» (conclusión 16).

Después de poner en valor la importancia y el potencial de la IA en la UE, se confirma que existen riesgos de vulneración de derechos fundamentales. Se afirma que para garantizar la compatibilidad de los sistemas de IA con la CDFUE «deben afrontarse retos como la opacidad, la complejidad, el sesgo, cierto grado de imprevisibilidad y un comportamiento parcialmente autónomo» (conclusiones 5 y 11). Se acoge favorablemente el Libro Blanco presentado por la Comisión y su propuesta de adaptar la legislación europea ya existente y proponer una nueva de carácter horizontal (conclusiones 6 y 12). Se enfatiza el «enfoque europeo» de la IA centrada en el respeto de los derechos fundamentales ya que «[e]n el mundo digital debe aplicarse el mismo grado de protección que en el mundo físico» (conclusión 7). Se recuerda que conforme al art. 52 CDFUE solo podrán introducirse limitaciones al ejercicio de los derechos de la Carta cuando sean necesarias para alcanzar un fin legítimo, sean proporcionales y estén establecidas por ley (Conclusión 10).

El Consejo identifica que el sesgo de los sistemas (conclusión 21), la vigilancia masiva (conclusión 18) o la opacidad son caracteres asociados a la IA que ponen en peligro los derechos a la igualdad, intimidad y tutela judicial efectiva. En relación con este último, se dice que deben «asegurarse vías de recurso efectivas para garantizar el derecho a un juicio justo, la presunción de inocencia y el derecho a la defensa» (conclusión 27).

---

[27] El expediente de tramitación parlamentaria está disponible en: https://oeil.secure.europarl.europa.eu/oeil/popups/ficheprocedure.do?reference=2020/2016(INI)&l=en

[28] Proyecto de Informe sobre la inteligencia artificial en el Derecho penal y su utilización por las autoridades policiales y judiciales en asuntos penales [2020/2016(INI)], p. 9.

[29] El documento está disponible en: https://data.consilium.europa.eu/doc/document/ST-11481-2020-INIT/es/pdf

2.5. La Agencia de los Derechos Fundamentales de la UE

La Agencia de Derechos Fundamentales de la UE —*Fundamental Rights Agency* (FRA)— tiene como fin institucional asesorar en materia de derechos fundamentales a las Instituciones de la UE. Aunque no participa directamente en el proceso legislativo ordinario, sus estudios dan una visión técnica y especializada que suele tener influencia en él. En diciembre del 2020 la Agencia publicó un estudio sobre la IA y los derechos fundamentales.[30]

El estudio selecciona algunos usos actuales de IA en la UE para, a modo de campo de prueba, identificar qué derechos fundamentales están en riesgo. Con esta metodología el estudio señala el marco legal de los derechos fundamentales en peligro y propone medidas concretas para minimizar los riesgos identificados.

El principal riesgo que se identifica asociado al uso de IA, al igual que el resto de Instituciones, es el de vulneraciones del derecho a la intimidad, a la igualdad y a la tutela judicial efectiva.[31]

Uno de los usos que se utilizan como campo de prueba para la identificación de estos riesgos es la vigilancia policial predictiva —*predictive policing*—.[32] El sesgo de la información o del algoritmo es un riesgo para el derecho a no padecer discriminación. La opacidad del algoritmo es un riesgo para el derecho a la tutela judicial efectiva. Y es que solo podrá pretenderse la tutela judicial *efectiva* frente al uso de sistema de IA cuando el individuo *(i)* sepa que se está utilizando un sistema de IA, *(ii)* sepa cómo y dónde pretender la tutela que el ordenamiento reconoce y *(iii)* pueda acceder y entender el contenido del sistema —*explainable IA*—.[33] En este sentido se destaca el derecho incorporado en el Reglamento de protección de datos a conocer cuándo una decisión se ha tomado por medios automatizados [considerando 71 y arts. 13.2.f), 14.2.g), 15.1.h) RGPD]. La opacidad también se considera un riesgo para el derecho a la presunción de inocencia.[34]

El estudio propone incorporar una evaluación de impacto —*impact assesment*— en el diseño de los sistemas de IA de forma que, a modo de *due dilligence*, se examine desde una fase inicial de desarrollo su compatibilidad con los derechos fundamentales. Y es que, de las entrevistas y trabajo de campo realizado para el estudio, se identifica un

[30] El estudio está disponible en: https://fra.europa.eu/en/publication/2020/artificial-intelligence-and-fundamental-rights
[31] *Artificial Intelligence and Fundamental Rights*, Fundamental Rights Agency, p. 57.
[32] *Artificial Intelligence and Fundamental Rights*, Fundamental Rights Agency, pp. 28, 34-36, 48-50, 69-71, 77.
[33] *Artificial Intelligence and Fundamental Rights*, Fundamental Rights Agency, pp. 10, 13, 75-76.
[34] *Artificial Intelligence and Fundamental Rights*, Fundamental Rights Agency, pp. 78.

desconocimiento y falta de conciencia de los desarrolladores de sistema de IA en relación con los derechos reconocidos en la Carta. [35]

Se constata en qué medida el marco legal e institucional creado en torno al derecho de protección de datos —v.gr. el delegado de protección de datos— ha exigido a las empresas que interioricen y consoliden determinados procesos empresariales de control. Se considera que imponer este análisis de impacto desde el diseño de los sistema de IA es la mejor manera de prevenir los riesgos asociados al uso de la IA.[36] En este sentido, se señala que el Derecho de la UE avanza en la dirección de incorporar obligaciones de diligencia debida y análisis de riesgos en relación con la vulneración de derechos fundamentales.

> En esta línea está el Reglamento sobre respeto de derechos humanos en las cadenas de suministro de determinados metales[37] o el proyecto de recomendaciones para la Comisión sobre diligencia debida de las empresas en proceso de tramitación en el Parlamento Europeo [2020/2129(INL)].[38]

III. ALGUNAS REFLEXIONES EN TORNO A LA PERSPECTIVA DE LA UE SOBRE IA

Se han examinado las posiciones de las distintas Instituciones de la UE y, con ello, la perspectiva desde la que la UE afronta la regulación de la IA. Es evidente —o por lo menos debería serlo— que los sistemas de IA deben respetar los derechos fundamentales. Aunque son proclamaciones plausibles y que forman parte del discurso político sobre el que se construye el «enfoque europeo», es exigible una mayor concreción.

Parece claro que determinadas características asociadas a la IA artificial suponen un riesgo para la dimensión subjetiva de los derechos fundamentales. Se trata de una cuestión que ya se ha incorporado en el discurso de política legislativa. Hecho esto, ahora debe profundizarse en el examen de los peligros identificados. Si los sistemas de IA son ya una realidad en la investigación criminal y existe un riesgo real de vulneración de derechos fundamentales, no bastan proclamaciones generales y abstractas: es exigible profundizar

---

[35] *Artificial Intelligence and Fundamental Rights,* Fundamental Rights Agency, p. 8.
[36] *Artificial Intelligence and Fundamental Rights,* Fundamental Rights Agency, pp. 8, 78.
[37] Reglamento (UE) 2017/821 del Parlamento Europeo y del Consejo, de 17 de mayo de 2017, por el que se establecen obligaciones en materia de diligencia debida en la cadena de suministro por lo que respecta a los importadores de la Unión de estaño, tantalio y wolframio, sus minerales y oro originarios de zonas de conflicto o de alto riesgo.
[38] Proyecto de informe con recomendaciones destinadas a la Comisión sobre diligencia debida de las empresas y responsabilidad corporativa [2020/2129(INL)]. Su tramitación parlamentaria puede seguirse en: https://oeil.secure.europarl.europa.eu/oeil/popups/ficheprocedure.do?lang=en&reference=2020/2129(INL) En este contexto son de especial interés los *Guiding Principles on Business and Human Rights* desarrollados por la ONU: https://www.ohchr.org/documents/publications/guidingprinciplesbusinesshr_en.pdf

en su análisis técnico-jurídico. En especial en un área en la que *el Estado tiene la doble condición de regulador y usuario*.

3.1. Un análisis técnico y multinivel de los derechos fundamentales en riesgo

El examen técnico-jurídico sobre los derechos fundamentales y el uso de sistemas de IA se desenvuelve en un ámbito de ejercicio del poder por el Estado especialmente intenso como lo son la investigación criminal y el posterior ejercicio del *ius puniendi*.[39] Es en este escenario en el que se tienen que identificar *(i)* las dimensiones concretas de los derechos fundamentales en riesgo; *(ii)* el usuario de los sistemas de IA —el Estado o los particulares—; y *(iii)* si determinados bienes jurídicos o derechos fundamentales constituyen un límite constitucionalmente legítimo a su ejercicio.

El Derecho de la UE está integrado por un completo sistema de derechos fundamentales incorporado en el Derecho primario —(TUE, TFUE y CDFUE)— y que es desarrollado legalmente a través de distintos reglamentos y directivas — a modo de nuestras Leyes Orgánicas (arts. 53 y 81 CE)—. El TJUE adquiere funciones de (*quasi*) tribunal constitucional interpretando los derechos y el ordenamiento comunitario conforme a ellos.[40]

Una vez identificados los derechos fundamentales y su núcleo esencial, debe examinarse si existen límites constitucionalmente admisibles a su ejercicio (art. 52 CDFUE). En nuestro caso, si la seguridad pública o la persecución del fenómeno delictivo como bienes jurídicamente protegidos o los derechos fundamentales de la víctima suponen limitaciones jurídicamente admisibles.

A todo lo anterior debe sumarse el análisis técnico a nivel nacional. Desde luego, no es inusual que exista un distinto nivel de protección de los derechos fundamentales a nivel europeo y nacional o que los estándares de limitación a su ejercicio sean diferentes. Esto acentúa un riesgo de choque entre el Derecho nacional y europeo —como se puso de manifiesto, v.gr., en el asunto *Melloni*— que pueda explicar una cierta reticencia del legislador europeo a concretar e incidir directamente en estas cuestiones.

3.2. Los riesgos asociados al uso de sistemas de IA para el derecho a la tutela judicial efectiva y la presunción de inocencia

En este contexto, es oportuno hacer un breve examen de los riesgos asociados al uso de sistemas de IA para los derechos y garantías procesales de naturaleza constitucional. Se dice, por ejemplo, que la opacidad del algoritmo supone un riesgo para el derecho a la

[39] Debe desterrarse toda referencia más o menos directa a la ética en la regulación. Aunque esta perspectiva ética se adopta únicamente respecto de los usos de bajo riesgo, parece preferible adoptar en todo caso un discurso jurídico-técnico. Se hace una crítica a la referencia a la ética en GONZÁLEZ FUSTER, G., *Artificial Intelligence and Law Enforcement*, op. cit., pp. 54-57, 69.

[40] En este sentido, vid. la reciente STJUE (Gran Sala) de 6 de octubre de 2020 (C-623/17) sobre la obligación de los proveedores de servicios de comunicaciones electrónicas de transmitir información a los servicios de inteligencia ingleses.

tutela judicial efectiva o para la presunción de inocencia (art. 6 CEDH, 47 y 48 CDFUE y art. 24 CE). Es evidente que la opacidad ocasiona una asimetría de información y de acceso a fuentes de prueba que puede afectar a la defensa técnica del investigado o imputado. Ahora bien, de un análisis de los distintos derechos subjetivos públicos que conforman el derecho a la tutela judicial efectiva o de defensa (art. 24.1 CE) no es evidente determinar qué concreta dimensión se vería vulnerada.[41] Tampoco es sencillo precisar desde nuestros parámetros constitucionales de qué modo la opacidad del algoritmo afecta a la presunción de inocencia (art 24.2 CE).

> El llamado juicio de indefensión supone: *(i)* la infracción de una norma procesal —requisito no siempre exigible por cuanto la indefensión puede provenir de la propia norma, que en su caso, sería inconstitucional—; *(ii)* la privación o limitación de las posibilidades de alegación o prueba contempladas legalmente; *(iii)* la falta de imputabilidad de esta privación al justiciable; *(iv)* que la indefensión sea definitiva y no subsanable; *(v)* la carga del justiciable de determinar la defensa privada o limitada y *(vi)* la existencia de un perjuicio material en su esfera jurídica.[42] En principio, y en la medida en que la parte no haya visto limitada sus posibilidades de alegaciones y prueba, la (mera) opacidad del algoritmo no supone una indefensión constitucionalmente relevante.

Como explica VEGAS TORRES, la presunción de inocencia se materializa en una regla de tratamiento del investigado en el proceso penal y en una regla de juicio que opera en el momento de dictar la sentencia.[43] La presunción de inocencia no es un mecanismo de fijación de hechos, por lo que no es técnicamente una presunción.[44] En realidad, la «presunción» opera a modo de *verdad interina —Interimswahrheiten—*: «se parte de» que el investigado es inocente.[45]

---

[41] DÍEZ-PICAZO GIMÉNEZ, I., "Comentario del art. 24" en ALZAGA VILLAMIL, O. (dir.), *Comentarios a la Constitución Española de 1978. Tomo III. Artículos 24 a 38*, Cortes Generales. Editoriales de Derecho Reunidas, Madrid, 1996, pp. 24-56. JARASS, G., "Art. 47" en *Charta der Grundrechte der Europäischen Union*, C. H. Beck, 4ª ed., München, 2021, Rn. 1-6, 22-63.

[42] DÍEZ-PICAZO GIMÉNEZ, I., "Comentario del art. 24", op. cit., pp. 50-56.

[43] VEGAS TORRES, J., *Presunción de inocencia y prueba en el proceso penal*, La Ley, Madrid, 1993, pp. 36-39. DÍEZ-PICAZO GIMÉNEZ, I., "Comentario del art. 24", op. cit., pp. 110-112. En el Derecho de la UE desarrolla este derecho, entre otras, en la Directiva (UE) 2016/343 del Parlamento Europeo y del Consejo, de 9 de marzo de 2016 por la que se refuerzan en el proceso penal determinados aspectos de la presunción de inocencia y el derecho a estar presente en el juicio.

[44] DE LA OLIVA SANTOS, A. "Prólogo" en VEGAS TORRES, J., *Presunción de inocencia*, op. cit., pp. 1-2. DÍEZ-PICAZO GIMÉNEZ, I., "Comentario del art. 24", op. cit., p. 112.

[45] «Técnicamente, y en el plano en el que nos estamos moviendo, la presunción de inocencia supone la afirmación de la inocencia del inculpado como *verdad provisional o interina* que entra en juego siempre que no se hayan conseguido fijar en el proceso, conforme a Derecho, todos los hechos de los que depende la responsabilidad criminal del acusado» VEGAS TORRES, J., *Presunción de inocencia*, op. cit., p. 212. «La "presunción de inocencia" debe ser situada entre las llamadas por la doctrina *verdades interinas o provisionales* […]» VAZQUEZ SOTELO, J. L., *Presunción de inocencia del imputado e íntima convicción del tribunal*, Bosch, Barcelona 1984, pp. 273-276. En este sentido las SSTS núm. 764/2017 de 27 noviembre [RJ 2017\5315]; num. 1066/2012 de 28 noviembre [RJ\2013\937]; núm. 151/2010 de 22 febrero [RJ 2010\1423], entre otras. En relación con las características que tiene que tener la prueba para poder enervar la verdad interina vid. GASCÓN INCHAUSTI, F., *Derecho procesal penal. Materiales para el estudio*, pp. 230-240. Disponible en: https://eprints.ucm.es/id/eprint/62310/

Los parámetros del algoritmo incorporan una serie de valoraciones que podrían afectar esta función de la presunción de inocencia. El algoritmo podría suponer que se parta de un determinado estado ideal que se da interinamente por cierto —v.gr. los afroamericanos delinquen más—.[46] Aunque el resultado que arroje un sistema de IA es (debería ser) un indicio más para la actuación humana (judicial o policial), la realidad demuestra que en la práctica se puede llegar a producir una cierta delegación o automatismo en la toma de decisiones. [47]

La afectación a la presunción de inocencia podría venir así de la mera existencia de los parámetros del algoritmo que incorporan verdades interinas —se «parte de»— o de la imposibilidad de no enervarlas porque no se conocen —no se sabe «de qué se parte»—. Esta afectación dependerá de la concreta fase procesal en la que se utiliza el sistema de IA —medidas de investigación, cautelares o en el juicio— y de la proyección que de la presunción de inocencia se predique en cada caso — v.gr. con escaso desarrollo por el TC respecto de la adopción de medidas cautelares personales —.[48] Como se ve, no es tan fácil relacionar conceptualmente la opacidad del algoritmo con la presunción de inocencia.

> Como explica DIEZ-PICAZO GIMÉNEZ, la presunción de inocencia como regla de tratamiento del investigado en el proceso penal impide que las medidas cautelares personales tengan una finalidad de anticipación de la pena o cualquier otra que no responda a los fines que les son propios (v.gr. 503 LECrim).[49] Es en este contexto en el que podría identificarse una vulneración de la presunción de inocencia: la adopción de las medidas no se basaría en indicios racionales de un *periculum in mora*, sino en determinadas valoraciones que se incorporan en el sistema de IA como verdades interinas.

La imposibilidad de dar respuestas a estas cuestiones con los parámetros doctrinales o jurisprudenciales actuales no debería ser un obstáculo insalvable; por lo menos si se entiende que los derechos fundamentales son un cuerpo vivo que debe adaptarse a los riesgos que para la dignidad del individuo existen en cada momento histórico. Ahora bien, en uno y otro caso, tiene que hacerse un esfuerzo intelectual y creativo de *integrar y relacionar conceptualmente* la transparencia o la existencia del propio algoritmo con el derecho fundamental en cuestión.

3.3. La contradicción de la prueba (art. 47 CDFUE) y la opacidad del sistema de IA

[46] Sobre las verdades interinas, UNGER, J., *System des österreichischen allgemeinen Privatrechts*, Breitkopf und Härtel, Leipzig, 1859, pp. 598-599.
[47] GASCÓN INCHAUSTI, F., "Desafíos para el proceso penal en la era digital", op. cit., pp. 201-205. MARTÍNEZ GARCÍA, E., BORGES BLÁZQUEZ, R., SIMÓ SOLER, E., "Inteligencia artificial y perspectiva de género en la justicia penal", *Diario La Ley*, nº 47, Sección Ciberderecho, 2021, pp. 9-10.
[48] DÍEZ-PICAZO GIMÉNEZ, I., "Comentario del art. 24", op. cit., pp. 117-119.
[49] DÍEZ-PICAZO GIMÉNEZ, I., "Comentario del art. 24", op. cit., p. 118.

Una fórmula «creativa» de hacer frente a la opacidad del algoritmo como uno de aquellos riesgos asociados al uso de la IA en el proceso penal puede encontrase en la reciente jurisprudencia del TJUE sobre la contradicción de la prueba como parte integrante del art. 47 CDFUE.

En las SSTJUE de Gran Sala de 6 de octubre de 2020 [ECLI:EU:C:2020:791] y de 2 de marzo de 2021 [ECLI:EU:C:2021:152] se examina la posible vulneración de derechos fundamentales en el sector de las comunicaciones derivada de la conservación generalizada e indiferenciada de datos de tráfico y de localización y su uso en el proceso penal. El TJUE realiza el juicio de ponderación propio de toda limitación de derechos fundamentales (arts. 7, 8 y 52 CDFUE) y concluye que la licitud de la conservación y su uso como medio de prueba en el proceso penal dependerá de que se persigan delitos graves o que supongan un riesgo para la seguridad nacional.[50]

Una vez identificados los elementos de los que depende su licitud, queda por determinar la cuestión relativa al tratamiento y uso en el proceso penal de la prueba obtenida *ilícitamente* —v.gr. respecto de delitos leves—.

En principio, el TJUE declara que esta cuestión forma parte del ámbito de la autonomía procesal de los EE. MM. Sin embargo, establece un límite derivado del principio de efectividad y declara que «un órgano jurisdiccional que considera que una parte no está en *condiciones de comentar eficazmente* un medio de prueba que pertenece a un ámbito que escapa al conocimiento de los jueces y que puede influir destacadamente en la apreciación de los hechos *debe declarar que existe una violación del derecho a un juicio justo y excluir ese medio de prueba a fin de evitar una violación de esta índole*».[51]

La validez de la prueba obtenida vulnerando derechos fundamentales se reconduce así a la posibilidad de que el acusado «esté en condiciones de comentar eficazmente un medio de prueba» y se integra en el derecho a un juicio justo (art. 47 CDFEU).[52] La consecuencia de esto es que «el principio de efectividad exige al juez penal nacional que *descarte la información y las pruebas que se han obtenido* […] cuando estas personas no estén en condiciones de comentar eficazmente tal información y tales pruebas». Su exclusión no deriva de la vulneración del derecho fundamental a la protección de datos o a la intimidad —o por lo menos no necesariamente, por cuanto es una cuestión regulada por los EE.

[50] Sentencia del Tribunal de Justicia (Gran Sala) de 2 de marzo de 2021 [ECLI:EU:C:2021:152], p. 50.
[51] STJUE (Gran Sala) de 6 de octubre de 2020 [ECLI:EU:C:2020:791], párr. 226. STJUE (Gran Sala) de 2 de marzo de 2021 [ECLI:EU:C:2021:152], párr. 44.
[52] Sentencia del Tribunal de Justicia (Gran Sala) de 2 de marzo de 2021 [ECLI:EU:C:2021:152]: «La necesidad de excluir la información y las pruebas obtenidas incumpliendo lo dispuesto en el Derecho de la Unión debe apreciarse atendiendo, en particular, al riesgo que la admisibilidad de dicha información y de dichas pruebas supone para el respeto del principio de contradicción y, por lo tanto, del derecho a un juicio justo».

MM.—; sino de la afectación que supone para el derecho a un juicio justo la imposibilidad de contradecir eficazmente la prueba.

De esta manera, es posible encontrar en el desarrollo de la prueba ilícita por parte del TJUE una forma de excluir cualquier elemento probatorio que derive directa o indirectamente de un sistema de IA opaco que impida a una parte estar en «condiciones de comentar eficazmente» su resultado. Esta línea jurisprudencial permitiría relacionar conceptualmente el derecho a un proceso justo con la prueba en el proceso penal obtenida a través de sistemas de IA.

> Desde una perspectiva nacional, en las garantías procesales que se integran en el derecho a la libertad puede encontrarse también otra forma de relacionar la opacidad del algoritmo con el derecho de defensa. El art. 520.2 d) LECrim, cuyo origen está en el art. 7 de la Directiva 2012/13/UE, reconoce el «[d]erecho a acceder a los elementos de las actuaciones que sean esenciales para impugnar la legalidad de la detención o privación de libertad». El TC ha desarrollado este derecho y afirma que su «finalidad inmediata [es] poder impugnar de forma efectiva la legalidad de la detención o prisión […] la contradicción e igualdad de armas implican de forma necesaria la previa información sobre los motivos de la privación de libertad y, muy especialmente, el acceso a las actuaciones esenciales para valorar la legalidad de la privación de libertad».[53] Con base en ello podría sostenerse que la opacidad del algoritmo impide «impugnar de forma efectiva la legalidad de la detención o prisión» y que el desconocimiento de las verdades interinas impide a las partes contradecir su resultado. El problema es que estamos ante una garantía procesal que el TC ha integrado en el derecho a la libertad (art. 17 CE) y no en el derecho a la tutela judicial efectiva (art. 24 CE), por lo que su eficacia práctica estaría reducida a la prisión provisional.

3.4. La externalización de la investigación penal: la eficacia de los derechos fundamentales entre particulares

Como explica GASCÓN INCHAUSTI, en la actualidad estamos ante un fenómeno de *externalización penal* en el que los particulares toman protagonismo en la investigación y persecución del delito. Una manifestación de esta externalización es la ejecución material de medidas de investigación electrónica por parte de empresas de telecomunicaciones o las investigaciones internas empresariales en aplicación de programas de cumplimiento penal.[54]

> En este ámbito se enmarca el deber de colaboración que incluye el art. 7 de la Ley Orgánica 7/2021 que transpone a nuestro ordenamiento la Directiva (UE) 2016/680 conforme al cual «[l]as Administraciones públicas, *así como cualquier persona física o jurídica*, proporcionarán a las autoridades judiciales, al Ministerio Fiscal o a la Policía Judicial los datos, informes, antecedentes y justificantes que les soliciten y que sean necesarios para la

[53] STC num. 180/2020 de 14 diciembre [ECLI:ECLI:ES:TC:2020:180].
[54] GASCÓN INCHAUSTI, F., "Desafíos para el proceso penal en la era digital", op. cit., pp. 193-199.

> investigación y enjuiciamiento de infracciones penales o para la ejecución de las penas».

En ese caso, el análisis técnico se desplaza a la eficacia de los derechos fundamentales entre particulares.[55] Lo contrario supondría aceptar la posibilidad de «que el Estado se aproveche de la externalización para dar cabida, por la puerta trasera, a una rebaja de estándares de protección de los derechos fundamentales, con el argumento de su diferente alcance en el ámbito de las relaciones entre particulares».[56]

Los derechos fundamentales son un sistema que garantiza al individuo un ámbito en el que su personalidad puede desarrollarse libremente (art. 10.1 CE). Y con este propósito deben adaptarse a una realidad histórica en la que las principales amenazas a la dignidad de la persona pueden venir de otros individuos. Ahora bien, como se sabe, los derechos fundamentales no se irradian ni con la misma eficacia ni de la misma forma frente al poder público y frente a los particulares. El análisis que ahora se propone pasa por identificar parcelas concretas de externalización penal y, en su caso, la irradiación que directa o indirectamente tienen los derechos fundamentales de naturaleza procesal horizontalmente.

El derecho a la tutela judicial efectiva, a la defensa o a la presunción de inocencia son derechos de prestación o de no interferencia cuyo sujeto pasivo es el Estado y, en consecuencia, solo pueden ser satisfechos y vulnerados por él. Por ello, en principio la irradiación de estos derechos entre particulares es nula.[57]

Ahora bien, como se ha dicho, los derechos fundamentales de naturaleza procesal deben adaptarse a su realidad histórica. Por un lado, deben incorporar prestaciones que en un mundo digital son imprescindibles para garantizar la efectividad de la tutela judicial o el derecho de defensa. Entre ellas, el acceso a fuentes de prueba o la preservación de la función de la presunción de inocencia cuando se aplica un sistema de IA (*vid. supra*). Por el otro, deben adecuarse las garantías de organización y de procedimiento y las obligaciones de optimización y de protección que derivan de su dimensión objetiva y que imponen al Estado adoptar aquellas medidas que garanticen su eficacia.[58] Entre ellas,

---

[55] En relación con la eficacia de los derechos de la CDFUE entre particulares vid. JARASS, G., "Art. 51" en *Charta der Grundrechte*, op. cit., Rn. 36-43.

[56] GASCÓN INCHAUSTI, F., "Desafíos para el proceso penal en la era digital", op. cit., p. 198.

[57] Es erróneo considerar que en el ámbito laboral el principio de indemnidad es una manifestación del efecto horizontal del derecho a la tutela judicial efectiva. Los particulares no pueden vulnerar el derecho a la tutela judicial efectiva. En su caso, pueden perturbar su ejercicio y, por eso, vulnerar la libertad genérica de otros individuos de solicitar la prestación garantizada por el derecho al Estado (arts. 10.1 y 17 CE).

[58] Según ALEXY, los derechos a acciones positivas del Estado pueden clasificarse en *(i)* derechos de protección, *(ii)* derechos a la organización y procedimiento y *(iii)* derechos prestacionales en sentido estricto. ALEXY, R., *Teoría de los derechos fundamentales* (Trad. BERNAL PULIDO, C.), Centro de Estudios Políticos y Constitucionales, 2ª ed., Madrid, 2007, p. 393. Las garantías de organización y procedimiento son aquellas que tienen como objeto crear «procesos u órganos cuya función es la de hacer posible la realización efectiva» del derecho. BASTIDA FREJEIDO, F. J., *et al.*, *Teoría General de los Derechos Fundamentales en la Constitución Española de 1978*, Tecnos, Madrid, 2004, p. 28. Los derechos

deben incorporarse mecanismos legales que aseguren su ejercicio y disfrute cuando estos puedan ser impedidos u obstaculizados por otros particulares en posición de poder.[59]

Por ello, y aunque los derechos fundamentales de naturaleza procesal no tengan eficacia horizontal, la externalización de la investigación penal supone para el Estado la obligación de incorporar herramientas legales que garanticen su ejercicio y disfrute en esta nueva realidad. Se estaría así ante una eficacia entre particulares indirecta o mediatizada por los remedios legales creados y aplicados por el legislador y los tribunales. Esta es una cuestión que desde luego todavía no es abordada frontalmente ni por el legislador europeo ni nacional.

3.5. Transparencia y protección jurídica de la IA

De nada sirve repetir el mantra de la transparencia y accesibilidad del algoritmo si no se invierten esfuerzos en identificar y erradicar las causas de esa oscuridad.[60] Los algoritmos, en tanto que uno de los elementos que forman los sistemas de IA, no son transparentes porque es justamente su condición de «secreto» lo que en la práctica los dota de protección jurídica. Como se sabe, existen graves obstáculos para proteger jurídicamente *todos* los elementos de un sistema de IA a través de derechos de propiedad industrial o intelectual. [61] Incluso aceptado esta posibilidad, muchas veces el régimen de protección legal es ineficiente o poco atractivo para los creadores.

En este contexto, en la práctica la forma más segura y eficiente —en algunos casos la única— de proteger jurídicamente un algoritmo y, con ello, el núcleo del sistema de IA es a través de la figura de los secretos empresariales (art. 1 Ley 1/2019). Aunque es evidente que los programas de código abierto (*open source*) son la mejor opción para asegurar la transparencia, estos no son siempre una opción real. Con el esquema actual, exigirles transparencia a los desarrolladores o usuarios de IA supone en algunos casos imponerles una renuncia a la protección que les otorga el propio sistema al mantener en secreto un activo intangible.

---

de protección deben entenderse como «los derechos del titular de derecho fundamental frente al Estado para que éste lo proteja de intervenciones de terceros». ALEXY, R., *Teoría*, op. cit., pp. 400-408.

[59] *El ejercicio* de los derechos puede ser impedido u obstaculizado fáctica o jurídicamente. No hay duda de que aquellos de naturaleza procesal también pueden serlo por otros particulares. Las referidas garantías y obligaciones que derivan de la dimensión objetiva del derecho fundamental impone al Estado articular herramientas que remedien estos peligros.

[60] *Artificial Intelligence and Fundamental Rights,* Fundamental Rights Agency, p. 77.

[61] Existen inconvenientes para conseguir la protección jurídica de todos los elementos de un sistema de IA. Conforme al art. 4 de la Ley de Patentes y al art. 52 del Convenio sobre concesión de Patentes Europeas, no son patentables los «métodos matemáticos». La Directiva 2009/24/CE sobre la protección jurídica de programas de ordenador determina claramente que «en la medida en que la lógica, los algoritmos y los lenguajes de programación abarquen ideas y principios, estos últimos no están protegidos con arreglo a la presente Directiva». En esta línea, el art. 96 de la Ley de Propiedad Intelectual no protege «las ideas y principios en los que se basan cualquiera de los elementos de un programa de ordenador incluidos los que sirven de fundamento a sus interfaces». Esto hace que en la práctica la mejor forma —y en algunos casos única— de proteger todos los elementos de un sistema de IA sea a través de la Ley de Secretos Empresariales o la Ley de Competencia Desleal.

El discurso y futura regulación debe superar el problema desde su origen. Si se quiere innovación e inversión del sector privado en sistemas de IA y al mismo tiempo transparencia, los esfuerzos deben centrarse en crear un sistema de protección que imponga la publicidad de su contenido.

Consciente de la dificultad de lo anterior, también es razonable plantearse si no se está sobredimensionando la importancia de tener acceso al algoritmo y si no es posible articular algún otro mecanismo que supere los problemas que genera su opacidad. Como lo hace la Propuesta de Reglamento de la Comisión en su art. 13, es posible plantearse la obligación de que el diseñador ponga a disposición del público algún documento o manual que explique en «lenguaje común» el funcionamiento del sistema de IA —las variables y el peso que cada una tiene en el resultado (*output*)—. No se impondría una transparencia absoluta, pero sí una claridad que permita su fiscalización por parte de las administraciones públicas y de los particulares afectados por él. Un mecanismo que asegure, en palabras del TJUE, las «condiciones [para] comentar *eficazmente* un medio de prueba». En ese caso, tendría que articularse algún procedimiento que permita a la administración pública o al juez durante la certificación o uso del sistema de IA verificar de algún modo que la información del manual es un fiel reflejo del algoritmo.

> Otra opción es mantener el algoritmo en secreto y crear algún mecanismo procesal que permita a las partes acceder a su contenido y les imponga estrictos deberes de confidencialidad [v.gr. como los arts. 283 bis a) a k) LEC]. Ahora bien, en la medida en que los sistemas de IA aspiran a tener una aplicación masiva por parte de los Cuerpos y Fuerzas de Seguridad, siempre se corre el riesgo de que los instrumentos sean utilizados de forma fraudulenta y la información pierda su carácter secreto.[62]

IV. CONCLUSIÓN

El uso de sistemas de IA para investigar y perseguir el delito es ya una realidad en nuestro país. Aunque no deben exagerarse sus aplicaciones reales en el proceso penal, estos sistemas pueden tener incidencia en cuestiones relacionadas con la práctica de diligencias de investigación, medidas cautelares, acceso a fuentes de prueba y valoración de la prueba.

La UE tiene la intención de regular el uso de la IA en el sector privado y público desde un enfoque caracterizado por el respecto de los derechos y valores fundamentales. Aunque son proclamaciones esperables que forman parte del discurso político, es exigible una mayor concreción y un análisis técnico, especialmente en un sector en el que Estado

[62] Este es el caso v.gr. de los *copyright trolls* que adquieren derechos de propiedad intelectual limitados para, a través de procedimiento judiciales, acceder información de terceros (direcciones IP). En este sentido son de interés las Conclusiones del Abogado General Spuznar de 17 de diciembre de 2020 en el asunto C-597/19 [ECLI:EU:C:2020:1063].

tiene la doble condición de regulador y usuario y en el que se espera una potente inversión de fondos públicos en los próximos años.

No es suficiente con declaraciones generales sobre el respeto de los derechos fundamentales que, aunque bienvenidas, deben considerarse siempre un punto de partida de toda propuesta de regulación a nivel europeo. Debe concretarse desde una perspectiva europea y nacional qué dimensión de los derechos fundamentales está en riesgo y si existen limitaciones legítimas a su ejercicio. Además, el fenómeno de la externalización en la investigación penal impone examinar la cuestión desde la eficacia de los derechos fundamentales entre particulares. La identificación de que existe un riesgo real de vulneración de derechos fundamentales en el uso de la IA en el proceso penal es lo suficientemente importante como para que se lleve a cabo una regulación adecuada y técnicamente detallada que sea proporcional a la gravedad de los usos.