DIRECTOR
David Vallespín Pérez

COORDINADOR
José María Asencio Gallego

COLECCIÓN DE ORGANIZACIÓN JUDICIAL

DAVID VALLESPÍN PÉREZ
ORGANIZADOR

INTELIGENCIA ARTIFICIAL Y PROCESO

EFICIENCIA VS. GARANTÍAS

COLABORADORES

César A. Villegas Delgado
Daniele Chiappini
David Vallespín Pérez
Diego Palomo Vélez
Diego Valdés Quinteros
Jordi Delgado Castro
José María Asencio Gallego
Leticia Fontestad Portalés
Lluís Boada i Guirao
Macarena Yáñez Ruiz
María Ángeles Pérez Marín
Nancy Carina Vernengo Pellejero
Noemí Jiménez Cardona
Pilar Martín Ríos
Raúl Carnevali Rodríguez
Roberto Cippitani
Santiago García Miguel
Valentina Colcelli

JURUÁ
EDITORIAL

La presente monografía sobre "Inteligencia Artificial y Proceso (Eficiencia vs Garantías)" es fruto de la participación colectiva de un elenco de figuras relevantes de la academia, la abogacía y la magistratura que, en el curso académico 2021/2022, bajo la supervisión científica del Prof. Dr. David Vallespín y la financiación del Observatori de Dret Públic (IDP Barcelona), han enfrentado el análisis de los principales retos y desafíos que plantea la aplicación de la IA en el ámbito de la Administración de Justicia.

El problema no es tanto que un robot pueda sustituir al juez humano (de hecho, ya existen diferentes experiencias que así lo demuestran), sino más bien tener claro si dicha sustitución nos va a abocar, para bien, a una sociedad más justa o si, por el contrario, acabará por situarnos ante un escenario de dictadura digital. En este contexto, parece que el futuro inmediato de la aplicación de la IA en la Administración de Justicia deberá pasar por la consecución de un justo y nada fácil equilibrio entre la eficiencia procesal y el respeto del modelo constitucional de juicio justo. Futuro en el que la presente obra colectiva enfrenta el análisis de la robotización de la valoración probatoria, los juicios telemáticos, los testigos virtuales y online, los smart contracts, los títulos ejecutivos electrónicos e inteligentes, la gobernanza algorítmica de la era digital, la mediación avatar, la gestión eficiente de los "puertos inteligentes", las exigencias éticas de la UE en materia de IA, la rendición de cuentas, la motivación de las resoluciones judiciales, la cibercriminalidad informática y las implicaciones de la IA en orden a la formación de la jurisprudencia.

Una obra, por tanto, escrita hoy, pero pensando en un futuro inmediato que, a buen seguro, puede resultar de utilidad a los estudiosos del Derecho en el marco de la revolución industrial 4.0 (constitucionalistas, procesalistas, civilistas, penalistas, administrativistas, mercantilistas e internacionalistas), así como también a todos aquellos profesionales del Derecho (en particular, los abogados) que deberán manejarse ante un escenario, con más o menos intensidad, de "robotización" judicial.

IDP OBSERVATORI DE DRET PÚBLIC

EDITORIAL JURUÁ Desde 1969 BRASIL

ISBN: 978-989-712-909-4

UTILICE EL LECTOR DE **QR** CODE DE SU TELÉFONO Y CONOZCA OTROS TÍTULOS

Cubierta: Bianca de Triana Franco

Imagen de la Cubierta: SHARP, William. O Juri Atento - 1905

COLECCIÓN DE ORGANIZACIÓN JUDICIAL

DIRECTOR: DAVID VALLESPÍN PÉREZ

COORDINADOR: JOSÉ MARÍA ASENCIO GALLEGO

INTELIGENCIA ARTIFICIAL Y PROCESO

EFICIENCIA VS. GARANTÍAS

Visite nuestra página *web*
www.editorialjurua.com
e-mail: *internacional@jurua.net*

La presente obra ha sido aprobada por el Conselho Editorial Científico de Editorial Juruá, adoptándose el sistema *blind view* (evaluación a ciegas). La evaluación innominada garantiza la exención e imparcialidad del cuerpo de juzgadores y la autonomía del Conselho Editorial, según las exigencias de las agencias e instituciones de evaluación, certificando la excelencia del material que publicamos y presentamos a la sociedad.

JURUÁ
EDITORIAL

Europa – Rua General Torres, 1.220 – Lojas 15 e 16 – Tel: +351 223 710 600
Centro Comercial D'Ouro – 4400-096 – Vila Nova de Gaia/Porto – Portugal
Brasil – R. Flávio Dallegrave, 7.665 – São Lourenço – Tel: +55 (41) 4009-3900
CEP: 82.210-310 – Curitiba – Paraná – Brasil

ISBN: 978-989-712-909-4

Depósito Legal: 513035/23

Editores: Luiz Augusto de Oliveira Junior
Francine Marie Carvalho de Oliveira

I61 Inteligencia artificial y proceso: eficiencia vs. garantías./ organização de David Vallespín Pérez./ Curitiba: Juruá, 2023.
280p.; 21cm (Colección de Organización Judicial)

1. Inteligência artificial. 2. Garantia. I. Vallespín Pérez, David (org.).

CDD 343.09944 (22.ed)
CDU 34:681.3

Datos Internacionales de Catalogación en la Publicación
Bibliotecaria: Maria Isabel Schiavon Kinasz, CRB9 / 626

David Vallespín Pérez
Organizador

INTELIGENCIA ARTIFICIAL Y PROCESO

EFICIENCIA VS. GARANTÍAS

Colaboradores:

César A. Villegas Delgado
Daniele Chiappini
David Vallespín Pérez
Diego Palomo Vélez
Diego Valdés Quinteros
Jordi Delgado Castro
José María Asencio Gallego
Leticia Fontestad Portalés
Lluís Boada i Guirao
Macarena Yáñez Ruiz
María Ángeles Pérez Marín
Nancy Carina Vernengo Pellejero
Noemí Jiménez Cardona
Pilar Martín Ríos
Raúl Carnevali Rodríguez
Roberto Cippitani
Santiago García Miguel
Valentina Colcelli

Porto
Editorial Juruá
2023

ASPECTOS ÉTICOS DEL USO DE LA INTELIGENCIA ARTIFICIAL EN LA TOMA DE DECISIONES EN EL DERECHO DE LA UNIÓN EUROPEA

Roberto Cippitani[1]

Resumen: *El uso creciente de la inteligencia artificial (en adelante, IA) está generando muchos interrogantes de naturaleza ético-jurídica, como los que se refieren al uso de la IA para tomar decisiones que tienen una gran relevancia jurídica (en el ámbito sanitario, contractual, administrativo, judicial, etc.). Las cuestiones asociadas a la IA derivan de cómo el Estado de Derecho puede enfrentar los retos que surgen de la «tecno-ciencia».*

En el ámbito del antemencionado enfoque jurídico y cultural, en Europa, especialmente en la Unión Europea, se está elaborando una normativa que trata de enfrentar las cuestiones que surgen del uso de la inteligencia artificial en la toma de decisiones jurídicamente relevantes. En particular en el caso de la IA, el objetivo de la Unión Europea es construir un marco de reglas específicas y de principios éticos generales de manera que las nuevas tecnologías se puedan desarrollar de manera coherente con los derechos e intereses fundamentales.

Palabras-clave: *Inteligencia Artificial, Estado de Derecho, Principios de ética de la tecno-ciencia.*

Sumario*: I. Documentos institucionales y normativa de la Unión Europea en materia de inteligencia artificial. II. Toma de decisiones jurídicamente relevante e inteligencia artificial. III. Una primera casuística sobre los riesgos que derivan del uso de la IA en la toma de*

1 Profesor de Bioderecho, Catedrático Jean Monnet, Università degli Studi di Perugia, Departamento de Derecho; Investigador asociado al Consiglio Nazionale delle Ricerche, IFAC; Profesor Profesor de INDEPAC - Instituto de Estudios Superiores en Derecho Penal (México) y de la Escuela Judicial del Poder Judicial del Estado de Oaxaca (México).

decisiones. IV. Principios éticos generales. V. Reglas específicas en materia de la toma de decisiones basadas en la inteligencia artificial. VI. Eticidad y respeto de los derechos e intereses fundamentales. VII. Legislación sobre la IA y colaboración transdisciplinaria. VIII. Bibliografía.

I DOCUMENTOS INSTITUCIONALES Y NORMATIVA DE LA UNIÓN EUROPEA EN MATERIA DE INTELIGENCIA ARTIFICIAL

La Unión Europea está elaborando un marco ético-jurídico como respuesta a uno de los temas tecnológicos más problemáticos, es decir el amplio uso hoy de la Inteligencia Artificial (en adelante «IA»)[2].

En la IA como en otros sectores la UE es pionera en la reglamentación de la tecnociencia[3]. La normativa sobre el tema de la IA se está desarrollando en base al debate y a la reflexión institucional a nivel continental y no solo en el seno de cada país europeo.

La cuestión de la utilización de algoritmos y en general de la inteligencia artificial son el objeto de documentos como la Resolución del Parlamento Europeo del 16 de febrero de 2017 con recomendaciones para la Comisión sobre «normas de Derecho civil sobre robótica»; la Comunicación de la Comisión Europea del 25 de abril de 2018 sobre «Inteligencia Artificial para Europa» (COM(2018) 237 final); la Comunicación de la Comisión Bruselas «Generar confianza en la inteligencia artificial centrada en el ser humano» del 8 de abril de 2019, COM(2019) 168 final y, más recientemente, en el Libro Blanco sobre Inteligencia Artificial del 19 de febrero de 2020 (COM(2020) 65 final).

En el marco del Consejo de Europa, el Comité de Expertos en Intermediarios de Internet publicó el estudio Algoritmos y Derechos Humanos en marzo de 2018 y poco después, el 3 de diciembre de 2018, la Comisión Europea para la Eficiencia de los Sistemas de Justicia (CEPEJ) aprobó la «Carta ética sobre la utilización de la inteligencia artificial en los sistemas de justicia y su entorno».

[2] En realidad, el término fue adoptado en 1956, por un grupo de científicos (McCarthy, Minsky, Rochester y Shannon) involucrados en el proyecto de investigación "Inteligencia Artificial" del Dartmouth College en los Estados Unidos. Vid. McCarthy et al. (31de agosto de 1955) "A Proposal for the Dartmouth Summer Research project on Artificial Intelligence". Universidad de Stanford Recuperado de: https://web.archive.org/web/20080930164306/ http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html

[3] En los Estados Unidos, en cambio, aunque el uso de algoritmos en las materias más dispares está bien desarrollado, no es posible encontrar reglamentos generales sino solo normas/prácticas individuales que se aplican a nivel local, como, por ejemplo, la adoptada por el Consejo de Nueva York, el 11 de diciembre de 2017, Ley Local en relación con los sistemas automatizados de decisión utilizados por los organismos.

Este debate institucional, que está en continua evolución, es la base para la adopción de instrumentos legales específicos o para la aplicación de normas y principios generales.

Desde este punto de vista, a los sistemas de IA se pueden aplicar otras fuentes legislativas de la Unión Europea, como las que se refieren a la protección de los datos personales (véase el artículo 22 del Reglamento (UE) n°2016/679); a la libre circulación de datos no personales[4]; a las directivas contra la discriminación[5], a la denominada directiva de máquinas[6] , a la directiva sobre responsabilidad del productor[7], a la directiva sobre derecho del consumidor[8] y a las directivas sobre salud y seguridad en el trabajo[9].

Además, la Comisión ha presentado recientemente una propuesta de Reglamento (del Parlamento europeo y del Consejo) por el que se establecen «normas armonizadas en materia de inteligencia artificial (ley de inteligencia artificial) y se modifican determinados actos legislativos de la Unión» (COM (2021) 206 final del 21 de abril 2021).

II TOMA DE DECISIONES JURÍDICAMENTE RELEVANTE E INTELIGENCIA ARTIFICIAL

En los documentos oficiales, la IA se considera como «sistemas que manifiestan un comportamiento inteligente, pues son capaces de analizar su

[4] Reglamento (UE) 2018/1807 del Parlamento Europeo y del Consejo, del 14 de noviembre de 2018, relativo a un marco para la libre circulación de datos no personales en la Unión Europea.

[5] Directiva 2000/43/CE: aplicación del principio de igualdad de trato de las personas independientemente de su origen racial o étnico; Informe de la Comisión al Consejo y al Parlamento Europeo – Aplicación de la Directiva 2000/43/CE, del 29 de junio de 2000, relativa a la aplicación del principio de igualdad de trato de las personas independientemente de su origen racial [COM(2006) 643 final del 30.10.2006]; Recomendación del Consejo, del 9 de diciembre de 2013, relativa a la adopción de medidas eficaces de integración de los gitanos en los Estados miembros (DO C 378 de 24.12.2013, pp. 1-7); Directiva 2000/78/CE del Consejo, del 27 de noviembre de 2000, relativa al establecimiento de un marco general para la igualdad de trato en el empleo y la ocupación ("Directiva de igualdad en el empleo"). [COM(2014) 2 final del 17.1.2014]

[6] Directiva 89/392/CEE del Consejo, del 14 de junio de 1989, relativa a la aproximación de las legislaciones de los Estados miembros sobre máquinas.

[7] Directiva 85/374/CEE del Consejo, del 25 de julio de 1985, relativa a la aproximación de las disposiciones legales, reglamentarias y administrativas de los Estados Miembros en materia de responsabilidad por los daños causados por productos defectuosos.

[8] Directiva (UE) 2019/771 del Parlamento Europeo y del Consejo del 20 de mayo de 2019 relativa a determinados aspectos de los contratos de compraventa de bienes, por la que se modifican el Reglamento (CE) n°2017/2394 y la Directiva 2009/22/CE y se deroga la Directiva 1999/44/CE Directiva 2011/83/UE del Parlamento Europeo y del Consejo, del 25 de octubre de 2011 , sobre los derechos de los consumidores, por la que se modifican la Directiva 93/13/CEE del Consejo y la Directiva 1999/44/CE del Parlamento Europeo y del Consejo y se derogan la Directiva 85/577/CEE del Consejo y la Directiva 97/7/CE del Parlamento Europeo y del Consejo.

[9] Directiva 89/391/CEE del Consejo, del 12 de junio de 1989, relativa a la aplicación de medidas para promover la mejora de la seguridad y de la salud de los trabajadores en el trabajo.

entorno y pasar a la acción –con cierto grado de autonomía– con el fin de alcanzar objetivos específicos»[10].

Esta capacidad de las tecnologías de IA se manifiesta en decisiones autónomas o en sugerir decisiones.

De hecho, sistemas de IA generan «información de salida como contenidos, predicciones, recomendaciones o decisiones que influyan en los entornos con los que interactúa» (artículo 4, n°1, propuesta de Reglamento del 2021).

Las decisiones tomadas con el apoyo de sistemas de IA pueden ser más eficientes y útiles a las personas y a la comunidad especialmente en el sector público.

Por ejemplo, como destaca el documento Plan coordinado sobre la IA de la Comisión Europea del 2018 (véase el Anexo, apartado G): «La IA puede mejorar la interacción ciudadano-administración pública a través de sistemas conversacionales (incluidos asistentes digitales y "chatbots" de la administración), servicios multilingües y traducción automática». Especialmente eso puede ocurrir en el «sector social y sanitario, para apoyar la toma de decisiones de los médicos o para respaldar el reconocimiento precoz de la marginación de los jóvenes». De otra parte, en el caso de la financiación pública «la decisión habilitada por IA puede simplificar la relación entre autoridades y beneficiarios a través de la integración de un interés público más amplio o consideraciones regulatorias en la toma de decisiones diaria (a través de comunicación dirigida, sugerencias de comportamiento, etc.)».

Sin embargo, como subraya el Libro Blanco sobre la Inteligencia Artificial, la IA, así como sucede con toda nueva tecnología, su uso presenta tanto oportunidades como amenazas (párr. 5).

La IA incrementa las posibilidades de hacer un seguimiento y un análisis de las costumbres cotidianas de las personas y, por tanto, existe «las autoridades estatales y otros organismos recurran a la IA para la vigilancia masiva, o las empresas la utilicen para observar cómo se comportan sus empleados». Además, recuerda la Comisión «El tratamiento de los datos, el modo en el que se diseñan las aplicaciones y la envergadura de la intervención humana pueden afectar a los derechos de libertad de expresión, protección de los datos personales, privacidad y libertad política».

Aunque los prejuicios y la discriminación sean inherentes a toda tecnología, en el caso de la IA, los efectos negativos pueden ser más amplios.

Otro aspecto negativo puede ser la opacidad («efecto caja negra»), la complejidad, la imprevisibilidad y un comportamiento parcialmente autónomo, pueden hacer difícil verificar el cumplimiento de la legislación de la UE sobre la

[10] Véase el Plan coordinado sobre la inteligencia artificial, cit., párr. 1; véase también la Comunicación, Inteligencia artificial para Europa, cit. párr. 1.

protección de los derechos fundamentales e impedir su cumplimiento efectivo. Eso limita el acceso efectivo a la justicia y la posibilidad para establecer las consecuencias jurídicas como la responsabilidad civil en cuanto «Es posible que las personas que hayan sufrido daños no dispongan de un acceso efectivo a las pruebas necesarias para llevar un caso ante los tribunales, por ejemplo, y tengan menos probabilidades de obtener una reparación efectiva en comparación con situaciones en las que los daños sean causados por tecnologías tradicionales. Estos riesgos aumentarán a medida que se generalice el uso de la IA» (Libro Blanco sobre la inteligencia artificial, cit., párr. 5).

El riesgo del uso de la IA aumenta cuando los sistemas diagnósticos y de apoyo a las decisiones humanas se refieren a la vida y la salud. En estos casos «La magnitud de las consecuencias adversas de un sistema de IA para los derechos fundamentales protegidos por la Carta es particularmente pertinente cuando este es clasificado como de alto riesgo». (véase el «considerando» 28 propuesta de Reglamento del 2021)

III UNA PRIMERA CASUÍSTICA SOBRE LOS RIESGOS QUE DERIVAN DEL USO DE LA IA EN LA TOMA DE DECISIONES

Ya algunas sentencias muestran en concreto los riesgos del uso de IA en la toma de decisiones de relevancia jurídica.

Por ejemplo, es el caso de las sentencias sobre los algoritmos COMPAS (Correctional offender management profiling for alternative sanctions) (Loomis vs. Wisconsin, en los Estados Unidos)[11] y HART (Harm Assessment Risk Tool, en el Reino Unido)[12] capaces de calcular la probabilidad de reincidencia de los crímenes. En dichos casos los jueces han establecido que los algoritmos llevaban a decisiones discriminatorias porque estaban basadas en índices sociales (como el desempleo, la falta de disponibilidad de vivienda y el abuso de sustancias, etc.) más difusos entre grupos con mayor tendencia a una conducta criminal. Eso lleva a predecir la probabilidad de que personas con antecedentes penales similares y que pertenecen a mismos grupos sociales puedan cometer un nuevo delito una vez que queden liberados de prisión.

[11] State vs. Loomis, 881 N.W.2d 749 (Wis. 2016). Vid. los comentarios de Simoncini, A., *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, in *BioLaw Journal*, 1/2019; Donati, F., *Intelligenza Artificiale e Giustizia*, in *Rivista AIC*, 1/2020; Muciacca, N., *Algoritmi e procedimento decisionale: alcuni recenti arresti della giustizia amministrativa*, in *Federalismi.it*. 10/2020; Cavallaro M.C. y Smorto, G., *Decisione pubblica e responsabilità dell'amministrazione nella società dell'algoritmo*, en *Federalismi.it*. 16/2019.

[12] OSWALD, M.; GRACE, J.; URWIN, S. Y BARNES, G.C., *Algorithmic Risk Assessment Policing Models: Lessons from the Durham HART Model and "Experimental" Proportionality*, en *Information & Communications Technology Law*, n. 27/2018, p. 223.

Otro caso es el de "Syri" ("*System Risk Indication*")[13], a través del cual el gobierno de los Países Bajos, desde 2014 hasta el 5 de febrero de 2020, establecía el riesgo de cometer fraude o abuso por parte de los beneficiarios de subsidios y de otras formas de asistencia pública. El sistema consideraba más riesgosos los beneficiarios de subvenciones que vivían en barrios con una alta densidad de residentes de bajos ingresos, migrantes y minorías. El Tribunal de Distrito de La Haya, llamado a verificar la legitimidad del uso de este algoritmo por parte de varias asociaciones no gubernamentales holandesas, opinó que Syri[14] no era trasparente y discriminaba injustificadamente a los ciudadanos menos pudientes.

En otros casos se ha observado que la decisión administrativa derivada del algoritmo no es siempre transparente y sin embargo puede afectar de manera razonable a los intereses de las personas.

Esto, por ejemplo, se ha verificado en relación con las decisiones sobre la movilidad de los profesores en Italia. La gestión de los procedimientos de movilidad mencionados se había realizado mediante el uso de un software desarrollado por una empresa (HPE Services s.r.l.) contratada por el Ministerio Italiano de Educación.

Algunos profesores se quejaron de que se les había asignado a provincias más alejadas de su lugar de residencia o de la indicada como su opción prioritaria, aunque había puestos disponibles en esas provincias.

Dichas decisiones constituyeron el objeto de unas sentencias de la justicia administrativa, especialmente de algunas recientes del Consejo de Estado italiano (véase Consiglio di Stato, Sez. VI, sentencias n°2270 y 8472/2019) que han representado la ocasión para identificar las directivas que deberían ser cumplidas por las administraciones públicas cuando se sirven de algoritmos[15].

IV PRINCIPIOS ÉTICOS

Para evitar los problemas puestos por IA, especialmente en la toma de decisiones jurídicamente importantes, la IA debe ser lícita, es decir cumplir las

[13] Asunto NJCM c.s/De Staat der Nederlanden (SyRI) before the District Court of The Hague (case number: C/09/550982/HA ZA 18/388).

[14] La decisión del Tribunal de Distrito se basó, además, en la intervención como *amicus curiae* de Philip Alston, Relator Especial de las Naciones Unidas. El resumen puede leerse en el siguiente link: https://www.ohchr.org/Documents/Issues/Poverty/Amicusfinalversionsigned.pdf.

[15] Véase Colcelli V., Burzagli L., *Public administration and technology in the time of artificial intelligence: the Italian case of Council State judgment n. 2019 2270*, en Arnold, R. y Danėlienė, I (Eds), *The Concept of Democracy as developed by Constitutional Justice. Constitutional Court of the Republic of Lithuania*, Vilnus, 2020, p. 137-148.

leyes y reglamentos aplicables[16], y también debe ser «ética», en el sentido de que no debe afectar intereses y valores protegidos por el ordenamiento jurídico[17].

Como en otros casos, el sistema jurídico de la Unión debe «reforzar la protección de los derechos fundamentales a tenor de la evolución de la sociedad, del progreso social y de los avances científicos y tecnológicos» (véase el preámbulo de la Carta de los Derechos fundamentales de la Unión Europea).

Lo que significa que la IA debe respetar principios generales y reglas específicas.

En el derecho europeo se pueden identificar algunos principios que sirven para enfrentar los problemas éticos en el ámbito de la ciencia y de la tecnología, como en el caso de la IA utilizada como soporte en la toma de decisiones.

a) Autodeterminación

Las personas involucradas en cualquier tipo de actividad (comercial, científica, sanitaria, etc.) tienen el derecho de expresar su consentimiento libero e informado en aplicación del principio de autodeterminación.

Incluso, cabe destacar que en las constituciones nacionales (véase por ejemplo el artículo 32 de la Constitución Italiana sobre el derecho a la salud), muchas fuentes europeas establecen el derecho al consentimiento como condición para desempeñar actividades concernientes a las personas: el tema está reglado por el artículo 3° párrafo 2 de la Carta UE; en la disciplina europea de la protección de los datos personales (véase el artículo 8 de la Carta de los Derechos Fundamentales de la Unión Europea y el Reglamento (UE) 2016/679); en el Convenio de Oviedo, que establece la reglamentación del consentimiento informado en las actividades en ámbito biomédico (véase especialmente el artículo 5)[18].

Asimismo, la persona tiene el derecho de recibir una información adecuada es decir « *all the necessary information ... that this should address the substantive aspects of the processing that the consent is intended to legitimise* »[19]. El hecho de que se utilicen sistemas de IA en un tratamiento es una de las informaciones que es necesario proporcionar, debido a los potenciales riesgos.

Aunque fuentes que se refieren a la IA no hablan directamente de la necesidad del consentimiento, lo hacen implícitamente cuando por ejemplo

[16] Véase la Comunicación de la Comisión «Generar confianza en la inteligencia artificial centrada en el ser humano», *passim*.

[17] Comisión europea, Comunicación sobre la inteligencia artificial para Europa, COM (2018) 237 final, p. 1.

[18] El derecho al consentimiento importa también al derecho de rechazar el consentimiento en cualquier momento (véase la jurisprudencia TEDH, Evans c. Reino Unido del 10 de abril de 2007).

[19] *Ibidem*.

establecen la obligación del respeto de los derechos fundamentales de las personas inclusa la protección de los datos personales.

El consentimiento previo no será obligatorio cuando el uso de la inteligencia artificial se ponga en marcha en actividades que pueden constituir una excepción como previsto, por ejemplo, en materia de salud pública, en ámbito judiciario o en otros casos en los cuales hay que proteger un interés colectivo.

b) Dignidad

En el derecho europeo a la autodeterminación de la persona está sujeto a otro principio fundamental de carácter ético-jurídico: la dignidad.

En el artículo 2 del Tratado UE y en el preámbulo de la Carta UE se prescribe que todas las actividades tienen que llevarse a cabo respetando la dignidad de las personas.

Como afirmado por el Abogado general Christine Stix-Hackl en sus observaciones (presentadas el 18 de marzo de 2004) concerniente el asunto «Omega» (véase más adelante) «La "dignidad humana" constituye la expresión del máximo respeto y valor que debe otorgarse al ser humano en virtud de su condición humana. Se trata de proteger y respetar la esencia y la naturaleza del ser humano como tal, la "sustancia" del ser humano. Por lo tanto, en la dignidad humana se refleja el propio ser humano; ampara sus elementos constitutivos. Sin embargo, la cuestión de cuáles son los elementos constitutivos de un ser humano remite inevitablemente a un ámbito prejurídico, es decir, en última instancia el contenido de la dignidad humana viene determinado por una determinada "imagen del ser humano"» (apartado 75).

La dignidad es un valor central en el marco ético-jurídico de la UE, en particular en el ámbito de la ciencia y de la tecnología (véase el artículo 3 de la Carta sobre la investigación biomédica) y especialmente en materia de IA (véase por ejemplo el Libro Blanco sobre la inteligencia artificial; véase también el preámbulo de la propuesta de Reglamento)

c) Prevención y precaución

Todos los sujetos, especialmente los profesionales como los científicos, están sujetos al deber de no perjudicar a otras personas. La consecuencia del incumplimiento de tales deberes es la aplicación de las sanciones del derecho civil (por ejemplo, la indemnización de los daños y perjuicios) u otras sanciones previstas por la ley y por el contrato.

Este deber de actuar con precaución es importante sobre todo en actividades como la ciencia y las tecnologías que pueden poner en riesgo la integridad de las personas y otros intereses fundamentales.

En particular, las fuentes jurídicas de la UE contemplan el caso de que se hayan identificado los efectos potencialmente peligrosos derivados de una actividad, pero la evaluación científica no permite determinar el riesgo con suficiente certeza (véase la Comunicación de la Comisión sobre el principio de precaución COM (2000) 1 final).

En este caso, es necesario actuar de acuerdo con el principio de «precaución».

La aplicación del principio de precaución debe comenzar con una evaluación científica, lo más completa posible, y cuando sea posible, identificando en cada etapa el grado de incertidumbre científica (Comisión Europea, 2000, COM (2000) 1 final, párrafo 4).

Por tanto, el principio de precaución es la base de un proceso de toma de decisiones de evaluación del riesgo. Este proceso tiene que identificar el grado de incertidumbre que conlleva la evaluación de la información científica y, a continuación, es necesario establecer el nivel de riesgo «aceptable» para la sociedad que, de todas maneras, no puede afectar a la dignidad y a la integridad de las personas. El riesgo debe estar presente (no es meramente hipotético) y no se puede descartar, aunque no sea seguro.

En el caso especifico de la IA, las fuentes de la UE tienen en cuenta que el uso de la inteligencia artificial puede provocar riesgos desde varios puntos de vista «que pueden estar vinculados a ciberamenazas, a la seguridad personal (por ejemplo, con relación a nuevos usos de la IA, como en el caso de los aparatos domésticos), a la pérdida de conectividad, etc., y pueden existir en el momento de comercializar los productos o surgir como resultado de la actualización de los programas informáticos y del aprendizaje automático del producto cuando este último se está utilizando» (Véase el Libro Blanco sobre la IA, párr. 5).

Según la antemencionada comunicación «Generar confianza en la inteligencia artificial centrada en el ser humano», en el ámbito de este sistema de confianza, los actores (desarrolladores, proveedores y usuarios) deberían seguir algunas directivas, elaborados por un grupo de expertos de alto nivel sobre la IA, nombrados por la Comisión, que representa a toda una serie de partes interesadas, entre las cuales la solidez técnica y la seguridad, que incluye la capacidad de resistencia a los ataques y a la seguridad, un plan de repliegue y la seguridad general, precisión, fiabilidad y reproducibilidad.

La Propuesta de Reglamento antemencionada tiene en consideración el riesgo potencial de la IA y regla sobre todo el caso de las obligaciones de los

proveedores y de los sistemas de «alto riesgo» identificados en el Anexo III de la Propuesta de Reglamento, como los que se refieren a la identificación biométrica y categorización de personas; gestión y funcionamiento de infraestructuras esenciales; educación y formación profesional; empleo, gestión de los trabajadores y acceso al autoempleo; acceso y disfrute de servicios públicos y privados esenciales y sus beneficios.

En particular, se consideran de alto riesgo, ámbitos relacionados a la toma de decisiones en materia civil, penal y administrativa como: asuntos relacionados con la aplicación de la ley (por ejemplo para determinar el riesgo de que una persona cometa una infracción o para evaluar una prueba)[20]; gestión de la migración, el asilo y el control fronterizo[21]; administración de justicia y procesos democráticos (es decir «los sistemas de IA destinados a ayudar a una autoridad judicial en la investigación e interpretación de hechos y de la ley, así como en la aplicación de la ley a un conjunto concreto de hechos»).

[20] En detalle, los sistemas de IA de alto riesgo en materia de aplicación de la ley se consideran a continuación: sistemas de IA para determinar el riesgo de que las personas cometan infracciones penales o reincidan en su comisión, así como el riesgo para las potenciales víctimas de delitos; sistemas de IA destinados a utilizarse por parte de las autoridades encargadas de la aplicación de la ley como polígrafos y herramientas similares, o para detectar el estado emocional de una persona física; sistemas de IA destinados a utilizarse por parte de las autoridades encargadas de la aplicación de la ley para detectar ultra falsificaciones; sistemas de IA destinados a utilizarse por parte de las autoridades encargadas de la aplicación de la ley para la evaluación de la fiabilidad de las pruebas durante la investigación o el enjuiciamiento de infracciones penales; sistemas de IA destinados a utilizarse por parte de las autoridades encargadas de la aplicación de la ley para predecir la frecuencia o reiteración de una infracción penal real o potencial con base en la elaboración de perfiles de personas físicas, de conformidad con lo dispuesto en el artículo 3, apartado 4, de la Directiva (UE) 2016/680, o en la evaluación de rasgos y características de la personalidad o conductas delictivas pasadas de personas físicas o grupos; sistemas de IA destinados a utilizarse por parte de las autoridades encargadas de la aplicación de la ley para la elaboración de perfiles de personas físicas, de conformidad con lo dispuesto en el artículo 3, apartado 4, de la Directiva (UE) 2016/680, durante la detección, la investigación o el enjuiciamiento de infracciones penales; sistemas de IA destinados a utilizarse para llevar a cabo análisis sobre infracciones penales en relación con personas físicas que permitan a las autoridades encargadas de la aplicación de la ley examinar grandes conjuntos de datos complejos vinculados y no vinculados, disponibles en diferentes fuentes o formatos, para detectar modelos desconocidos o descubrir relaciones ocultas en los datos.

[21] Es decir: sistemas de IA destinados a utilizarse por parte de las autoridades públicas competentes como polígrafos y herramientas similares, o para detectar el estado emocional de una persona física; sistemas de IA destinados a utilizarse por parte de las autoridades públicas competentes para evaluar un riesgo, como un riesgo para la seguridad, la salud o relativo a la inmigración ilegal, que plantee una persona física que pretenda entrar o haya entrado en el territorio de un Estado miembro; sistemas de IA destinados a utilizarse por parte de las autoridades públicas competentes para la verificación de la autenticidad de los documentos de viaje y los documentos justificativos de las personas físicas y la detección de documentos falsos mediante la comprobación de sus elementos de seguridad; sistemas de IA destinados a ayudar a las autoridades públicas competentes a examinar las solicitudes de asilo, visado y permisos de residencia, y las reclamaciones asociadas con respecto a la admisibilidad de las personas físicas solicitantes.

c) Proporcionalidad

Otro principio que hay que destacar es el de «proporcionalidad» por el cual la actividad científica debe llevarse a cabo de manera proporcional, es decir a través de medios no excesivos respecto a los fines que se deben alcanzar. Los medios y los fines deben ser legítimos y deben evitar riesgos que pueden afectar los intereses protegidos por el ordenamiento jurídico [22].

La proporcionalidad es un principio considerado también por la jurisprudencia del Tribunal Europeo de los Derechos Humanos[23] que se puede aplicar incluso a los temas biomédicos disciplinado por el Convenio de Oviedo[24].

Los documentos europeos, por ejemplo, afirman que el uso de sistemas de inteligencia artificial para luchar contra la delincuencia es admisible solo si son necesarios y proporcionales[25].

V REGLAS ESPECÍFICAS EN MATERIA DE LA TOMA DE DECISIONES BASADAS EN LA INTELIGENCIA ARTIFICIAL

Asimismo, los principios generales en materia de ética de la ciencia y de la tecnología, el marco normativo que está diseñando la Unión Europea ya permite identificar reglas específicas para enfrentar los problemas que surgen de la aplicación de la IA, como en el caso de las decisiones que tienen una relevancia jurídica.

Dichos principios pueden ser resumidos a continuación:

[22] Se habla de proporcionalidad por ejemplo en la jurisprudencia *Volker* (Tribunal de justicia, 9 de noviembre de 2010, asuntos C-92/09 y C-93/09, *Volker und Markus Schecke y Eifert*, Rec. 2010, pág. I-11063) que se ocupa de la publicación en Internet de los datos de los beneficiarios de subvenciones comunitarias, establecida por reglamento comunitario (véase el artículo 44 *bis* del Reglamento (CE) nº1290/2005). Como el Tribunal argumenta, el hecho de que el tratamiento de los datos esté previsto por una norma legislativa, no significa que no haya necesidad de respetar los principios comunitarios, especialmente el de proporcionalidad, en base al cual «la limitación de los derechos consagrados por los artículos 7 y 8 de la Carta es proporcionada a la finalidad legítima perseguida» (punto 72).

[23] Cf. TEDH, W c. Reino Unido, sent. 8 de julio de 1987, Series A, n°121-A, p. 27, § 60 (b) and (d); Olsson v. Sweden, 24 March 1988, Series A n°130, pp. 31-32, § 67. Por la jurisprudencia del Tribunal de Estrasburgo, véanse, en particular, TEDH, sentencia *Gillow c. Reino Unido*, del 24 de noviembre de 1986, serie A nº109, pár. 55, y la sentencia *Österreichischer Rundfunk y otros*, apartado 83)

[24] Explanatory Report to the Convention on Human Rights and Biomedicine, Paragraph 159. Véase Andorno, R., *The Oviedo Convention: A European Legal Framework at the Intersection of Human Rights and Health Law*, in JIBL, Vol 02, I. 2005, pp. 133-143.

[25] Véase el Libro Blanco sobre la Inteligencia artificial del 2020, cit., párr. 6

a) No exclusividad y no automaticidad

Una decisión (judicial o administrativa) no debe basarse exclusivamente y de manera automática sobre el resultado de un dato técnico o científico, sino pasar por el filtro del decisor (humano).

Se puede citar también, en el Derecho italiano, los artículos 194 y siguientes del Código procesal civil, que prevé que la pericia del experto se pueda admitir como prueba en la medida en que: i) la consulta respete los límites de la investigación impuestos por el juez/decisor; ii) se redacte un informe especial de cada operación realizada; iii) se respete el principio del contrainterrogatorio.

Como expresión de dicho principio de no exclusividad, se puede citar también el artículo 22 del Reglamento de la UE en materia de protección de datos personales que establece, en su primer párrafo que: «Todo interesado tendrá derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado, incluida la elaboración de perfiles, que produzca efectos jurídicos en él o le afecte significativamente de modo similar».

Finalmente se refieren al principio del que se está hablando, las sentencias que consideran los procedimientos de toma de decisiones a través de algoritmos, como la antemencionada jurisprudencia del Consejo de Estado italiano (véase Cons. di Stato, Secc. VI, sentencias n. 2270 y 8472/2019).

El Tribunal Constitucional francés (Conseil Constitutionnel, sentencia nº2018-765 del 12 de junio de 2018)[26] destaca que las decisiones automáticas no se pueden poner en marcha especialmente en el caso de decisión de informaciones «sensibles!» es decir las que se refieren el origen étnico, las creencias religiosas, políticas o filosóficas, la afiliación a un sindicato, los datos sanitarios, genéticos o biométricos o relacionados de otro modo con la inclinación sexual de las personas (véase también el artículo 22, párr. 4, Reglamento (UE) nº2016/679).

La necesidad de una vigilancia humana está prevista en la propuesta de Reglamento sobre la IA (Véase el «considerando» 48 y el artículo 14). Sobre este punto la antemencionada Comunicación de 2019 «Generar cofinancia...» (párr. 2.2.i) destaca cómo «La supervisión humana ayuda a garantizar que un sistema de IA no socave la autonomía humana ni cause otros efectos adversos». La supervisión debe realizarse a través de mecanismos de gobernanza, tales como el enfoque de la participación humana (human-in-the-loop), la supervisión humana (human-on-the-loop), o el control humano (human-in-command). Además, las autoridades tienen que ejercer sus propias competencias de supervisión conforme a sus mandatos.

[26] El texto se puede leer en la página: https://www.conseil-constitutionnel.fr/decision/2018/2018765DC.htm.

b) Transparencia

La legislación vigente requiere que las personas tengan el derecho de conocer cuáles son las razones que fundamentan las decisiones de las autoridades. Este derecho a la trasparencia y a la información completa está previsto en muchas normas como los artículos 13, 14 y 15 del GDPR y del artículo 41 de la Carta Europea de Derechos Fundamentales de la UE. el Plan coordinado sobre la inteligencia artificial del 2018 destaca como los sistemas de IA deben tener en cuenta «que las administraciones públicas están obligadas a actuar según lo prescrito por la ley, que necesitan motivar sus decisiones y que sus actos están sujetos a revisión judicial por los tribunales administrativos» (véase el párr. 2.5 «Crear el espacio de datos europeo es esencial para la IA en Europa, incluso para el sector público»). Además, «los seres humanos deben entender cómo la IA toma decisiones. Europa puede convertirse en un líder mundial en el desarrollo y el uso de la IA para el bien y la promoción de un enfoque centrado en el ser humano y principios de ética por diseño» (Plan coordinado sobre la inteligencia artificial, párr. 2.6, intitulado «Desarrollar directrices de ética con una perspectiva global y garantizar un marco jurídico favorable a la innovación»).

Hay que cumplir con el principio de trasparencia incluso en el caso de decisiones que se basan en datos técnico-científicos, como destaca la jurisprudencia sobre los algoritmos (véase en particular la del Conseil Consitutional francés, del Consiglio di Stato italiano, pero también la sentencia en el caso Syri).

Del mismo modo, la Comunicación del 2019 sobre «Generar confianza en la inteligencia artificial centrada en el ser humano» (véase párr. 2.2.iv) prevé que «es importante registrar y documentar tanto las decisiones tomadas por los sistemas como la totalidad del proceso (incluida una descripción de la recogida y el etiquetado de datos, y una descripción del algoritmo utilizado) que dio lugar a las decisiones». Hay que poner en marcha medidas para la explicabilidad del proceso de toma de decisiones algorítmico.

Por otra parte, en la Resolución del Parlamento sobre la robótica se afirma que la «transparencia y la inteligibilidad de los procesos decisorios» se deben considerar como los principales principios éticos aplicables. De hecho «siempre ha de ser posible justificar cualquier decisión que se haya adoptado con ayuda de la inteligencia artificial y que pueda tener un impacto significativo sobre la vida de una o varias personas (...); siempre debe ser posible reducir los cálculos del sistema de inteligencia artificial a una forma comprensible para los humanos; (...) los robots avanzados deberían estar equipados con una 'caja negra' que registre los datos de todas las operaciones efectuadas por la máquina, incluidos, en su caso, los pasos lógicos que han conducido a la formulación de sus decisiones» (párr. 12).

c) Responsabilidad

En coherencia con el principio de no exclusividad, las decisiones que se basan en datos técnico-científicos deben ser atribuibles a sujetos jurídicos y especialmente a personas.

Esos sujetos tendrán la responsabilidad (civil, administrativa y penal) en caso de incumplimiento de las reglas y principios (véase ampliamente el Libro Blanco sobre la Inteligencia artificial de 2020, especialmente p. 15; véase también el Libro Blanco «Inteligencia artificial para Europa», espec. el párr. 3.3 «Garantizar un marco ético y jurídico adecuado»; véase la Resolución del Parlamento europeo del 2017, puntos L y M).

En la Comunicación «Generar confianza en la inteligencia artificial centrada en el ser humano» (véase párr. 2.2.vii) donde se afirma a este respecto: «Deben instaurarse mecanismos que garanticen la responsabilidad y la rendición de cuentas de los sistemas de IA y de sus resultados, tanto antes como después de su implementación. La posibilidad de auditar los sistemas de IA es fundamental, puesto que la evaluación de los sistemas de IA por parte de auditores internos y externos, y la disponibilidad de los informes de evaluación, contribuye en gran medida a la fiabilidad de la tecnología. La posibilidad de realizar auditorías externas debe garantizarse especialmente en aplicaciones que afecten a los derechos fundamentales, por ejemplo, las aplicaciones críticas para la seguridad».

Muy atenta al tema de la responsabilidad por las consecuencias del uso de sistemas de inteligencia artificial es la Resolución de 2017 del Parlamento concerniente a la robótica (véase párr. 49-59) que exige que este sistema de responsabilidad no se base en la responsabilidad objetiva (que sin embargo requiere «probar que se ha producido un daño o perjuicio y el establecimiento de un nexo causal entre el funcionamiento perjudicial del robot y los daños o perjuicios causados a la persona que los haya sufrido», véase párr. 54) sino en la gestión del riesgo, es decir, que las personas deben ser capaces de minimizar los riesgos y gestionar el impacto negativo (párr. 55); además se afirma la necesidad de un esquema de seguros obligatorios (párr. 57) y de un fondo en caso de ausencia del seguro (párr. 58), como ya está previsto para los vehículos.

VI ÉTICA Y RESPETO DE LOS DERECHOS E INTERESES FUNDAMENTALES

Ahora bien, la ética de la IA debe consistir en el respeto de derechos humanos y otros intereses considerados fundamentales por el ordenamiento jurídico de la UE (véase el «considerando» 28 de la propuesta de Reglamento sobre la IA; véase también la Resolución del Parlamento sobre la robótica, párr. 10)).

Incluso el derecho a la dignidad humana (véase también jurisprudencia Loomis (en el caso COMPAS) y HART), los sistemas de IA no pueden violar o poner en peligro la vida privada y familiar de las personas, especialmente las más vulnerables; la libertad de expresión y de información; la libertad de reunión y de asociación; los derechos de consumidores y los trabajadores; el derecho a la salud; la seguridad; el medioambiente.

Por lo que concierne el tema del uso de la inteligencia artificial en los procesos de toma de decisiones administrativas y judiciales, los sistemas de IA deben respetar el derecho a la tutela judicial efectiva y a un juez imparcial, los derechos de la defensa y la presunción de inocencia, y el derecho a una buena administración (véase también el considerando 28 antemencionado). Por otra parte, la Comunicación «Generar confianza en la inteligencia artificial centrada en el ser humano» (párr. 2.2.v) recuerda que la toma de decisiones asistida por la IA puede provocar discriminaciones, por lo que establece: «Los conjuntos de datos utilizados por los sistemas de IA (tanto para el entrenamiento como para el funcionamiento) pueden verse afectados por la inclusión de sesgos históricos involuntarios, por no estar completos o por modelos de gobernanza deficientes. La persistencia en estos sesgos podría dar lugar a una discriminación (in)directa. También pueden producirse daños por la explotación intencionada de sesgos (del consumidor) o por una competencia desleal. Por otra parte, la forma en la que se desarrollan los sistemas de IA (por ejemplo, la forma en que está escrito el código de programación de un algoritmo) también puede estar sesgada. Estos problemas deben abordarse desde el inicio del desarrollo del sistema».

La protección efectiva de los derechos humanos está asociado al principio de transparencia antemencionado, especialmente en presencia de situaciones de vulnerabilidad, como en caso de la gestión de «la migración, el asilo y el control fronterizo afectan a personas que con frecuencia se encuentran en una situación especialmente vulnerable y dependen del resultado de las actuaciones de las autoridades competentes» (véase. el «considerando» 39 de la propuesta de Reglamento sobre la IA).

La protección de los derechos fundamentales es tan importante en el sistema jurídico-ético de la UE, y especialmente en ámbito de las tecnologías, que no se considera admisible una derogación ni siquiera cuando hay que tener en cuenta otros intereses fundamentales, como el «orden público»[27] y la «seguridad pública»[28].

---

[27] el concepto de «orden público» requiere «aparte de la perturbación del orden social que constituye cualquier infracción de la ley, que exista una amenaza real, actual y suficientemente grave que afecte a un interés fundamental de la sociedad» Véanse sentencias Zh. y O., C-554/13, EU:C:2015:377.

[28] Para Tribunal de Justicia ese concepto comprende «tanto la seguridad interior de un Estado miembro como su seguridad exterior, y que, en consecuencia, el hecho de poner en peligro el funcionamiento de las instituciones y de los servicios públicos esenciales, así como la

Según la Opinión del Grupo Europeo sobre la ética en la ciencia y en la tecnología (European Group on Ethics in Science and New Technologies, en adelante «EGE») sobre las «Ethics of Security and Surveillance Technologies» (Opinión n°28 del 20 de mayo de 2014), no es aceptable un sacrificio de los derechos fundamentales para garantizar la seguridad (el llamado «trade-off», véase el párrafo 4).

Como ejemplo se utiliza el uso de algoritmos que identifican los rostros (y otros sistemas biométricos) elaborados en base a perfiles que se refieren a específicos grupos de personas. En este caso el sistema de IA realiza una «estigmatización por diseño» y por tanto una violación de la dignidad humana.

Cuando el uso de la IA puede afectar un derecho fundamental, los criterios inspiradores son los previstos en el artículo 52 de la Carta UE cuando establece que las limitaciones a un derecho fundamental deben ser necesarias y proporcionadas con el objetivo de garantizar otros derechos o intereses fundamentales. De manera que se adopten soluciones estrictamente necesarias y proporcionales a la amenaza, minimizando los efectos negativos sobre las personas y la sociedad en su conjunto

VII LEGISLACIÓN SOBRE LA IA Y COLABORACIÓN TRANSDISCIPLINARIA

Como opina la jurisprudencia del Tribunal Europeo de los Derechos Humanos, una legislación que tiene que ocuparse de temas asociados a la ciencia y a la tecnología debe surgir de un debate no solo científico, sino también más amplio (véase la jurisprudencia Evans c. Reino Unido de 2007)[29] y debe tener en cuenta la rápida evolución en estos ámbitos (véase por ejemplo S. H. y otros c. Austria del 2011[30], en particular el apartado n°117)[31].

Lo que es aún más evidente en un sector, el de la IA, cuya evolución es notable. Es por ello por lo que la elaboración de un marco jurídico debe basarse en una colaboración entre las instituciones con los sujetos que participan en la construcción de sistemas de IA y los que los utilizan.

supervivencia de la población, además del riesgo de una perturbación grave de las relaciones exteriores o de la coexistencia pacífica de los pueblos, o, incluso, la amenaza a intereses militares, pueden afectar a la seguridad pública». Véase, en ese sentido, la sentencia Tsakouridis, C-145/09, EU:C:2010:708, apartados 43 y 44. véase también el «considerando» n°19 del Reglamento (UE) 2018/1807 del Parlamento europeo y del Consejo del 14 de noviembre de 2018 relativo a un marco para la libre circulación de datos no personales en la Unión Europea.

29 TEDH, sent. 10 de abril de 2007, *Evans c. Reino Unido* [GC], no. 6339/05.

30 TEDH, sent. 3 de noviembre de 2011, *S.H. c. Austria* [GC], no 57813/00.

31 Véase también TEDH, sent. 11 de julio de 2002, *Christine Goodwin c. Reino Unido*, no. 28957/95, § 74, ECHR 2002-VI; Id. sent. 28 de mayo de 2002, Stafford c. Reino Unido [GC], no. 46295/99, § 68, ECHR 2002-IV.

En una lógica colaborativa entre los actores, la vía europea para la construcción de una IA fiable se basa en un enfoque multidisciplinario de los problemas[32], donde, por un lado, los conocimientos tecnológicos por sí solos ya no son suficientes para solucionar los problemas que surgen del uso de la IA y, por otro lado, las normas jurídicas impuestas de manera unilateral constituirían un obstáculo al desarrollo tecnológico y a los beneficios que pueden derivar de ello.

El primer paso para la creación de un sistema de confianza dedicado a la IA consiste en la adopción de un enfoque que identifique una metodología para abordar el problema ético en el ámbito científico y tecnológico. La vía trazada es un ejemplo de sinergia entre desarrolladores de aplicaciones tecnológicas, juristas y expertos en ética, realizado con la intención de abordar el problema ético y jurídico en el desarrollo de tecnologías que explotan la IA desde el momento de su programación.

La Comisión europea está convencida de que, adoptando este enfoque metodológico, el marco normativo de la Unión puede constituir la referencia mundial para la IA centrada en el ser humano[33]. Es lo que ya está pasando en otros ámbitos éticos de la tecnociencia, como la protección de los datos personales[34] (y también en los casos de la conservación del medioambiente y del bienestar de los animales; de la seguridad; de la evaluación de los aspectos éticos de la ciencia financiada, etc.)[35].

Para garantizar que los sistemas de IA sean éticos, en el sentido que no afectan los derechos e intereses fundamentales, incluso el necesario control del funcionamiento y el control judicial sería importante construir estos sistemas de manera adecuada.

Se trata de un tema destacado en las fuentes europeas sobre la IA, donde se afirma que los problemas éticos (por ejemplo, la potencial discriminación) deben abordarse desde el inicio del desarrollo del sistema (véase la Comunicación «Generar confianza en la inteligencia artificial centrada en el ser humano» (párr. 2.2.v).

Las reglas en un sistema de confianza se originarían en fuentes jurídicas basadas en el debate entre instituciones y sociedad, como se ha visto, pero también en instrumentos jurídicos flexibles como códigos de conducta, elaborados por las organizaciones y partes interesadas; como la normalización

32 LEIKAS J., KOIVISTO R., *et al.*, "Ethical Framework for Designing Autonomous Intelligent Systems", *Journal of Open Innovation: Technology, Market, and Complexity*, Vol. 5, n°1, 2019, p. 18. doi:10.3390/joitmc5010018.

33 Véase la comunicación "*Generar confianza en la inteligencia artificial centrada en el ser humano*", *cit.*, párr. 2

34 CIPPITANI, R., *La protección de datos personales y el Derecho de la integración*, en Pizzolo, C. (Coord.), Integración regional y Derechos humanos. Puntos de convergencia, Astrea, Buenos Aires, 2021, pp. 175-209.

35 VÉASE CIPPITANI, R., *Diálogo entre cortes y elaboración de principios éticos de la ciencia*, en FIGUEIREDO, M.; ARCARO CONCI, L.G.; GERBER, K. A., *A jurisprudência e o diálogo entre tribunais*, Lumen Juris, Rio de Janeiro, 2016, pp. 123-152.

técnica relativas al diseño, la fabricación o las prácticas empresariales[36]; o a través de la certificación[37].

Para que un sistema de IA sea comprensible, predecible y controlable *ex post*, debe ser capaz de garantizar el principio de transparencia y el Estado de derecho *ex ante*, es decir, en el momento del diseño algorítmico.

Los requisitos de una IA fiable deben «traducirse» en procedimientos que deben integrarse en la arquitectura de los sistemas de IA.[38].

La construcción de una inteligencia artificial lícita y ética conlleva la necesidad de incorporar elementos éticos/jurídicos desde el momento en que se diseña el algoritmo, vigilando su desarrollo para evitar distorsiones de distinta naturaleza en el momento operativo y final.

El aspecto crítico se centra en que, como en otros asuntos de ética de la tecnociencia, los problemas que hay que enfrentar y solucionar tienen una dimensión no solo nacional o continental, sino mundial.

Como demuestra la jurisprudencia del Tribunal de Justicia en materia de protección de datos personales (vid. por ejemplo los asuntos Schrems[39] y Schrems II[40] en la relación entre la legislación de la UE y de los EEUU) las enfoques sobre los aspectos éticos de la tecnociencia en muchos casos siguen siendo distantes[41].

---

36 La normalización técnica puede funcionar como un sistema de gestión de la calidad para los usuarios de la IA, los consumidores, las organizaciones, las instituciones de investigación y los gobiernos, ofreciendo a todos estos agentes la capacidad de reconocer y fomentar una conducta ética a través de sus decisiones de compra. Entre los que se pueden citar las normas ISO o la serie de normas IEEE P7000, aunque en el futuro podría resultar adecuado crear un sello de "IA fiable" que, a partir de las normas técnicas especificadas, confirme por ejemplo que el sistema cumple los requisitos de seguridad, solidez técnica y explicabilidad.

37 Como afirma la Comisión europea «La etiqueta voluntaria permitirá a los agentes económicos interesados en mostrar que los productos y servicios provistos de IA que ofrecen son fiables. Además, permitirá a los usuarios distinguir fácilmente si los productos y servicios en cuestión respetan ciertos referentes objetivos y normalizados a escala de la UE, que van más allá de las obligaciones legales aplicables normalmente. Ello contribuirá a incrementar la confianza de los usuarios en los sistemas de IA y fomentará una adopción generalizada de esta tecnología. Esta opción conlleva la creación de un nuevo instrumento jurídico para establecer un marco de etiquetado voluntario para los desarrolladores y/ o implementadores de los sistemas de IA que no se consideren de alto riesgo. Si bien la participación en el sistema de etiquetado debe ser voluntaria, una vez que el desarrollador o implementador opte por usar la etiqueta, todos los requisitos serán vinculantes. La combinación de estas imposiciones ex ante y ex post debe garantizar que se cumplan todos los requisitos». Véase el Libro Blanco sobre la inteligencia artificial cit., p. 29, en https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_es.pdf.

38 *Ivi*, p. 26.

39 Tribunal de Justicia, sent. 6 de octubre 2015, C-362/14, Schrems, ECLI:EU:C:2015:650.

40 Tribunal de Justicia, sentencia del 16 de julio de 2020, Facebook Ireland et Schrems (C-311/18), ECLI:EU:C:2020:559

41 CIPPITANI, R., *La protección de datos personales y el Derecho de la integración*, en Pizzolo, C. (Coord.), *Integración regional y Derechos humanos. Puntos de convergencia*, Astrea, Buenos Aires, 2021, pp. 175-209

VIII BIBLIOGRAFÍA

OSWALD, M.; GRACE, J.; URWIN, S. Y BARNES, G.C., *Algorithmic Risk Assessment Policing Models: Lessons from the Durham HART Model and "Experimental" Proportionality, en Information & Communications Technology Law*, n. 27/2018, p. 223.

VÉASE COLCELLI V., BURZAGLI L., *Public administration and technology in the time of artificial intelligence: the Italian case of Council State judgment n. 2019/2270*, en Arnold, R. y Danėlienė, I (Eds), *The Concept of Democracy as developed by Costitutional Justice, Constitutional Court of the Republic of Lithuania*, Vilnus, 2020, p. 137-148.

ANDORNO, R., *The Oviedo Convention: A European Legal Framework at the Intersection of Human Rights and Health Law*, in JIBL, Vol 02, I, 2005, pp. 133-143.

LEIKAS J., KOIVISTO R., *et al.*, "Ethical Framework for Designing Autonomous Intelligent Systems", *Journal of Open Innovation: Technology. Market, and Complexity*, Vol. 5, n°1, 2019, p. 18, doi:10.3390/joitmc5010018.

CIPPITANI, R., *La protección de datos personales y el Derecho de la integración*, en Pizzolo, C. (Coord.), Integración regional y Derechos humanos. Puntos de convergencia, Astrea, Buenos Aires, 2021, pp. 175-209.

VÉASE CIPPITANI, R., *Diálogo entre cortes y elaboración de principios éticos de la ciencia*, en FIGUEIREDO, M.; ARCARO CONCI, L.G.; GERBER, K. A., *A jurisprudência e o diálogo entre tribunais, Lumen Juris*, Rio de Janeiro, 2016, pp. 123-152.

CIPPITANI, R., *La protección de datos personales y el Derecho de la integración*, en Pizzolo, C. (Coord.), *Integración regional y Derechos humanos. Puntos de convergencia*, Astrea, Buenos Aires, 2021, pp. 175-209