

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/339687454>

Derechos Humanos e Inteligencia Artificial

Chapter · March 2020

CITATIONS
0

READS
2,291

1 author:



Ana Gascon Marcen
University of Zaragoza

37 PUBLICATIONS 8 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



IT Law [View project](#)



Human Rights Protection in Europe [View project](#)

Setenta años de Constitución Italiana y cuarenta años de Constitución Española

Volumen V

Retos en el siglo XXI

DIRECTORES:

ANTONIO PÉREZ MIRAS

GERMÁN M. TERUEL LOZANO

EDOARDO C. RAFFIOTTA

MARIA PIA IADICICCO

COORDINADORA:

SILVIA ROMBOLI



eucons



CENTRO DE ESTUDIOS POLÍTICOS Y CONSTITUCIONALES

	Páginas
5.1.11 Prospettive per una Unione «sociale» europea: i vincoli di bilancio e la garanzia dei diritti sociali in Italia. <i>Giuseppe Eduardo Polizzi</i>	191
5.1.12 Il principio di equilibrio finanziario e i vincoli di bilancio: l'impatto delle politiche fiscali europee sui diritti sociali. <i>Antonio Tipaldi</i>	205
5.1.13 La extrafiscalidad en la jurisprudencia del Tribunal Constitucional español. <i>José Miguel Martín Rodríguez</i>	221
5.1.14 Constitucionalismo indígena latinoamericano y derechos sociales. <i>M.ª Concepción Pérez Villalobos</i>	233
5.1.15 Precisiones necesarias para comprender el proceso constituyente iniciado en Cuba con el Proyecto de Constitución de 22 de julio de 2018. <i>Antonio-Filiu Franco Pérez</i>	253
PARTE 2. LA ERA DE LA TECNOLOGÍA Y EL CAMBIO CLIMÁTICO	
5.2.1 Il costituzionalismo nella società tecnologica. <i>Tommaso Edoardo Frosini</i>	273
5.2.2 La democrazia liberale alla prova della rivoluzione tecnologica: una sfida tra dati e date. <i>Marco Schirripa</i>	291
5.2.3 L'amministrazione e la sfida dei <i>Big data</i> . <i>Giovanna De Minico</i> ...	301
5.2.4. El derecho constitucional en la era de la inteligencia artificial, los robots y los drones. <i>Joaquín Sarrión Esteve</i>	321
5.2.5 Derechos humanos e inteligencia artificial. <i>Ana Gascón Marcén</i> ..	335
5.2.6 Dificultades específicas del e-government en el seno de Estados descentralizados. Los casos español e italiano. <i>Alfonso Sánchez García</i>	351
5.2.7 Retos del Derecho constitucional a la luz del cambio climático. <i>Borja Sánchez Barroso</i>	367
PARTE 3. LA SEGURIDAD Y LA LEGALIDAD SANCIONADORA	
5.3.1 El concepto constitucional de seguridad en permanente redefinición. <i>Enrique Guillén López</i>	387
5.3.2 Algunas consideraciones acerca de las referencias a la dignidad en el Derecho penal. <i>Clara Moya Guillem</i>	405
5.3.3 ¿El Derecho Penal como instrumento de tutela de los derechos fundamentales, sociales y laborales? Un cuestionamiento desde el análisis del delito de discriminación laboral. <i>Ignacio Valentín Mayoral Narros, David Pavón Herradón</i>	421
5.3.4 Derecho Penal y Derecho Administrativo sancionador a la luz de la Constitución española: franja de convergencia y divergencia. <i>Ignacio Valentín Mayoral Narros, María José Molina García</i>	437

DERECHOS HUMANOS E INTELIGENCIA ARTIFICIAL

ANA GASCÓN MARCÉN*
Universidad de Zaragoza

SUMARIO: 1. Introducción. 2. Derechos humanos afectados por el uso de la inteligencia artificial. 3. La respuesta de las organizaciones internacionales. 3.1 Las Naciones Unidas. 3.2 Las organizaciones europeas. 3.2.1 El Consejo de Europa. 3.2.2 La Unión Europea. 4. Conclusiones. 5. Referencias bibliográficas.

1. INTRODUCCIÓN

En un mundo profundamente digitalizado, la inteligencia artificial está cada vez más presente en nuestras vidas, incluso en muchas ocasiones sin que nos demos cuenta. La inteligencia artificial tiene un potencial enorme para mejorar el bienestar de las personas a través del diagnóstico médico o de sistemas logísticos, de transporte o energía más eficientes en pos de un desarrollo más sostenible (UN Global Pulse, 2012). Actualmente existe una carrera de diferentes Estados por consolidarse como los líderes en la materia, ya que no quieren quedarse atrás en un ámbito de innovación que puede cambiar muchos aspectos de nuestras vidas y

* Además de Profesora Contratada Doctora Interina en la Universidad de Zaragoza, la autora es funcionaria en excedencia del Consejo de Europa, no obstante, las opiniones expresadas en este trabajo son responsabilidad únicamente de la autora y en ningún caso representan la opinión de la organización.

producir beneficios económicos considerables (Dutton, 2018). Sin embargo, unidos a esas ventajas se presentan también riesgos que suponen un desafío para la efectiva protección y el disfrute de los derechos humanos.

La inteligencia artificial puede tener un impacto positivo en el ejercicio de determinados derechos, pero otros pueden verse menoscabados por su uso como: la libertad de expresión (incluyendo el derecho a comunicar o recibir información libremente), la privacidad, la prohibición de la discriminación, y el derecho a la tutela judicial efectiva, entre otros, como se analizará en el apartado 2. Las organizaciones internacionales dedicadas a la protección de los derechos humanos están actualmente reflexionando sobre estos riesgos y las posibles soluciones para ayudar a sus Estados miembros a responder a estos retos, así en el apartado 3 se presentan las acciones llevadas a cabo en el marco universal por las Naciones Unidas (NNUU) y en el ámbito europeo por el Consejo de Europa y la Unión Europea (UE).

2. DERECHOS HUMANOS AFECTADOS POR EL USO DE LA INTELIGENCIA ARTIFICIAL

En este apartado, se expondrá de una manera panorámica cómo el uso de la inteligencia artificial puede afectar a determinados derechos, comenzando por la libertad de expresión, que incluye el derecho a comunicar o recibir información libremente. A la hora de ejercer este derecho en línea es común apoyarse en ciertas plataformas que hacen uso de la inteligencia artificial para hacer más eficientes sus servicios.

En el caso de buscadores como Google, su algoritmo hace uso de la inteligencia artificial para optimizar la indexación de la información, lo mismo ocurre con plataformas de vídeo como Youtube o redes sociales como Facebook. Esto supone que utilizando la información que tienen de sus usuarios estas empresas personalizan los resultados que éstos obtienen. Como se trata de empresas que buscan maximizar sus resultados, los algoritmos no van a tener en cuenta criterios como la pluralidad de la prensa que es básica en una sociedad democrática, sino que van a buscar exponer a sus usuarios a ideas que les gusten, que compartan y que en cierta manera refuercen lo que ya piensan sobre un tema creando así burbujas informativas. Incluso, en ciertos casos, algunas páginas tienen un algoritmo para mostrar contenido cada vez más extremo para que los usuarios sigan viendo vídeos o entrando en enlaces (Tufekci, 2018).

Esto hace que sufra la libertad de comunicar o recibir información libremente, dado que la información que se recibe está preseleccionada por la inteligencia artificial, sin que en ocasiones ni siquiera se sea consciente de ello. Esto

está muy relacionado con el fenómeno de las noticias falsas o «fake news» que pueden ser extendidas, e incluso creadas por medio de la inteligencia artificial y conlleva consecuencias muy serias en la participación política (Nemitz, 2018).

Además, muchas de estas empresas fundamentan su base de negocio en la economía de los datos, lo que también hace que la inteligencia artificial pueda interponerse en el disfrute del derecho a la privacidad. Se ha demostrado que a partir de interacciones en principio inocuas en redes sociales se pueden deducir características de una persona tan sensibles como sus preferencias sexuales o políticas. Aplicando inteligencia artificial a ingentes cantidades de datos se pueden encontrar patrones y tendencias muy difíciles de deducir para un ser humano, pero también se corre el riesgo de crear perfiles equivocados de una persona, de desanonimizar datos o de ir mucho más allá de lo que la persona haya consentido al ceder los datos. Los perfiles que se pueden realizar uniendo datos de una persona y deduciendo otros nuevos pueden ser decisivos en que una persona consiga un empleo, pueda suscribir un seguro por enfermedad o acceder a una hipoteca.

Los programas que utilizan inteligencia artificial se presumen como neutros y se entiende que, al decidir sobre un determinado aspecto, una máquina lo hará de manera más objetiva. No obstante, esos programas se entrenan con datos que en sí mismos pueden incluir prejuicios reforzando así la discriminación, pero haciéndola más difícil de identificar. Por ejemplo, en Estados Unidos un informe de ProPublica reveló que un programa que se usaba en los tribunales de Estados Unidos para calcular la probabilidad de reincidencia de los acusados de la comisión de un delito utilizaba una fórmula que era particularmente propensa a señalar falsamente a los acusados negros como futuros delincuentes, etiquetándolos erróneamente de esta manera casi el doble de veces que a los acusados blancos y, además, los acusados blancos fueron etiquetados erróneamente como de bajo riesgo con más frecuencia que los acusados negros (Angwin et al., 2016). Ni siquiera es necesario que el algoritmo tome un criterio como la raza como factor, dado que esto está claramente prohibido, pero se pueden usar otros indicios relacionados (*proxies*) como la dirección para que el resultado sea indirectamente discriminatorio. Otro ejemplo es el de un programa desarrollado por Amazon para buscar talento y filtrar los currículos para la contratación de nuevos empleados que fue descartado por descubrirse que discriminaba a las mujeres (Dastin, 2018). Este ejemplo es relevante porque este tipo de programas se utilizan cada vez más en la selección de personal y pueden tener efectos que limiten el acceso de determinados grupos al mercado de trabajo reforzando situaciones históricamente discriminatorias en vez de corregirlas.

A esto se une la dificultad en muchos casos para recurrir este tipo de decisiones porque hace falta una investigación en profundidad para demostrar que se está siendo víctima de discriminación porque estos programas están protegidos por las leyes de propiedad intelectual, y no se tiene por tanto acceso a su código. Incluso hay algunos que optimizan su funcionamiento de manera continuada y los parámetros que utiliza y la manera de llegar a los resultados va cambiando, es lo que se conoce como *machine learning*.

La tutela judicial efectiva también se puede ver mermada porque no se sea consciente de que un contenido publicado, por ejemplo, ha sido filtrado o bloqueado por un programa que usa inteligencia artificial. Desde los diferentes gobiernos cada vez se fomenta más que los intermediarios recurran a métodos automáticos para eliminar contenidos ilegales, ya sean pornografía infantil, contenidos que promueven el terrorismo o que vulneran las normas de protección de la propiedad intelectual. Se trata normalmente de mecanismos bastante eficientes, pero que presentan problemas como que el contenido puede estar protegido por la libertad de expresión que se extiende a los discursos ofensivos, chocantes o perturbadores. Esto tiene que ver con una tendencia peligrosa a pasar el control de la aplicación de la ley de las autoridades estatales a las empresas (Angelopoulos et al., 2016). Además, un algoritmo puede no entender cuando un contenido protegido por derechos de autor se está utilizando conforme a una excepción legítima o como en una expresión que literalmente podría parecer una arenga a la violencia se está haciendo uso de la ironía, el humor o el sarcasmo. Esto puede dar lugar al sobrebloqueo o sobrefiltrado de información. En muchos casos es complejo recurrir estas acciones porque: no se es consciente de ser víctima de las mismas, los mecanismos para apelar no existen, o incluso ellos mismos están gestionados a través de inteligencia artificial.

3. LA RESPUESTA DE LAS ORGANIZACIONES INTERNACIONALES

Los Estados están compitiendo para desarrollar estrategias para fomentar la investigación en inteligencia artificial, pero normalmente el énfasis no se pone en una aproximación desde los derechos humanos sino desde el progreso científico y económico. No obstante, la doctrina especializada lleva muchos años denunciando la necesidad de crear marcos adecuados ya sea éticos o jurídicos que rijan estos desarrollos. Incluso las grandes multinacionales tecnológicas están empezando a publicar sus cartas de principios relativos a la inteligencia artificial.

Las organizaciones internacionales de protección de los derechos humanos son conscientes de los riesgos enunciados en el apartado anterior y están tomando medidas para investigar el impacto real de la inteligencia artificial y elaborar las recomendaciones necesarias que faciliten a sus Estados miembros reaccionar ante las mismas. La labor de estas organizaciones puede presentar un valor añadido por su amplia experiencia en la protección de los derechos humanos y como foros de debate que fomentan la colaboración de sus Estados miembros. En el plano internacional se describirá la acción de las NNUU y en el plano europeo del Consejo de Europa y de la UE.

3.1 Las Naciones Unidas

En el marco de las NNUU, son múltiples las organizaciones y órganos que exploran las oportunidades y retos de la inteligencia artificial. En octubre de 2015, la 70.^a reunión de la Asamblea General celebró un evento sobre cómo hacer frente a los desafíos de la seguridad internacional por la aparición de la inteligencia artificial.

La Unión Internacional de Telecomunicaciones, organismo especializado de las NNUU para las tecnologías de la información y la comunicación, se ha convertido en una de las plataformas clave para explorar este tema. Organizó la *AI for Good Global Summit* en 2017, 2018 y 2019, que se ha centrado en las estrategias para garantizar un desarrollo fiable, seguro e inclusivo de estas tecnologías y el acceso equitativo a sus beneficios. El Foro de la Gobernanza de Internet de NNUU también ha celebrado en 2017 y 2018 múltiples sesiones dedicadas a la inteligencia artificial, un tema que cada año recibe más atención en este marco. Si bien hay que entender la propia naturaleza de este foro que sirve para fomentar el diálogo entre las distintas partes interesadas pero que no se enfoca a la consecución de resultados concretos. El Instituto Interregional de Investigación de la Criminalidad y la Justicia de las NNUU lanzó su programa sobre inteligencia artificial y robótica en 2015 y en septiembre de 2017 creó con el apoyo del Gobierno de los Países Bajos y la Ciudad de La Haya el Centro de Inteligencia Artificial y Robótica. El objetivo de este centro es concentrar la experiencia en inteligencia artificial de las NNUU, pero desde el enfoque de la lucha contra el crimen y la seguridad.

En octubre de 2017, se produjo una reunión conjunta del Consejo Económico y Social y la Segunda Comisión de la Asamblea General para considerar

el papel y el impacto de la inteligencia artificial en el desarrollo sostenible.¹ En la Resolución 73/179 de 17 de diciembre de 2018, la Asamblea observó con preocupación que la elaboración de perfiles, la adopción automatizada de decisiones y las tecnologías de aprendizaje automático, a veces denominadas inteligencia artificial, pueden, sin las debidas salvaguardias, dar lugar a decisiones que afecten al disfrute de los derechos humanos, incluidos los derechos económicos, sociales y culturales, y reconoció la necesidad de aplicar el derecho internacional de los derechos humanos al diseño, la evaluación y la reglamentación de esas prácticas, alentando al Consejo de Derechos Humanos y a la Oficina del Alto Comisionado de las NNUU a seguir este tema. El trabajo quizás más centrado en analizar el impacto de la inteligencia artificial en los derechos humanos lo constituye el Informe del Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión sobre la materia (Kaye, 2018). En él, se analizan las características especiales que hacen que la inteligencia artificial presente problemas específicos como son: la automatización de la toma de decisiones, la gran capacidad de análisis de datos que tiene y su adaptabilidad. El Relator apuesta por un enfoque de la inteligencia artificial basado en los derechos humanos no sólo por parte de los Estados sino también de las empresas. Para concretizar los Principios Rectores sobre las Empresas y los Derechos Humanos de NNUU en este campo es necesario que las empresas asuman compromisos orientados a respetar los derechos humanos de sus usuarios en todas las aplicaciones de la inteligencia artificial; se abstengan de causar o contribuir a cualquier efecto negativo para los derechos humanos de la inteligencia artificial y prevenir o mitigar todo efecto negativo relacionado con sus operaciones; hagan operaciones de diligencia debida sobre los sistemas de inteligencia artificial para determinar y resolver las posibles repercusiones en los derechos humanos; apliquen estrategias de prevención y mitigación; hagan un seguimiento continuo de las actividades relacionadas con la inteligencia artificial, en particular mediante consultas con los interesados y consultas públicas; y ofrezcan mecanismos de reclamación accesibles para remediar los efectos negativos en los derechos humanos de los sistemas de inteligencia artificial.

¹ El Secretario General de las NNUU aseveró en su Estrategia en materia de nuevas tecnologías que la inteligencia artificial encierra una promesa extraordinaria para el bienestar humano, pero también puede generar más desigualdad y más violencia. Ofrece nuevas y poderosas maneras de cumplir con los Objetivos de Desarrollo Sostenible, pero también puede exacerbar la desigualdad en y entre las naciones, y contribuir al desempleo, asimilar y profundizar los prejuicios. Por ello, se comprometió en nombre de la Organización a: profundizar las capacidades internas de las NNUU y su exposición a las nuevas tecnologías; aumentar la comprensión, la promoción y el diálogo; apoyar el diálogo sobre marcos normativos y de cooperación; y aumentar el apoyo de NNUU al desarrollo de la capacidad de los gobiernos.

El Relator considera que proponer una ley o regulación integral de la inteligencia artificial puede ser inadecuado para un campo tan innovador, dado que tendría que ser muy general y podría terminar siendo demasiado vaga o demasiado restrictiva. En ese sentido, la regulación sectorial puede ser preferible, si bien, los Estados deben garantizar que la inteligencia artificial se desarrolle de conformidad con las normas de derechos humanos. Cualquier esfuerzo por desarrollar una política o regulación estatal en el campo de la inteligencia artificial debe garantizar este tipo de consideraciones. El Relator propone medidas concretas para cumplir con los principios expuestos *supra* como son: las evaluaciones de impacto de la inteligencia artificial sobre los derechos humanos (Reisman et al., 2018); las auditorías externas que presentan problemas al enfrentarse a algoritmos protegidos pero que podrían elaborarse incluso sin acceso directo al código (por ejemplo, a través de *zero-knowledge proofs*) o permitiéndoselo sólo a algunos terceros certificados; y los mecanismos de refuerzo de la autonomía individual. El Relator también plantea dar publicidad al uso de la inteligencia artificial a través de obligaciones de información y notificación, igual que se requieren avisos públicos de los servicios de videovigilancia, ya que a través del uso del lenguaje natural una persona puede llegar a ignorar completamente que está interactuando con una máquina (véase Google Duplex). Además, el Relator considera esencial asegurar medidas correctivas cuando se comprueben violaciones de los derechos humanos. Su informe concluye con recomendaciones tanto para los Estados como para las empresas que diseñan y utilizan inteligencia artificial.

3.2 Las organizaciones europeas

3.2.1 EL CONSEJO DE EUROPA

El Consejo de Europa es un líder en lo referente a la interacción entre tecnologías de la información y la comunicación y derechos humanos como lo demuestra su trabajo en la gobernanza de Internet (Gascón Marcén, 2014). En el marco de la organización, se ha creado un coordinador de inteligencia artificial que se encarga de supervisar la coherencia de la actividad de la misma en este ámbito dado que existen iniciativas en materia de: lucha contra la discriminación, bioética, derechos de los niños, cultura, protección de datos personales, democracia, educación, lucha contra la corrupción y el delito, igualdad de género, sociedad de la información y justicia.

La inteligencia artificial ocupó un lugar importante en la deliberación de su Comité de Expertos sobre los Intermediarios de Internet que publicó un estudio sobre la dimensión de los derechos humanos de las técnicas de procesamiento automatizado de datos y sus posibles implicaciones regulatorias (MSI-NET, 2017). En él se examina el impacto del uso de algoritmos en derechos como la tutela judicial efectiva, la privacidad y protección de datos, la libertad de expresión, la libertad de reunión y asociación, la prohibición de la discriminación, los derechos sociales y el acceso a servicios públicos, y el derecho a elecciones libres, entre otros, subrayándose la necesidad de cara a su regulación de poner énfasis en la transparencia, la responsabilidad y los marcos éticos y la evaluación mejorada del riesgo. Los autores de este estudio consideran que el debate público sobre las múltiples dimensiones de los algoritmos en materia de derechos humanos va muy por detrás de la evolución tecnológica y debe fortalecerse rápidamente para garantizar que los derechos humanos y los intereses de las personas se salvaguarden de manera eficaz y sostenible de acuerdo con los valores establecidos en el Convenio Europeo de Derechos Humanos y el resto de tratados internacionales de protección de los derechos humanos. Las conclusiones de este trabajo llevaron al Consejo de Europa a crear un nuevo Comité de Expertos sobre las dimensiones de derechos humanos del procesamiento automatizado de datos y diferentes formas de inteligencia artificial (MSI-AUT) que prepara un proyecto de recomendación del Comité de Ministros sobre los impactos de los sistemas algorítmicos en los derechos humanos. El enfoque general de esta recomendación es neutral en cuanto a la innovación, al tiempo que llama la atención sobre los riesgos específicos para los derechos humanos derivados de los sistemas algorítmicos y proporciona orientación a los Estados miembros sobre cómo minimizar esos riesgos y maximizar sus beneficios sociales. Las directrices que son la parte sustantiva de la recomendación hacen referencia a las obligaciones de los Estados miembros, por un lado, y la responsabilidad de las entidades del sector privado. El MSI-AUT debe elaborar también un estudio sobre el concepto de responsabilidad de los sistemas de toma de decisiones de inteligencia artificial dentro de un marco de derechos humanos. Los resultados de este Comité deben estar listos a final de 2019 para ser examinados por los representantes de los Estados miembros de la organización, aunque ya dio sus primeros frutos cuando elaboró una declaración sobre las capacidades de manipulación de los procesos algorítmicos adoptada por el Comité de Ministros el 13 de febrero de 2019. Un tema tremendamente relevante para la democracia, cuya importancia se vio confirmada con el escándalo de Cambridge Analytica.

En el campo de la justicia, se recurre cada vez más al uso de programas de inteligencia artificial en los sistemas judiciales de los Estados europeos, por lo que la Comisión Europea para la Eficiencia de la Justicia (CEPEJ) decidió examinar los desafíos y oportunidades relacionados con el uso de la inteligencia artificial, incluidas las llamadas herramientas de justicia predictiva, en los sistemas judiciales. La CEPEJ adoptó en diciembre de 2018, la Carta ética europea sobre el uso de la inteligencia artificial en los sistemas judiciales y su entorno. Es la primera Carta adoptada por una organización europea sobre este tema y además va acompañada de un estudio científico, un glosario y una revisión de diferentes aplicaciones de inteligencia artificial que pueden contribuir a mejorar el trabajo de los profesionales jurídicos. Los cinco principios establecidos por la Carta son:

1. principio de respeto a los derechos fundamentales²;
2. principio de no discriminación³;
3. principio de calidad y seguridad⁴;
4. principio de transparencia, imparcialidad y equidad⁵; y
5. principio «bajo el control del usuario»⁶.

El Comité Consultivo del Convenio para la protección de las personas con respecto al tratamiento automatizado de datos de carácter personal (conocido como Convenio 108 y tras su revisión como Convenio 108+) también propuso unas guías generales sobre la inteligencia artificial, algunas de las cuáles van destinadas a los desarrolladores de programas que utilicen esta tecnología y otras a los legisladores.⁷ Más allá de las recomendaciones⁸ y los dictámenes, la inteligencia

² Debe asegurarse que el diseño e implementación de herramientas y servicios de inteligencia artificial sean compatibles con los derechos fundamentales.

³ Hay que prevenir específicamente el desarrollo o la intensificación de cualquier discriminación entre individuos o grupos de individuos.

⁴ Respecto al procesamiento de decisiones y datos judiciales, deben utilizarse fuentes certificadas y datos intangibles con modelos concebidos de manera multidisciplinaria, en un entorno tecnológico seguro.

⁵ Hay que hacer que los métodos de procesamiento de datos sean accesibles y comprensibles, y autorizar auditorías externas.

⁶ Excluye un enfoque prescriptivo y garantiza que los usuarios sean actores informados y tengan el control de sus elecciones.

⁷ Consultative Committee of the Convention for the protection of individuals with regard to automatic processing of personal data, *Draft Guidelines on Artificial Intelligence*, T-PD(2018)17, de 17 de octubre de 2018. Estas guías se basaron en el informe: Mantelero, A., *Artificial Intelligence and Data Protection: Challenges and Possible Remedies*, 2018.

⁸ La Comisaría de Derechos Humanos del Consejo de Europa también publicó un comentario sobre derechos humanos centrado en salvaguardar los derechos humanos en la era de la inteligencia artificial. En su opinión, la inteligencia artificial tiene el potencial de ayudar a los seres humanos a maximizar su tiempo, libertad y felicidad. Al mismo tiempo, puede llevarnos hacia una sociedad distópica. Por lo tanto, encontrar el equilibrio adecuado entre el desarrollo tecnológico y la protección de los derechos humanos es

artificial empieza a ser también objeto del derecho convencional del Consejo de Europa. El ejemplo más concreto es que en la modernización del Convenio 108 se ha incorporado un nuevo derecho para los sujetos de datos en su art. 9 que consiste en que todo individuo tendrá derecho a no estar sujeto a una decisión que le afecte significativamente basándose únicamente en un procesamiento de datos automatizado sin que se tenga en cuenta su opinión. El Consejo de Europa se propone elaborar en los próximos años un convenio internacional que limite los impactos negativos de la inteligencia artificial en los derechos humanos.

3.2.2 LA UNIÓN EUROPEA

La UE ha visto la cantidad de recursos que otras grandes potencias económicas mundiales están invirtiendo en innovación como China o Estados Unidos y ha decidido no quedarse atrás por lo que anunció su Estrategia de Inteligencia Artificial para Europa en 2018. La Comisión decidió incrementar sus inversiones anuales en inteligencia artificial en un 70% en el marco del programa de investigación e innovación Horizonte 2020, alcanzando así los 1.500 millones de euros para el período 2018-2020. Junto con la Estrategia de la Comisión, la mayoría de los Estados miembros firmaron una Declaración que busca crear una alianza para la inteligencia artificial en Europa cuyo objetivo es llegar de manera conjunta a invertir más de 20.000 millones de euros al año durante la próxima década en este ámbito. La UE considera que es una oportunidad para ella convertirse en un líder de la inteligencia artificial ética, uniendo a la innovación científica los valores europeos de democracia, derechos humanos y estado de derecho.

En lo relativo a los desafíos jurídicos y legales que plantea el uso extendido de la inteligencia artificial, el Grupo europeo de ética de la ciencia y de las nuevas tecnologías (grupo consultivo de la Comisión) publicó en marzo de 2018 una Declaración sobre inteligencia artificial, robótica y sistemas autónomos.⁹ Ba-

un asunto urgente, del cual depende el futuro de la sociedad en la que queremos vivir. Especialmente preocupantes en este sentido sería la capacidad de la inteligencia artificial para vulnerar el derecho a la intimidad y el derecho a la igualdad y limitar la libertad de expresión y reunión (MIJATOVIĆ, 2018).

En el marco también del Consejo de Europa, el Comité Europeo de Cooperación Jurídica (CDCJ) está trabajando actualmente en los mecanismos de resolución de disputas en línea y está considerando las posibles aplicaciones de la inteligencia artificial en tales sistemas y su cumplimiento del derecho a un proceso equitativo y a un recurso efectivo (arts. 6 y 13 del Convenio Europeo de Derechos Humanos).

⁹ En él, se establecían los principios éticos y requisitos democráticos que debían ser considerados para la creación de nuevos estándares sobre inteligencia artificial como: la dignidad humana; la autonomía; la responsabilidad; la justicia, equidad y solidaridad; la democracia; el estado de derecho; la sostenibilidad; la seguridad, integridad física y mental de las personas; y la protección de datos personales y la privacidad.

sándose en el trabajo de ese Grupo, la Comisión decidió reunir a todas las partes interesadas pertinentes en un Grupo de Expertos de Alto Nivel en Inteligencia Artificial, para contribuir al desarrollo de un proyecto de directrices éticas en este ámbito.¹⁰ Estas directrices debían tratar cuestiones tales como el futuro del trabajo, la equidad, la seguridad, la protección, la inclusión social y la transparencia de los algoritmos. En términos más generales, en ellas se examina el impacto en los derechos fundamentales, en particular, a la intimidad, la dignidad, la protección de los consumidores y la lucha contra la discriminación. El 8 de abril de 2019 se presentaron las Directrices éticas para una IA fiable.

Su objetivo es conseguir una inteligencia artificial fiable que tiene tres componentes: debe ser lícita, es decir, cumplir todas las leyes y reglamentos aplicables; ha de ser ética, de modo que se garantice el respeto de los principios y valores éticos; y debe ser robusta. El Capítulo I trata de garantizar el propósito ético de la inteligencia artificial, estableciendo cuáles son los derechos, principios y valores fundamentales que debe cumplir. En ese sentido, la inteligencia artificial debe estar centrada en el ser humano y se debe desarrollar, implementar y utilizar con un propósito ético. Tiene que regirse por los derechos fundamentales, los valores sociales y los principios éticos de respeto de la autonomía humana, prevención del daño, equidad y explicabilidad. Es necesario, además, prestar especial atención a las situaciones que involucran a grupos vulnerables o a situaciones con asimetrías de poder o información (como entre empleadores y empleados, o empresas y consumidores). El Capítulo II desarrolla una guía concreta enumerando los requisitos para la inteligencia artificial fiable y ofreciendo una visión general de los métodos técnicos y no técnicos que se pueden utilizar para su implementación. Estos requisitos que deben incluirse desde la primera fase de diseño son: acción y supervisión humanas; solidez técnica y seguridad; gestión de la privacidad y de los datos; transparencia; diversidad, no discriminación y equidad; bienestar ambiental y

La Agencia de los Derechos Fundamentales de la UE también ha empezado a investigar el impacto de la inteligencia artificial en el disfrute de los derechos y ha identificado la lucha contra la discriminación, como un área particularmente afectada por los avances tecnológicos. En su opinión, cuando se utilizan algoritmos para la toma de decisiones, existe un potencial de discriminación contra los individuos, lo que sería contrario al artículo 21 de la Carta de los Derechos Fundamentales de la UE (AGENCIA DE LOS DERECHOS FUNDAMENTALES, 2018).

¹⁰ Su presidente (que proviene del mundo empresarial) se mostró reacio a regular la inteligencia artificial de un modo prematuro para no limitar la innovación y competitividad en la materia. En su opinión, Europa debería proponer principios horizontales amplios para el uso ético de la inteligencia artificial basados en los valores fundamentales europeos, pero no crearse una regulación genérica sino más bien ir adaptando la legislación a medida que surjan los problemas y se planteen casos ante los tribunales (DELC-KER, 2018), aunque podría argumentarse que esto puede chocar con un valor europeo fundamental como es la seguridad jurídica.

social; y rendición de cuentas. Debe hacerse un esfuerzo por facilitar la capacidad de auditoría de los sistemas de inteligencia artificial, especialmente en contextos o situaciones críticas, además, de definir los métodos de explicación de los sistemas y asegurar un proceso específico para la gobernanza de la rendición de cuentas. El Capítulo III proporciona una lista para la evaluación concreta pero no exhaustiva que debe adaptarse a los usos específicos.

La Comisión se ha mostrado consciente de que, si bien la autorregulación puede proporcionar un primer conjunto de índices de referencia con respecto a los cuales sea posible valorar las aplicaciones y resultados que van apareciendo, las autoridades públicas deben garantizar que los marcos reglamentarios para el desarrollo y el uso de las tecnologías de inteligencia artificial estén en consonancia con esos valores y derechos fundamentales. La Comisión se mostró preparada para supervisar la evolución de la situación y, en caso necesario, revisar los marcos jurídicos existentes para adaptarlos a retos específicos y en particular para garantizar el respeto de los valores básicos y los derechos fundamentales de la UE (Comisión Europea, 2018, 18). Ursula von der Leyen (2019, 13), en su agenda como candidata a la presidencia de la Comisión Europea, prometió, en sus primeros 100 días en el cargo, presentar una propuesta de legislación sobre un enfoque europeo coordinado sobre las implicaciones humanas y éticas de la inteligencia artificial, lo cual parece un tanto precipitado dada la complejidad del tema.

De hecho, ya existe normativa de la UE en un caso concreto relativa a la automatización de la toma de decisiones que puede realizarse a través de la inteligencia artificial y es en el Reglamento General de Protección de Datos (RGPD)¹¹. Según algunos autores, este Reglamento planteará grandes desafíos para la industria, pero tendrá un impacto positivo al crear oportunidades para que se tome la iniciativa en Europa en el diseño de algoritmos y marcos de evaluación que eviten la discriminación y permitan explicar el funcionamiento de los algoritmos, lo que se conoce como explicabilidad (Goodman y Flaxman, 2016). Sin embargo, hay autores que se muestran mucho más escépticos, ya que, según ellos, hay varias razones para dudar tanto de la existencia legal como de la viabilidad de un «derecho a explicación» porque el Reglamento sólo exige que los interesados reciban información significativa (pero debidamente limitada) (arts. 13-15) sobre la lógica involucrada, así como la importancia y las consecuencias previstas de los sistemas de toma de decisiones automatizados, lo que, en opinión de los autores debería entenderse más

¹¹ Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE, DOUE L 119, 4.5.2016, pp. 1-88.

como un «derecho a estar informado». Además, la ambigüedad y el alcance limitado del «derecho a no estar sujeto a la toma de decisiones automatizada» contenida en el art. 22 (del cual se deriva el presunto «derecho a explicación») plantea dudas sobre la protección realmente otorgada a los interesados. Estos problemas muestran que el Reglamento carece de un lenguaje preciso, así como de derechos y garantías explícitos y bien definidos contra la toma de decisiones automatizada, y, por lo tanto, corre el riesgo de no servir de base a mecanismos coercitivos efectivos (Wachter, S. et al., 2016)¹².

4. CONCLUSIONES

La inteligencia artificial ya no es sólo un tema de la ciencia ficción que nos lleve a plantearnos qué pasará cuando los ordenadores cobren conciencia de su propia existencia y poder y eliminen a la raza humana de la faz de la tierra, sino que es algo real que está, por ejemplo, en nuestros bolsillos a través de los teléfonos «inteligentes». Se trata de una tecnología que ya afecta a nuestras vidas y a cómo las personas se relacionan, trabajan, votan y viven.

Pueden encontrarse muchos ejemplos de cómo la inteligencia artificial afecta al ejercicio de los derechos humanos, por lo que es perentorio analizar su impacto en profundidad y estudiar si los actuales marcos de protección de los derechos humanos pueden responder a los retos que plantea. No parece apropiado aprobar leyes generales de inteligencia artificial, porque esta puede revestir muchas formas diferentes y aplicarse de maneras muy distintas, pero sí es necesario dar respuesta a las cuestiones prácticas que plantea, por ejemplo, en materia de protección de datos. En este sentido, el objetivo del RGPD de permitirnos entender la lógica de los algoritmos que toman decisiones que nos afectan es positivo pero la viabilidad técnica del mismo es cuestionable porque hay muchos algoritmos que ni siquiera sus programadores entienden en detalle cómo funcionan de ahí la conocida metáfora de la «caja negra» y a esto se añade que en muchos casos su funcionamiento cambia, aprende y se refina por lo que no permanece estable en el tiempo cuando se combina con *deep learning*. A nivel técnico, se está empezando a trabajar

¹² Algunos de los académicos más relevantes en la materia piden a la UE que haga más en el campo de la inteligencia artificial, por ejemplo, que desarrolle una nueva agencia de supervisión responsable de la protección del bienestar público a través de la evaluación científica y la supervisión de productos, software, sistemas o servicios de inteligencia artificial y un observatorio europeo para la inteligencia artificial, así como que incentive financieramente el desarrollo y uso de tecnologías de inteligencia artificial que sean socialmente positivas y favorables para el medio ambiente (FLORIDI ET AL., 2018).

en la explicabilidad y es un ámbito que debe mejorarse para garantizar la seguridad jurídica cuando tiene un impacto en los derechos y obligaciones de las personas.

En este sector tecnológico, es importante tomar decisiones incluyendo a todas las partes interesadas, las empresas, la academia, la sociedad civil y también los grupos más vulnerables a la ampliación de la brecha tecnológica. Es positivo recurrir a la autorregulación o a la corregulación, pero sin excluir que determinados aspectos deben regularse por ley. Además, el recurrir a la ética no puede considerarse como una manera de escapar a la regulación (Wagner, 2018), sino como un elemento que ayude a las empresas y a los Estados en la toma de decisiones dando prioridad a aquellas que produzcan un mayor bien social. Igual que en el ámbito de la privacidad se está avanzando con el RGPD hacia la privacidad desde el diseño y por defecto, en el diseño de programas de inteligencia artificial debería tomarse una perspectiva pro derechos humanos.

Los mecanismos prácticos propuestos por el Relator especial de NNUU van en la buena dirección. Particularmente interesante es la propuesta de evaluaciones de impacto en los derechos humanos, sobre todo en el caso de aquellos programas utilizados por autoridades públicas que pueden afectar al acceso al sistema de salud o judicial, donde es de suma importancia evitar los resultados discriminatorios de manera directa e indirecta. También tiene que seguirse trabajando en los métodos para permitir la auditoria de los algoritmos. Otro aspecto básico es el de reforzar y hacer efectivos los mecanismos de tutela judicial cuando el uso de la inteligencia artificial vulnera los derechos humanos de las personas, para lo que es un prerequisite que las personas sepan que se han tomado decisiones que les afectan basadas en el uso de inteligencia artificial y entender la lógica que hay detrás para lo que debe avanzarse en desarrollar la transparencia, la responsabilidad y la rendición de cuentas.

La UE va por detrás respecto a China o Estados Unidos en lo que al desarrollo de la inteligencia artificial se refiere y es poco probable que les alcance. Sin embargo, sí puede tener un impacto a nivel internacional en lo que se refiere a los estándares que debe cumplir la inteligencia artificial, como ya ha ocurrido en el ámbito de la protección de datos. Es necesario concienciar a las empresas de cómo esto puede hacer más compleja la innovación, pero también presenta un valor añadido.

Hay que permanecer atentos al trabajo de la Comisión Europea y del Consejo de Europa dado que en un momento de ebullición tecnológica hay que introducir este tipo de salvaguardas de los derechos humanos desde el principio.

5. REFERENCIAS BIBLIOGRÁFICAS

- AGENCIA DE LOS DERECHOS FUNDAMENTALES DE LA UNIÓN EUROPEA (2018), *#Big-Data: Discrimination in data-supported decision making*, Oficina de Publicaciones de la Unión Europea, Luxemburgo.
- ANGELOPOULOS, C.; BRODY, A.; HINS, A. W.; HUGENHOLTZ, B.; LEERSSEN, P.; MARGONI, T.; MCGONAGLE, T.; VAN DAALLEN, O.; HOBOKEN, J. V. (2016), *Study of fundamental rights limitations for online enforcement through self-regulation*, Institute for Information Law (IViR), Amsterdam.
- ANGWIN, J.; LARSON, J.; MATTU, S.; KIRCHNER, L. (2016), «Machine Bias: There's Software Used Across the Country to Predict Future Criminals, And It's Biased Against Blacks», *ProPublica*, 23 de mayo de 2016, disponible en: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> (última consulta: 10/11/2018).
- COMISIÓN EUROPEA (2018), *Comunicación de la Comisión al Parlamento Europeo, al Consejo Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones sobre Inteligencia artificial para Europa*, Bruselas, 25 de abril de 2018, COM(2018) 237 final.
- DASTIN, J. (2010), «Amazon scraps secret AI recruiting tool that showed bias against women», *Reuters*, disponible en: <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G> (última consulta: 10/11/2018).
- DELCKER, J. (2018), «Europe's AI ethics chief: No rules yet, please», *Politico.eu*, 30 de octubre de 2018, disponible en: <https://www.politico.eu/article/pekka-ala-pietila-artificial-intelligence-europe-shouldnt-rush-to-regulate-ai-says-top-ethics-adviser/> (última consulta: 10/11/2018).
- DUTTON, T. (2018), «Artificial Intelligence Strategies», *Medium*, 28 de junio de 2018, disponible en: <https://medium.com/politics-ai/an-overview-of-national-ai-strategies-2a70ec6edfd> (última consulta: 10/11/2018).
- EUROPEAN GROUP ON ETHICS IN SCIENCE AND NEW TECHNOLOGIES (2018), *Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems*, Oficina de Publicaciones de la Unión Europea, Bruselas, 9 de marzo de 2018.
- FLORIDI, L.; COWLS, J.; BELTRAMETTI, M.; CHATILA, R.; CHAZERAND, P.; DIGNUM, V.; LUETGE, C.; MADELIN, R.; PAGALLO, U.; ROSSI, F.; SCHAFER, B.; VALCKE, P.; VAYENA, E. (2018), «AI4People – An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations», *Minds and Machines*, Atomium European Institute for Science, Media and Democracy, Bruselas.
- GASCÓN MARCÉN, A. (2014), «La gobernanza de Internet y el Consejo de Europa», *Direito, governança e tecnologia: princípios, políticas e normas do Brasil e da Espanha*, Conceito Editorial, Florianópolis, pp. 115 y ss.

- GOODMAN, B. y FLAXMAN, S. (2016), «European Union Regulations on Algorithmic Decision Making and a Right to Explanation», *2016 ICML Workshop on Human Interpretability in Machine Learning*, New York, NY: ArXiv e-prints.
- KAYE, D. (2018), *Informe del Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión*, Asamblea General de Naciones Unidas, A/73/348.
- MIJATOVIĆ, D. (2018), *Safeguarding human rights in the era of artificial intelligence*, 3 de julio de 2018, disponible en: <https://www.coe.int/en/web/commissioner/-/safeguarding-human-rights-in-the-era-of-artificial-intelligence> (última consulta 20/12/2018).
- MSI-NET (2017), *Study on the human rights dimensions of automated data processing techniques and possible regulatory implications*, Consejo de Europa, Estrasburgo.
- NEMITZ, P. (2018), *Constitutional democracy and technology in the age of artificial intelligence*, Phil. Trans. R. Soc. A 376: 20180089.
- PASQUALE, F. (2015), *The Black Box Society. The Secret Algorithms That Control Money and Information*, Harvard University Press.
- REISMAN, D., SCHULTZ, J., CRAWFORD, K Y WHITTAKER, M. (2018), *Algorithmic impact assessments: a practical framework for public agency accountability*, AI NOW.
- SECRETARIO GENERAL DE LAS NACIONES UNIDAS (2018), *Estrategia en materia de nuevas tecnologías*, Naciones Unidas.
- TUFEKCI, Z. (2018), «YouTube, the Great Radicalizer», *The New York Times*, 10 de marzo de 2018, p. SR6.
- UN GLOBAL PULSE (2012), *Big data for development: Opportunities & challenges*, United Nations Global Pulse, disponible en: www.unglobalpulse.org/sites/default/files/BigDataforDevelopment-UNGlobalPulseJune2012.pdf (última consulta: 10/11/2018).
- VON DER LEYEN, U. (2019), *A Union that strives for more. My agenda for Europe. Political Guidelines for the Next European Commission 2019-2024*, disponible en: <https://www.europarl.europa.eu/resources/library/media/20190716RES57231/20190716RES57231.pdf> (última consulta: 03/08/2019).
- WACHTER, S.; MITTELSTADT, B.; FLORIDI, L. (2016), «Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation», *International Data Privacy Law*, disponible en: <https://ssrn.com/abstract=2903469> (última consulta: 10/11/2018).
- WAGNER, B. (2018), «Ethics as an Escape from Regulation: From ethics-washing to ethics-shopping?», Hildebrandt (Ed.), *Being Profiled. Cogitas ergo sum*, Amsterdam University Press, Amsterdam.