



Derechos humanos e inteligencia artificial

HUMAN RIGHTS AND ARTIFICIAL INTELLIGENCE

Andrea Elena Grigore

Universidad de Sevilla

andreaelenagrigore@gmail.com 0000-0001-6134-8420

Recibido: 13 de diciembre de 2021 | Aceptado: 21 de marzo de 2022

RESUMEN

La inteligencia artificial (IA) 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. 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 se presentan las acciones llevadas a cabo en el marco universal por las Naciones Unidas y en el ámbito europeo por el Consejo de Europa y la Unión Europea. La Inteligencia Artificial es una herramienta sumamente poderosa y necesaria para ayudarnos a abordar los inmensos retos que debemos superar como especie. Pero si a la hora de desarrollarla pasamos por alto los factores sociales, laborales y éticos, la Inteligencia Artificial puede convertirse en arma a favor de la desigualdad, el control y la destrucción.

ABSTRACT

Artificial intelligence (AI) can have a positive impact on the exercise of certain rights, but others can be undermined by its use such as: freedom of expression (including the right to freely communicate or receive information), privacy, the prohibition of discrimination, and the right to effective judicial protection. International organizations dedicated to the protection of human rights are currently reflecting on these risks and possible solutions to help their member states respond to these challenges, as well as the actions carried out in the universal framework by the United Nations and at the European level by the Council of Europe and the European Union. Artificial Intelligence is an extremely powerful and necessary tool to help us address the immense challenges we must overcome as a species. But if when developing it we ignore social, labor and ethical factors, Artificial Intelligence can become a weapon in favor of inequality, control and destruction.

PALABRAS CLAVE

Inteligencia artificial
Ética
Derechos humanos

KEYWORDS

Artificial intelligence
Ethics
Human rights

POSICIÓN DE LA AUTORA: En el contexto actual es preciso más que nunca poner de manifiesto el valor de los derechos humanos, respecto al continuo uso de la IA, que se considera como elemento estratégico del siglo XXI. Afecta a todos los ámbitos: desde la geopolítica hasta los intereses comerciales, pasando por el crecimiento económico, la lucha contra el cambio climático o la mejora de la atención sanitaria, lo que en cierto modo nos hace incluso dependientes de ella. No es objetiva, neutral o universal; por el contrario, está profundamente arraigada en la cultura y la realidad económica de quienes la crean, en su mayoría, hombres blancos y con recursos. Nuestro futuro como humanidad depende de cómo la desarrollemos para que el impacto sea positivo. Todas las transformaciones sociales transcurren a través de una mejoría y avance imprescindible de los derechos humanos. La comunidad académica tiene el deber de reflexionar y participar en investigaciones sobre los efectos sociales, económicos y políticos de la personalización del contenido de la IA.

I. CONTEXTO ACTUAL

“La injusticia, en cualquier parte, es una amenaza a la justicia en todas partes”. Ya lo apuntaba Martin Luther King referente a los derechos civiles en los 70.

Según NIKKEN¹ la noción de derechos humanos se corresponde con la afirmación de la dignidad de la persona frente al Estado. El poder público debe ejercerse al servicio y el bienestar del ser humano: no puede ser empleado lícitamente para ofender atributos inherentes a la persona y debe ser vehículo para que ella pueda vivir en sociedad en condiciones cónsonas con la misma dignidad que le es consustancial.

Una de las características resaltantes del mundo contemporáneo es el reconocimiento de que todo ser humano, por el hecho de serlo, es titular de derechos fundamentales que la sociedad no puede arrebatarse lícitamente. Estos derechos no dependen de su reconocimiento por el Estado ni son concesiones suyas; tampoco dependen de la nacionalidad de la persona ni de la cultura a la cual pertenezca. Son derechos universales que corresponden a todo habitante de la tierra. La expresión más notoria de esta gran conquista es el artículo 1 de la Declaración Universal de Derechos Humanos: todos los seres humanos nacen libres e iguales en dignidad y derechos y, dotados como están de razón y conciencia, deben comportarse fraternalmente los unos con los otros (Nikken, 1994).

¿Cuál es la relevancia de la inteligencia artificial (IA) para los derechos humanos? Definir la IA no es fácil, ya que el concepto de inteligencia *per se* no es del todo preciso. En términos coloquiales, IA se usa cuando una máquina es capaz de imitar las funciones cognitivas propias de la mente humana, como: creatividad, sensibilidad, aprendizaje, entendimiento, percepción del ambiente y uso del lenguaje. Esta tecnología es una herramienta extraordinariamente poderosa que podría tener un efecto transformador e

1. Ex Presidente del Consejo Directivo del IIDH y Ex Presidente de la Corte Interamericana de Derechos Humanos. Profesor de la Universidad Central de Venezuela.

innovador en muchos aspectos de la vida cotidiana, desde el transporte y la fabricación hasta la atención médica y la educación.

Su uso está aumentando en todos estos sectores, así como en el sistema de justicia, la policía y el ejército. La IA puede aumentar la eficiencia, generar nuevos conocimientos sobre las enfermedades y acelerar el descubrimiento de medicamentos novedosos. La inteligencia artificial se usa mucho para crear recomendaciones personalizadas para los consumidores, basadas, por ejemplo, en sus búsquedas y compras previas o en otros comportamientos en línea. La IA es muy importante en el comercio, para optimizar los productos, planear el inventario, procesos logísticos, etc. El desarrollo de tecnologías inteligentes impacta profundamente en la sociedad.

El desarrollo de la IA se ha acelerado gracias a los incentivos tecnológicos, financieros y humanos actuales. Existen avances importantes de las tecnologías inteligentes en ámbitos como agricultura, manufactura, medicina, educación, vehículos autónomos y entretenimiento. Además, sus implicaciones de carácter social, económico, ético y legal son temas de reflexión y debate alrededor del mundo.

Pero su uso indebido, intencional o no, también puede dañar los derechos de las personas. En los ámbitos ético y legal se debaten las asignaciones de responsabilidades y obligaciones en situaciones donde intervienen sistemas inteligentes. Por ejemplo, si un automóvil autónomo atropella a un peatón, la responsabilidad del acto podría recaer en el dueño del vehículo; pero también debe reconocerse que éste no tiene un control completo del automóvil y la responsabilidad también podría recaer en el dueño intelectual de la IA. Los sistemas inteligentes pueden aprender prejuicios desde los mismos datos. Un ejemplo es el de experimentos en procesamiento automático del lenguaje, que muestran cómo los sistemas aprenden a asociar la palabra mujer con profesiones más cercanas a las humanidades y al hogar, mientras que a la palabra hombre, la relacionan con profesiones con un mayor componente matemático.

Otro caso ilustrativo es el de Tay, un bot para Twitter desarrollado por Microsoft (un bot es un programa que ejecuta tareas en internet), capaz de interactuar automáticamente con usuarios y aprender de ellos. En pocas horas, Tay aprendió expresiones racistas, xenofóbicas, misóginas y fascistas y tuvo que ser bloqueado de la red social. Sin un modelo transparente, es muy difícil hacer valer condiciones sociales con estas tecnologías, como el respeto hacia otros usuarios. Para prevenir que los sistemas hereden tantos prejuicios y que a su vez representen al grueso de la población, se hacen esfuerzos para que exista igualdad de género en los desarrolladores (Hunt, 2016)

En el aspecto ético, destacan los requisitos que deben tener los sistemas con el fin de prevenir prejuicios y reacciones negativas en la sociedad. En particular, la transparencia y la confiabilidad juegan un papel crucial. Asimismo, el acceso equitativo de mujeres y hombres ha sido impulsado por algunos de los actores con mayor injerencia en el campo, especialmente en América Latina. Esto puede fomentar que los sistemas cuenten con una visión menos sesgada y más representativa del grueso de la población. Por último, el futuro de la IA concierne también al de sus interacciones con otras áreas del conocimiento y otras tecnologías. En particular, el cómputo cognitivo se beneficiará de los avances en el entendimiento del funcionamiento del cerebro humano

y las neurociencias, así como del desarrollo de procesadores, sistemas de almacenamiento de datos y los ambientes de realidad virtual y realidad aumentada.²

La Alta Comisionada de las Naciones Unidas para los Derechos Humanos, Michelle Bachelet, enfatizó en GINEBRA (15 de septiembre de 2021) la urgente necesidad de establecer una moratoria a la venta y el uso de los sistemas de inteligencia artificial (IA) que amenazan gravemente a los derechos humanos, hasta que se hayan puesto en vigor las salvaguardas adecuadas. Bachelet pidió además que se prohíban los programas de AI cuyo uso no cumpla con la normativa internacional de derechos humanos.

“La inteligencia artificial puede ser una fuerza para el bien, que ayude a las sociedades a superar algunos de los mayores retos de nuestro tiempo. Pero las tecnologías de IA también pueden tener efectos nocivos e incluso catastróficos, cuando se emplean sin prestar la debida atención a su capacidad de vulnerar los derechos humanos”, declaró Bachelet.

“Mientras mayor sea el riesgo para los derechos humanos, más estrictos deben ser los requisitos legales para el uso de la tecnología de IA”, añadió la Alta Comisionada. “Pero dado que la evaluación de los peligros y cómo superarlos puede llevar todavía algún tiempo, los Estados deberían implementar desde ahora moratorias sobre el uso de las tecnologías que presentan un alto potencial de riesgo”.³

Como parte de su labor en materia de tecnología y derechos humanos, la Oficina de Derechos Humanos de la ONU ha publicado un informe en el que se analiza cómo la IA –incluidas la elaboración automática de perfiles, la toma de decisiones y otras tecnologías de aprendizaje para las máquinas– afecta al derecho a la intimidad y a otros derechos, incluidos los relativos a la salud, la educación, la libertad de movimiento, la libertad de reunión y asociación pacífica, y la libertad de expresión.

“La inteligencia artificial llega ahora a casi todos los rincones de nuestras vidas, e incluso alcanza a nuestros estados emocionales. Los sistemas de IA se usan para decidir quién recibe servicios públicos, quién tiene oportunidad de obtener un empleo y, por supuesto, afectan a la información que la gente ve y que puede compartir en Internet”, señaló la Alta Comisionada.

En el informe se examina cómo a menudo los Estados y las empresas se han precipitado a incorporar los programas informáticos de IA sin ejercer la diligencia debida. Ha habido muchos casos de personas que han sido injustamente tratadas a causa de la IA, ya sea porque se les ha denegado prestaciones del seguro social debido a la aplicación de instrumentos deficientes de IA o porque han sido arrestadas como consecuencia de errores en los sistemas de reconocimiento facial⁴.

2. Meetings of the Expert Group on Business-to-Government Data Sharing, Digital Single Market, 2019. <https://ec.europa.eu/digital-single-market/en/news/meetings-expert-group-businessgovernment-data-sharing>

3. <https://www.ohchr.org/SP/HRBodies/HRC/Pages/NewsDetail.aspx?NewsID=27469&LangID=5>(consultado diciembre 2021)

4. OECD, (2019). Recommendation of the Council on Artificial Intelligence. Pub. L. No. OECD/LEGAL/0449. <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>

En el informe se describe detalladamente cómo los sistemas de IA se basan en grandes bases de datos, que contienen información personal compilada, intercambiada, combinada y analizada con métodos diversos y a veces opacos. Los datos que alimentan y orientan a los sistemas de IA pueden ser deficientes, discriminatorios, obsoletos o poco pertinentes. El almacenamiento de datos a largo plazo entraña también riesgos particulares, ya que en el futuro esa información podría aprovecharse de formas difíciles de prever en la actualidad.⁵

“Dado el rápido y continuo crecimiento de la IA, llenar el inmenso vacío de asunción de responsabilidades sobre cómo se recogen, almacenan, comparten y utilizan los datos es uno de los cometidos más urgentes que afrontamos en materia de derechos humanos”, dijo Bachelet.

Las deducciones, predicciones y actividades de seguimiento que llevan a cabo las herramientas de IA, incluida la búsqueda de pautas de conducta humana, plantean también graves interrogantes. Las bases de datos sesgadas en los que a menudo se basan los sistemas de IA pueden conducir a la adopción de decisiones discriminatorias, un riesgo que es más acentuado cuando se trata de grupos que ya se encuentran marginados.

“El riesgo de discriminación inherente a las decisiones basadas en la IA –decisiones capaces de cambiar, definir o perjudicar a las personas– es muy real. Por eso es necesario evaluar y monitorear sistemáticamente las repercusiones de los sistemas de IA, para poder identificar y atenuar los riesgos en materia de derechos humanos”, añadió Bachelet.

“Las tecnologías biométricas, que se están convirtiendo en un comodín para los Estados, las organizaciones internacionales y las empresas de tecnología, constituyen un ámbito que necesita urgentemente más directrices de derechos humanos”, señala el informe.

Esas tecnologías, que comprenden el reconocimiento facial, se usan cada vez más para identificar a determinadas personas a distancia y en tiempo real, lo que puede permitir un seguimiento ilimitado de los individuos. El informe reitera el llamamiento a una moratoria de su uso en espacios públicos, al menos hasta que las autoridades demuestren que no existen problemas significativos con la precisión o los efectos discriminatorios, y que esos sistemas de IA cumplen con las normas aplicables a la protección de datos y a la privacidad.

También es preciso que las empresas y los Estados incrementen la transparencia en lo que se refiere al desarrollo y la aplicación de los sistemas de IA.

“La complejidad del contexto estadístico, los algoritmos y los modelos en los que se basa la elaboración y el funcionamiento de los sistemas de IA, así como el carácter secreto que le confieren los gobiernos y los agentes privados, contribuyen a que la población en general no alcance a comprender la repercusión de los sistemas de IA sobre la sociedad y los derechos humanos”, apunta el informe.

“No podemos seguir reaccionando tardíamente a los efectos de la IA, ni permitir que se use de manera ilimitada, sin fronteras ni supervisión, para enfrentarnos luego a sus

5. Human Rights Council, (13 September–1 October 2021). Annual report of the United Nations High Commissioner for Human Rights and reports of the Office of the High Commissioner and the Secretary-General .

casi inevitables consecuencias sobre los derechos humanos. La capacidad de la IA para servir a la población es innegable, pero también lo es su capacidad de contribuir a violaciones de derechos humanos en gran escala, de manera casi indetectable. Es necesario adoptar urgentemente medidas para imponer límites basados en los derechos humanos a la utilización de la IA por el bien de todos”, terminó diciendo Bachelet.⁶

II. MARCO LEGAL

Las preocupaciones por los dilemas éticos en el marco de la inteligencia artificial (IA), constituye un aspecto ampliamente debatido en las condiciones actuales. Estas inquietudes se centran en el empleo de herramientas de Big Data, y en los riesgos políticos que se derivan de su uso mal intencionado, al ser empleadas para el fomento de la manipulación y la distorsión de procesos sociales.

Para abordar las preocupaciones relacionadas con la inteligencia artificial y los DD.HH, Human Rights Watch, Access Now, Amnistía Internacional y una coalición de organizaciones de derechos humanos y tecnología emitieron una declaración histórica que propone estándares de derechos humanos para el aprendizaje automático, presentando la Declaración de Toronto sobre la protección del derecho a la igualdad y la no discriminación en los sistemas de aprendizaje automático.

La Declaración de Toronto es un documento histórico que busca aplicar las normas internacionales de derechos humanos existentes al desarrollo y uso de sistemas de aprendizaje automático (o “inteligencia artificial”).

La Declaración hace un llamamiento a los gobiernos y a las empresas para garantizar que las aplicaciones de aprendizaje automático e inteligencia artificial respeten los principios de igualdad y no discriminación. El documento articula normas de derechos humanos que deben cumplir los sectores público y privado para garantizar que los algoritmos utilizados se apliquen equitativa y justamente, y que aquellos que creen sus derechos han sido violados tengan una vía significativa para la reparación. Si bien ha habido un diálogo sólido sobre ética e inteligencia artificial, la Declaración enfatiza la importancia y aplicabilidad de la ley de derechos humanos, que está diseñada para proteger los derechos y proporcionar remedios donde los seres humanos son perjudicados.

La Declaración se centra en el aprendizaje automático y los derechos a la igualdad y la no discriminación, pero muchos de los principios se aplican a otros sistemas de inteligencia artificial. Además, el aprendizaje automático y la inteligencia artificial afectan a una amplia gama de derechos humanos, incluidos el derecho a la privacidad, la libertad de expresión, la participación en la vida cultural, el derecho a la reparación y el derecho a la vida.

El aprendizaje automático (AA) es un subconjunto de la inteligencia artificial. Se puede definir como “proporcionar a los sistemas la capacidad de aprender y mejorar de forma automática a partir de las experiencias, sin necesidad de programación explícita”.

6. Declaración de la Alta Comisionada sobre los riesgos de la inteligencia artificial para la privacidad en el Consejo de Europa (14 de septiembre de 2021).

Uno de los riesgos más importantes que presenta el aprendizaje automático es el peligro de amplificar la discriminación y los sesgos existentes contra ciertos grupos, a menudo comunidades marginadas y vulnerables, que ya batallan para que se les trate con dignidad y respeto. Cuando se utilizan datos históricos para entrenar a los sistemas de aprendizaje automático sin salvaguardias, estos sistemas pueden reafirmar o incluso aumentar el sesgo estructural existente. También se pueden causar daños discriminatorios cuando las decisiones tomadas al diseñar los sistemas de IA producen resultados sesgados, ya sea de forma deliberada o no. Un ejemplo concreto y actual es la polémica creada porque la IA podría llegar a ser discriminatoria.

Un ejemplo claro de ello es la discriminación algorítmica –por errores de datos– en la concesión de hipotecas o la evaluación de inquilinos. En 2019, en el estado de Connecticut (EE UU) una agencia inmobiliaria que disponía de un software de evaluación individualizada de inquilinos para automatizar la concesión de alquileres consideró que una madre no podía alquilar un piso porque su hijo, con diversidad funcional, había sido años antes arrestado por robo erróneamente. El software de la empresa incluía a todas las personas que habían sido arrestadas, pero no si estas habían sido finalmente acusadas o no, o si ese arresto se había retirado por un error, como fue este caso. Pese a la batalla judicial entre una asociación de derechos de la vivienda y la empresa, la Corte del distrito federal determinó que la empresa de evaluación automatizada de inquilinos estaba amparada bajo la regulación de la FHA o Administración Federal de Viviendas. El caso está nuevamente en curso a día de hoy y ejemplos así se pueden extrapolar tanto en España que en el resto de Europa.

El informe anual “*Automating Society*” de la organización europea *AlgorithmWatch* pone de manifiesto las tendencias de la toma automatizada de decisiones mediante algoritmos en la Unión Europea y sus Estados miembros. En España, entre otros, el SEPE (el Servicio de Empleo Público de España) utiliza un sistema automatizado para calcular los beneficios por desempleo, así como distribuir las ofertas de empleo, entrevistas y cursos de formación. Si bien se desconoce el impacto de este sistema automatizado y si sus resultados serían distintos a los realizados por una persona, a nivel técnico sí se sabe que existe un reto, que es el de la limitada reconciliación de datos: es decir, en el momento en que haya un error en un dato o una casilla en un archivo, dejará de ser compatible con el resto de archivos y el software lo asignará como nulo. El efecto en el desembolso del paro puede ser importante si no se previene este riesgo.

Cómo observación al respecto podemos señalar que, en gran medida, las normas de derechos humanos no estaban presentes en el debate sobre la ética de la IA; en varias ocasiones se mencionaba la importancia de los derechos humanos, pero por lo general no era más que una referencia pasajera. Y en algunos lugares, la discusión sobre la ética de la IA estaba comenzando a tomar un giro ya conocido: se argumentaba que la ética de la IA debería depender de la cultura. Esto abre la puerta a diferentes estándares y diferentes niveles de protección para los derechos de las personas; y en el contexto de una tecnología digital como la IA, a la que no le interesan las fronteras, no tiene mucho sentido.

Era evidente que afirmar el papel central del marco de derechos humanos en la ética de la IA era una prioridad, centrando el foco de atención en la igualdad y la no discrimi-

nación en el aprendizaje automático porque se trata de un tema ya urgente que causa cada vez más problemas en la vida real.

La Declaración de Toronto fue redactada después de discusiones y entrevistas con docenas de expertos en IA, empresas y derechos humanos, y otros temas. Se realizaron amplias consultas sobre el borrador, y el documento se adoptó el 16 de mayo, al inicio de RightsCon 2018. La Declaración tiene tres secciones principales.

Primero, establece el deber de los Estados de prevenir la discriminación en el diseño o la implementación de los sistemas de aprendizaje automático en contextos públicos o mediante asociaciones público-privadas. Esta sección incluye principios para identificar riesgos en el uso de sistemas de aprendizaje automático, para garantizar la transparencia y la rendición de cuentas (incluida la divulgación al público de dónde se usan los sistemas de aprendizaje automático), para hacer cumplir la supervisión (incluidos los mecanismos para la supervisión independiente) y para promover la igualdad. A nivel Internacional, buscar un punto de encuentro sobre el que construir una visión común del desarrollo tecnológico debe estar en todas las agendas diplomáticas. La política exterior ha de aspirar a una Declaración de derechos humanos actualizada a la era de la inteligencia artificial. Los Estados pueden elaborar su rol en la Gobernanza e la IA refiriéndose a las normas internacionales de derechos humanos, y desarrollando mecanismos para la rendición de cuentas, la trsparencia y la reparación de violaciones y abusos.

En segundo lugar, la Declaración describe las responsabilidades de los actores privados en el contexto del desarrollo y la implementación de sistemas de AA. Se basa en el marco de diligencia debida en materia de derechos humanos (que se definió originalmente en los Principios Rectores de la ONU sobre las Empresas y los Derechos Humanos). Estas responsabilidades incluyen detectar posibles resultados discriminatorios a través de la identificación y evaluación de riesgos, tomar medidas efectivas para prevenir y mitigar la discriminación (entre ellas, someter a los sistemas a auditorías externas independientes cuando haya un riesgo considerable de que se violen los derechos humanos), y ser transparentes, lo que incluye la publicación de especificaciones técnicas, muestras de los datos de entrenamiento utilizados y fuentes de datos.

En tercer lugar, la Declaración afirma el derecho a un recurso efectivo y a llamar a cuentas a los responsables de las violaciones. Exhorta a los gobiernos a que garanticen estándares de debido proceso para el uso del aprendizaje automático en el sector público, que actúen con cautela al utilizar sistemas de AA en el sistema penal, que definan líneas claras de rendición de cuentas para el desarrollo y la implementación de aplicaciones de AA y que aclaren cuáles son los organismos o personas jurídicamente responsables de las decisiones tomadas mediante el uso de dichos sistemas.

Redactar y publicar la Declaración es solo el primer paso para que el marco de derechos humanos se convierta en un componente fundamental del ámbito de la ética de la IA y los datos, el cual está experimentando un rápido desarrollo. La Declaración establece principios para la política y la práctica, pero para hacerlos efectivos necesitamos guías prácticas de implementación que ayuden a los ingenieros y gerentes de productos a aplicar los principios en su trabajo.

Además, es importante que la sociedad civil dé su más amplio respaldo a la Declaración de Toronto para afirmar la centralidad de los derechos humanos en el debate sobre la ética de la IA. Esta sería una señal importante de que se espera que las empresas de tecnología también apoyen la Declaración y se comprometan a velar por que las nuevas tecnologías no diluyan los derechos humanos existentes. Acercarse más a la comunidad tecnológica para entablar un diálogo entre los defensores de derechos humanos y los ingenieros ayudará a incorporar los derechos humanos en el uso y desarrollo de la IA.

Las personas tenemos derecho a poder confiar en el progreso y a no temer a la tecnología. Tenemos que exigir que esta respete nuestros valores democráticos y de convivencia. Nuestro país, España, ha sido pionero a nivel mundial en presentar una Carta de Derechos Digitales de la Ciudadanía, un marco de referencia para poder navegar este entorno digital en el que vivimos, que busca contribuir a una reflexión global sobre el papel que ha de ocupar la persona en el desarrollo tecnológico, que garantice una digitalización humanista.

Finalmente, debemos impulsar la información equilibrada sobre las tecnologías de IA: al hablar de ellas como tecnologías que se pueden usar para el bien, pero de las que también se puede abusar. Describir la IA mediante visiones distópicas o utópicas no conduce a un debate sensato, y nos distrae de abordar los riesgos reales y de aprovechar las oportunidades de utilizar la IA en nuestro beneficio.

III. CONCLUSIONES

Los repetidos escándalos públicos de las grandes empresas y las redes sociales han conseguido poner de manifiesto la necesidad de promover e inculcar un interés amplio y genuino en las cuestiones éticas en torno a la IA, no solo entre los académicos, sino también en muchas empresas que están empezando a desarrollar un movimiento que buscaba responder de forma proactiva a los riesgos asociados con la IA.

La inteligencia artificial es una tecnología estratégica que ofrece numerosas ventajas a los ciudadanos, las empresas y la sociedad en su conjunto, siempre y cuando sea antropocéntrica, ética y sostenible y respete los derechos y valores fundamentales. La IA aporta importantes mejoras de la eficiencia y la productividad que pueden reforzar la competitividad de la industria europea y mejorar el bienestar de los ciudadanos. También puede contribuir a encontrar soluciones a algunos de los problemas sociales más acuciantes, como la lucha contra el cambio climático y la degradación medioambiental, los retos relacionados con la sostenibilidad y los cambios demográficos, la protección de nuestras democracias y, cuando sea necesario y proporcionado, la lucha contra la delincuencia.

En un contexto de feroz competencia mundial, se requiere un enfoque europeo sólido basado en la Estrategia Europea para la IA presentada en abril de 2018⁷. Para aprovechar las oportunidades que ofrece la inteligencia artificial y abordar los retos que

7. Inteligencia artificial para Europa [COM(2018) 237 final].

presenta, la UE debe actuar conjuntamente y determinar de qué manera, a partir de los valores europeos, promoverá su desarrollo y adopción.⁸

La Comisión se ha comprometido a facilitar el avance científico, preservar el liderazgo tecnológico de la UE y garantizar que las nuevas tecnologías estén al servicio de todos los europeos, de manera que mejoren sus vidas al mismo tiempo que respetan sus derechos. La presidenta de la Comisión, Ursula von der Leyen, anunció en sus orientaciones políticas un enfoque europeo coordinado en torno a las implicaciones éticas y humanas de la inteligencia artificial y un análisis sobre cómo mejorar la utilización de los macrodatos en la innovación.⁹

En la Unión Europea, el Libro Blanco de IA de 2020 identificaba dos niveles de riesgo –bajo, alto– en las aplicaciones de IA (que no de las decisiones automatizadas en general). Ante esta propuesta, la Comisión de Ética de Datos de Alemania propuso ampliarlo a cinco niveles de riesgo para dotarle de mayor precisión y “justicia algorítmica”. En la reciente publicación de la Estrategia de IA en abril de 2021, la UE ha ampliado el número de niveles hasta cuatro (riesgo inaceptable, alto, limitado, mínimo). Sin embargo, todavía hoy en día las auditorías de las ADM suelen realizarse de forma limitada y, además, no existe una metodología común de cómo prevenir la discriminación algorítmica.

El crecimiento económico sostenible y el bienestar social presentes y futuros de Europa se valen cada vez más de los valores creados por los datos. La inteligencia artificial es una de las partes más importantes de la economía de los datos. Hoy en día, la mayor parte de los datos son relativos a los consumidores y se almacenan y tratan en infraestructuras ubicadas en nubes centralizadas. Frente a esto, una enorme proporción de los datos del futuro, que serán mucho más abundantes, procederá de la industria, las empresas y el sector público, y se almacenará en diversos sistemas, entre los que destacan los dispositivos informáticos que operan en el borde de la red. Este hecho ofrece nuevas oportunidades a Europa, que cuenta con una posición sólida en la industria digitalizada y las aplicaciones de comunicación empresarial, pero con una posición relativamente frágil en las plataformas de consumidores.

En otras palabras, la inteligencia artificial es una combinación de tecnologías que agrupa datos, algoritmos y capacidad informática. Los avances en computación y la creciente disponibilidad de datos son, por tanto, un motor fundamental en el pronunciado crecimiento actual de la inteligencia artificial. Europa puede aunar su potencial tecnológico e industrial con una infraestructura digital de gran calidad y un marco regulador basado en sus valores fundamentales para convertirse en líder mundial de la innovación en la economía de los datos y sus aplicaciones, tal como se establece en la Estrategia Europea de Datos.¹⁰

8. LIBRO BLANCO sobre la inteligencia artificial - un enfoque europeo orientado a la excelencia y la confianza https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_es.pdf

9. https://ec.europa.eu/commission/commissioners/sites/default/files/commissioner_mission_letters/president-elect_von_der_leyens_mission_letter_to_thierry_breton.pdf

10. COM(2020) 66 final a European strategy for data, de 19 de febrero de 2020. Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones

Sobre estos cimientos, puede desarrollar un ecosistema de inteligencia artificial que acerque las ventajas de la tecnología a la sociedad y a la economía europea en su conjunto:

- a los ciudadanos, para que obtengan nuevos beneficios, como una mejor atención sanitaria, una menor cantidad de averías de los aparatos domésticos, unos sistemas de transporte más seguros y limpios, o mejores servicios públicos;
- al desarrollo empresarial, por ejemplo, mediante una nueva generación de productos y de servicios en áreas en las que Europa es particularmente fuerte (maquinaria, transporte, ciberseguridad, agricultura, economía verde y circular, atención sanitaria y sectores de gran valor añadido, como la moda y el turismo); y
- a los servicios de interés público, por ejemplo, mediante una reducción de los costes de la prestación de servicios (transporte, educación, energía y gestión de los residuos), una mayor sostenibilidad de los productos¹¹, o proporcionando a los servicios y fuerzas de seguridad las herramientas adecuadas para que aseguren la protección de los ciudadanos¹² garantizando correctamente el respeto de sus derechos y libertades. Teniendo en cuenta el enorme impacto que puede tener la inteligencia artificial en nuestra sociedad y la necesidad de que suscite confianza, resulta clave que la inteligencia artificial europea se asiente en nuestros valores y derechos fundamentales, como la dignidad humana y la protección de la privacidad.

El tema de los derechos humanos domina progresivamente la relación de la persona con el poder en todos los confines de la tierra. Su reconocimiento y protección universal representa una revalorización ética y jurídica del ser humano como poblador del planeta más que como poblador del Estado. Los atributos de la dignidad de la persona humana, donde quiera que ella esté y por el hecho mismo de serlo prevalecen no solo en el plano moral sino en el legal, sobre el poder del Estado, cualquiera sea el origen de ese poder y la organización del gobierno. Es esa la conquista histórica de estos tiempos (Nikken, 1994).

11. La inteligencia artificial y la digitalización en general son motores clave de las ambiciones contempladas en el Pacto Verde Europeo. No obstante, se estima que la huella medioambiental actual del sector de las TIC se sitúa por encima del 2 % del conjunto de emisiones mundiales. La Estrategia Digital Europea que acompaña al presente Libro Blanco propone medidas de transformación ecológica para el sector digital.

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

Bibliografía

- Nikken, P. (1994). El Concepto de Derechos Humanos, Estudios Básicos de Derechos Humanos, IIDH, San José.
- Brooks R. A., (2003).Cuerpos y máquinas, de los robots humanos a los hombres robots, trad. De Guillermo Solana, Barcelona: Ediciones B.
- Steels, L., (1994). The Artificial Life Roots of Artificial Intelligence, en Artificial Life.
- Hunt, E (2016). Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter.
- Bellosos Martín, N. (2018). La necesaria presencia de la ética en la robótica: La roboética y su incidencia en los derechos humanos. Cuadernos de Programa de Pós-Graduação. Vol. 13, nº. 2.
- Pérez Orozco, B (2017). Entrevista con Brent Mittelstadt.
[-https://www.foroconsultivo.org.mx/INCYTU/documentos/Completa/INCYTU_18-012.pdf](https://www.foroconsultivo.org.mx/INCYTU/documentos/Completa/INCYTU_18-012.pdf)
- Meetings of the Expert Group on Business-to-Government Data Sharing, Digital Single Market, (2019).<https://ec.europa.eu/digital-single-market/en/news/meetings-expert-group-businessgovernment-data-sharing>
- OECD, (2019). Recommendation of the Council on Artificial Intelligence. , Pub. L. No. OECD/LEGAL/0449 . <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>.
- HumanRightsCouncil,(2021).AnnualreportoftheUnitedNationsHighCommissionerforHumanRightsandreports of the Office of the High Commissioner and the Secretary-General <https://www.ohchr.org/en/reports-secretary-general-and-high-commissioner-human-rights>
- Declaración de la Alta Comisionada sobre los riesgos de la inteligencia artificial para la privacidad en el Consejo de Europa, (2021). <https://www.ohchr.org/es/press-releases/2021/09/artificial-intelligence-risks-privacy-demand-urgent-action-bachelet>