

Francisco José Santamaría Ramos

Implicaciones éticas y legales relacionadas con el uso creciente de la IA

Abstract: Related Ethical and Legal Implications with the Increasing Use of AI. This paper examines the ethical implications of artificial intelligence, emphasizing the importance of a responsible and ethical approach in its development and application. Through the analysis of the Montreal Declaration and the UNESCO Recommendation, reflections are made on the regulation and the role ethics should play in the design and development of artificial intelligence, bearing in mind that the core objective of this technology is to be a benefit to society while respecting fundamental human values and rights.

Keywords: Artificial Intelligence, Morality, Ethics, Human Rights.

RESUMEN: 1. Introducción – 2. Definiendo la inteligencia artificial – 3. La Declaración de Montreal: Análisis y Crítica – 3.1. Principio de bienestar – 3.2. Principio de respeto a la autonomía – 3.3. Principio de protección de la privacidad y la intimidad – 3.4. Principio de solidaridad – 3.5. Principio de participación democrática – 3.6. Principio de equidad – 3.7. Principio de inclusión de la diversidad – 3.8. Principio de prudencia – 3.9. Principio de responsabilidad – 3.10. Principio de desarrollo sostenible – 4. UNESCO: Recomendación sobre la ética de la inteligencia artificial – 4.1. Introducción – 4.2. Evaluación del impacto ético – 4.3. Gobernanza y administración éticas – 4.4. Política de datos – 4.5. Desarrollo y cooperación internacional – 4.6. Medio ambiente y ecosistemas – 4.7. Género – 4.8. Cultura – 4.9. Educación e investigación – 4.10. Comunicación e información – 4.11. Economía y trabajo – 4.12. Salud y bienestar social – 5. Conclusiones.

1. Introducción

La sociedad actual asiste expectante a una revolución de carácter tecnológico que no sólo promete unos grandes beneficios al ser humano, sino que también plantea una serie de preguntas sobre cómo se van a implementar esos cambios y, sobre todo, cuáles van a ser las consecuencias de estos.

No puedo sino coincidir con el economista e ingeniero alemán Klaus Schwab que asistimos a: “La cuarta revolución industrial donde se plasma que el ser humano asiste a una nueva

transformación de la humanidad que está cambiando por completo la forma de vivir, trabajar y relacionarse”¹.

La presente revolución es consecuencia necesaria de “la confluencia de numerosos avances tecnológicos en diversos campos: impresión 3D, vehículos autónomos, biotecnología, computación cuántica, etc”². Desde el punto de vista que más nos interesa, para el presente trabajo, el avance tecnológico de importancia es la Inteligencia Artificial.

Es más, en la actualidad podemos hablar de las denominadas tecnologías: “NBIC (nano-bio-info-cogno): nanotecnologías, biotecnologías, tecnologías de la información y ciencias cognitivas (IA, ciencias de datos, robótica, interfaces cerebro-máquina, biología sintética, nanotecnología)”³.

Atrás queda octubre de 1950 cuando Alan Turing se hizo la pregunta que cambió radicalmente el mundo de la informática: “¿Pueden pensar las máquinas?”⁴. En menos de un siglo hemos pasado de pensar que el Test de Turing era una mera fábula a *democratizar* el uso de la inteligencia artificial a través de sistemas como *Chat GPT* o *Midjourney*, entre otros.

Hablar sobre cómo la inteligencia artificial ha penetrado a nivel social, supone examinar el impacto de la presente tecnología en diversos aspectos de la vida humana y, al mismo tiempo, supone hablar de los desafíos y oportunidades de una tecnología que, a pesar de que cada vez nos sorprende más, no podemos olvidar que estamos en presencia de una tecnología emergente.

Como el lector puede suponer, el presente trabajo no puede examinar de forma detallada todas y cada una de las áreas sociales en las que la inteligencia artificial va a impactar. Sin embargo, lo que sí puedo hacer es enumerar las que, en mi opinión, considero más importantes, aunque no por ello las únicas:

¹ K. Schwab, *La cuarta revolución industrial*, Penguin Random House Grupo Editorial, Barcelona, 2016.

² F.J. Santamaría Ramos, “Robótica, Inteligencia Artificial y Desempleo”, in J.P. Pampillo Baliño y S. Botero Gómez (coordinadores), *Justicia Social global: Perspectivas, reflexiones y propuestas desde Iberoamérica*, Tirant Lo Blanch, Ciudad de México, 2022, pp. 425-458.

³ T. Ausín, “¿Por qué ética para la Inteligencia Artificial? Lo viejo, lo nuevo y lo espurio”, in *Sociología y Tecnología*, 11 (2021), p. 2.

⁴ A.M. Turing, “Computing machinery and intelligence”, in *Mind*, (1950), n. 49, pp. 433-460.

- *Empleo*: Sin duda, el uso de la inteligencia artificial en los mercados laborales va a marcar un antes y un después en la historia del ser humano y en la sociedad del siglo XXI. Hoy en día, la pérdida de empleos en algunos sectores es una realidad. Incluso, sin ser absolutamente pesimista, incluso sin pensar que el ser humano se puede ver desprovisto de su activo laboral, no podemos negar que la inteligencia artificial trae consigo desplazamiento laboral e incluso la necesidad de adaptación y capacitación del empleado para obtener un *trabajador del siglo XXI* acorde a las expectativas del mercado laboral.

- *Brechas*: Un uso desigual de la inteligencia artificial puede generar brechas tanto sociales como económicas. Hablar de brechas supone hablar de una desigualdad desde el punto de vista del acceso a las tecnologías y, sobre todo, a las oportunidades. Todo ello, nos conduciría a hablar de la posibilidad real de que se produzca una concentración tanto de pobreza como de poder en manos de aquellos que controlen sistemas basados en la inteligencia artificial.

- *Cambios sociales y culturales*: Las interacciones sociales e incluso la cultura también son sectores en donde la penetración de la inteligencia artificial va a penetrar en mayor o menor medida. Sin duda alguna, su impacto va a poder ser observado en cuestiones como el ámbito laboral, el ocio y el entretenimiento así como la forma en la que los seres humanos nos relacionamos con los demás.

- *Formación y educación*: Hablamos de cómo la penetración de la inteligencia artificial va a modificar la forma en la que se adquieren habilidades y conocimientos así como cuestiones relativas a la accesibilidad y equidad en materia formativa y/o educativa.

- *Ética, moral y derechos humanos*: La inteligencia artificial es una tecnología que impacta de lleno, debido a cuestiones como la toma de decisiones y la responsabilidad, en dilemas de carácter moral, ético, así como en una cuestión esencial: los derechos humanos.

Es justo en este último punto donde vamos a focalizar el objeto de discurso del presente trabajo. Vamos a tratar de analizar las cuestiones éticas y morales que implica la penetración de la

inteligencia artificial en la sociedad. Asimismo, trataremos de analizar el impacto de esta tecnología en los derechos humanos.

A medida que la sociedad empieza a convivir con los sistemas de inteligencia artificial, el ser humano se empieza a hacer preguntas e incluso a cuestionar el avance de la presente tecnología. ¿Debemos limitar el desarrollo de la inteligencia artificial? ¿Las máquinas harán todo el trabajo por nosotros? ¿Hasta qué punto debemos dejar que las inteligencias artificiales tomen decisiones sobre nosotros? La sociedad, los seres humanos, nos hacemos este tipo de preguntas. Y son preguntas totalmente lícitas. Al fin y al cabo “todo el mundo está interesado en la ética. Todos tenemos nuestras propias ideas sobre lo que está bien y lo que está mal y cómo podemos diferenciarlo”⁵.

Si bien la Real Academia de la Lengua Española define la ética como “aquel conjunto de normas morales que rigen la conducta de la persona en cualquier ámbito de la vida” yo prefiero personalmente entender la ética como “aquella dimensión de la filosofía que reflexiona sobre la moralidad”⁶.

Esta conceptualización de la ética me parece mucho más apropiada, en tanto en cuanto, la ética se encarga de estudiar la moral. ¿Y qué es la moral? Podemos decir que la moral es “una dimensión del mundo vital compuesta de valoraciones, actitudes, normas y costumbres que orientan o regulan el obrar humano”⁷.

Como podemos observar, el campo de la ética cobra importancia en un ámbito tecnológico de tanta trascendencia como es la inteligencia artificial. Al fin y al cabo, hemos de tener en cuenta que las personas que diseñan y desarrollan un sistema de inteligencia artificial se encuentran sometidas al ámbito de la moral y, por tanto, en cierto modo, se encuentran sometidas al campo de la ética.

A lo largo de las presentes líneas trataremos de analizar cómo la ética puede ser aplicada al campo de la inteligencia artificial. Para ello procederemos a analizar dos de los grandes textos in-

⁵ D. Robinson, C. Garrat, *Ética: una guía ilustrada*, Editorial Tecnos, Madrid, 2021, p. 11.

⁶ A. Cortina, *Ética sin moral*, Editorial Tecnos, Madrid, 2000, p. 29.

⁷ J. de Zan, *La ética, los derechos y la justicia*, Fundación Konrad Adenauer, Montevideo, 2004, p. 19.

ternacionales en la materia: la Declaración de Montreal para un desarrollo responsable de la inteligencia artificial y la Recomendación de la UNESCO, sobre la ética de la inteligencia artificial.

Sin embargo, antes de embarcarnos en el análisis de los citados textos, debemos comenzar por definir qué es la inteligencia artificial.

2. Definiendo la inteligencia artificial

Debemos comenzar diciendo que no existe una definición *pacífica* sobre qué es la inteligencia artificial. Los propios avances tecnológicos sobre la materia hacen que, de alguna forma, estemos en constante revisión del término o del concepto.

En 2017 el Parlamento Europeo, realiza una aproximación al concepto de inteligencia artificial estableciendo las características que debe tener un “robot inteligente”:

- Capacidad de adquirir autonomía mediante sensores y/o mediante el intercambio de datos con su entorno (interconectividad) y el análisis de dichos datos.
- Capacidad de aprender a través de la experiencia y la interacción.
- Forma del soporte físico del robot.
- Capacidad de adaptar su comportamiento y acciones al entorno⁸.

En 2018, la Declaración de Montreal para un desarrollo responsable de la inteligencia artificial define el concepto como:

Una serie de técnicas que permiten que una máquina imite el aprendizaje humano, en concreto para aprender, hacer predicciones, tomar decisiones y percibir lo que la rodea. En el caso de un sistema informático, la inteligencia artificial se aplica a los datos digitales⁹.

⁸ Parlamento Europeo (2017), Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica (2015/2103(INL)). Recuperado de https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_ES.html#title2.

⁹ Centro de Investigación en Ética (CRE), Universidad de Montreal (2018), Declaración de Montreal para un desarrollo responsable de la inteligencia artificial. Recuperado de https://declarationmontreal-iaresponsable.com/wp-content/uploads/2023/01/ES-UdeM_Decl-IA-Resp_LA-Declaracion_v4.pdf.

En 2020, la Comisión Europea definía la Inteligencia Artificial como: “una combinación de tecnologías que agrupa datos, algoritmos y capacidad informática”¹⁰.

En 2021, la UNESCO, en su Recomendación sobre la ética de la inteligencia artificial ya es consciente de que no es posible proporcionar una única definición del concepto y que incluso la definición debe ser actualizada en función de los avances tecnológicos, considera que la inteligencia artificial puede ser definida como “sistemas capaces de procesar datos e información de una manera que se asemeja a un comportamiento inteligente, y abarca generalmente aspectos de razonamiento, aprendizaje, percepción, predicción, planificación o control”¹¹.

En el mismo año que la UNESCO, la Comisión Europea lanzaba una propuesta de normas armonizadas, comúnmente denominada “Ley de Inteligencia Artificial” que definía el concepto de “Sistema de inteligencia artificial” como:

El software que se desarrolla empleando una o varias de las técnicas y estrategias que figuran en el anexo I¹² y que puede, para un conjunto determinado de objetivos definidos por seres humanos, generar información de salida como contenidos, predicciones, recomendaciones o decisiones que influyan en los entornos con los que interactúa¹³.

¹⁰ Comisión Europea (2020), *Libro blanco sobre la inteligencia artificial – un enfoque europeo orientado a la excelencia y la confianza*. Recuperado de https://commission.europa.eu/document/download/d2ec4039-c5be-423a-81ef-b9e44e79825b_es?filename=commission-white-paper-artificial-intelligence-feb2020_es.pdf.

¹¹ UNESCO (2021), *Recomendación sobre la ética de la inteligencia artificial*. Recuperado de https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa.

¹² Las técnicas y estrategias de dicho anexo son las siguientes:

“Estrategias de aprendizaje automático, incluidos el aprendizaje supervisado, el no supervisado y el realizado por refuerzo, que emplean una amplia variedad de métodos, entre ellos el aprendizaje profundo.

Estrategias basadas en la lógica y el conocimiento, especialmente la representación del conocimiento, la programación (lógica) inductiva, las bases de conocimiento, los motores de inferencia y deducción, los sistemas expertos y de razonamiento (simbólico).

Estrategias estadísticas, estimación bayesiana, métodos de búsqueda y optimización”.

¹³ Comisión Europea (2021), Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en ma-

Recientemente, la presente definición ha sido modificada para alinearla con el trabajo de organizaciones internacionales que trabajan en la materia. Asimismo, se ha encomendado a la Comisión Europea que trabaje en aras a desarrollar directrices sobre cómo debe aplicarse la definición de “sistema de inteligencia artificial”.

En todo caso, el proyecto final establecido por la Unión Europea, el 26 de enero de 2024, define como “sistema de inteligencia artificial”:

Un sistema basado en una máquina diseñado para funcionar con distintos niveles de autonomía y que pueden mostrar adaptabilidad después del despliegue y que, por razones explícitas, u objetivos implícitos, infiere, a partir de los insumos que recibe, como generar resultados tales como predicciones, contenidos, recomendaciones o decisiones que puedan influir físicamente o en entornos virtuales¹⁴.

Definición que finalmente, en el Reglamento 2024/1689 quedó de la siguiente forma:

Un sistema basado en una máquina que está diseñado para funcionar con distintos niveles de autonomía y que puede mostrar capacidad de adaptación tras el despliegue, y que, para objetivos explícitos o implícitos, infiere de la información de entrada que recibe la manera de generar resultados de salida, como predicciones, contenidos, recomendaciones o decisiones, que pueden influir en entornos físicos o virtuales¹⁵.

teria de inteligencia artificial (Ley de inteligencia artificial) y se modifican determinados actos legislativos de la Unión. Recuperado de <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:52021PC0206>.

¹⁴ Comisión Europea (2024), 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. Recuperado de <https://data.consilium.europa.eu/doc/document/ST-5662-2024-INIT/en/pdf>.

¹⁵ Parlamento y Consejo de la Unión Europea (2024), Reglamento (UE) 2024/1689 del Parlamento Europeo y del consejo de 13 de junio de 2024 por el que se establecen normas armonizadas en materia de inteligencia artificial y por el que se modifican los Reglamentos (CE) n.o 300/2008, (UE) n.o 167/2013, (UE) n.o 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE)

Desde el punto de vista doctrinal, también observamos que lograr una definición *pacífica* del concepto inteligencia artificial no es fácil.

Algunos autores consideran que puede definirse como:

la ciencia y la ingeniería para fabricar máquinas inteligentes, especialmente programas informáticos inteligentes. Está relacionada con la tarea similar de usar computadoras para comprender la inteligencia humana, pero la IA no tiene que limitarse a métodos que son biológicamente observables¹⁶.

Otros autores incluso hablan de distintos tipos de definiciones en función del pensamiento y la racionalidad ante la actuación:

- Sistemas que piensan como los humanos.
- Sistemas que actúan como los humanos.
- Sistemas que piensan racionalmente.
- Sistemas que actúan racionalmente¹⁷.

Otros autores prefieren utilizar el concepto *agente*:

Un agente es algo capaz de comportamiento inteligente. Puede ser un robot o un programa de ordenador. Los agentes físicos, como los robots, tienen una interpretación clara. Toman cuerpo como aparatos físicos que interactúan con un entorno físico. La mayoría de las investigaciones de IA, sin embargo, se centran en los agentes *virtuales* o de *software* que existen como modelos que ocupan un entorno virtual dentro de un ordenador¹⁸.

La presente definición creo que es bastante acertada para definir el estado actual de la técnica. Lo que hoy en día podemos denominar como inteligencia artificial no es ni más ni menos que un sistema que tiene la capacidad de comportarse de una forma inteligente y que puede tener corporeidad y, por tanto,

2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828 (Reglamento de Inteligencia Artificial)

¹⁶ J. McCarthy (2007), "What is Artificial Intelligence?", in *Universidad de Stanford*. Recuperado de <http://jmc.stanford.edu/articles/whatisai/whatisai.pdf>.

¹⁷ S. Russel, P. Norvig, *Artificial Intelligence: A Modern Approach*, Pearson Education, Hoboken, 2003.

¹⁸ H. Brighton, H. Selina, *Inteligencia Artificial: Una guía ilustrada*, Editorial Tecnos, Madrid, 2019.

capacidad de interactuar con su entorno físico o bien, pueden ser incorpóreas y, por tanto, ser únicamente un software con comportamiento inteligente.

No obstante, también nos podemos encontrar con definiciones que diferencian no por la cuestión corpórea sino por la consciencia. Así encontramos definiciones como la siguiente:

[...] Podemos denominar inteligencia artificial blanda (agentes artificiales blandos) o robótica a aquel tipo de agentes que, aunque son capaces de tomar decisiones por sí mismas y, además, son capaces de interactuar con su entorno, aún no son capaces de tomar conciencia de sí mismas. [...] Hablar de un agente artificial fuerte, supone hablar de un agente que, en potencia, podría tener todas las habilidades y talentos del ser humano y no sólo eso, sino que también podrían encontrarse en posesión de otros talentos que no se encontrasen en el ser humano. Es decir, en potencia, un agente artificial fuerte, podría llegar a superar incluso al ser humano. [...] ¹⁹.

Quizá sea demasiado prematuro poder hablar de la posibilidad de que un agente artificial adquiera autoconciencia y, por tanto, que una inteligencia artificial sea capaz no sólo de detectar todo lo que sucede a su alrededor, sino que también sea capaz de descubrir aquello que sucede dentro de ella misma. A pesar de ello es una definición que probablemente tengamos que aplicar. Y seguramente no tardemos demasiado en hacerlo.

No obstante, considero que con las presentes definiciones y a pesar de que puedan existir muchas más, el lector puede tomar conciencia de lo difícil que es lograr definir el concepto de inteligencia artificial.

Creo que, en este sentido, podemos tomar como definición de aproximación la siguiente: una inteligencia artificial es un agente con capacidad de comportamiento inteligente, es decir, puede comportarse y razonar como un ser humano, con independencia de si es un agente virtual o un agente con corporeidad.

¹⁹ F.J. Santamaría Ramos, “Agentes Artificiales”, in F.J. Santamaría Ramos (coordinador), *Tecnoretos del Derecho*, Tirant Lo Blanch, Valencia, 2021, pp. 421-440.

3. La Declaración de Montreal: Análisis y Crítica

En el año 2017 la Universidad de Montreal comenzó a desarrollar y redactar una Declaración para el desarrollo responsable de la inteligencia artificial, conocida comúnmente como Declaración de Montreal y que vio la luz finalmente en el año 2018.

Los objetivos esenciales de la presente declaración son los siguientes:

- Crear un marco ético para el desarrollo y la implementación de la inteligencia artificial.
- Guiar la transición digital para que todos puedan beneficiarse de esta revolución tecnológica.
- Abrir un foro de discusión nacional e internacional a fin de lograr conjuntamente un desarrollo equitativo, inclusivo y ecológicamente sostenible de la inteligencia artificial²⁰.

A lo largo de las presentes líneas vamos a tratar de analizar y, en su caso, criticar los principios y valores éticos insertos en la presente declaración para tratar de dilucidar si en el campo de la inteligencia artificial es posible canalizar una serie de valores éticos esenciales que realmente respeten los intereses de la humanidad.

Comencemos, por tanto, enumerando los principios de la presente declaración:

- Principio de bienestar.
- Principio de respeto a la autonomía.
- Principio de protección de la privacidad y la intimidad.
- Principio de solidaridad.
- Principio de participación democrática.
- Principio de equidad.
- Principio de inclusión de la diversidad.
- Principio de prudencia.
- Principio de responsabilidad.
- Principio de desarrollo sostenible.

Pasemos a analizar detalladamente cada uno de estos principios.

²⁰ Universidad de Montreal (2018), *Declaración de Montreal para un desarrollo responsable de la inteligencia artificial*. Recuperado de https://declarationmontreal-iaresponsable.com/wp-content/uploads/2023/01/ES-UdeM_Decl-IA-Resp_LA-Declaration_v4.pdf.

3.1. Principio de bienestar

Sin entrar en su contenido completo, podemos decir que el presente principio se puede resumir en la siguiente frase: “El desarrollo y el uso de los sistemas de inteligencia artificial deben permitir que se cultive el bienestar de todos los seres conscientes”.

Parto de la base, y creo que es necesario dejarlo claro, de que todos y cada uno de los principios establecidos en la presente declaración están pensados y diseñados como mecánicas para que las inteligencias artificiales puedan ser utilizadas siempre en beneficio de la humanidad.

Dicho esto, existen una serie de preguntas o de retos, dentro del presente sector, que considero que es necesario plantearse sino replantearse dado que cuando uno lee los principios, si bien los considera loables, no deja de tener la sensación de que les falta definición, de que les falta contenido para poder ser aplicados de una forma clara, concisa y contundente.

Por ejemplo, en el segundo punto del presente principio se establece que una inteligencia artificial “debe permitir que las personas vayan tras sus deseos, siempre y cuando no causen daño a otros seres conscientes”. Asimismo, su cuarto apartado establece que una inteligencia artificial “no debe generar males, a menos que nos permitan alcanzar un bienestar superior”.

Analicemos detenidamente estos dos puntos. En primer lugar, el concepto de daño es un concepto extremadamente abstracto. ¿Qué es causar daño a otro ser o a otra persona? Sinceramente, no creo que a la hora de construir el presente principio se estuviese pensando el en principio jurídico de no causar daño a otro (*alterum non laedere*) sino que más bien se está pensando en otro tipo de daño, más no estrictamente el de carácter jurídico. Debemos entender, por tanto, que hablamos de daño es su acepción más común o cotidiana.

Por tanto, considero que debemos entender por daño, cualquier cuestión que pueda afectar de forma negativa a un ser consciente. En este sentido, podríamos hablar de cualquier acción física, psicológica, económica o incluso discriminatoria y tanto de forma intencional como accidental.

El contexto, como podemos observar, es muy amplio e incluso puede dar lugar a situaciones contradictorias que puedan

impedir que la propia inteligencia artificial sea incapaz de tomar una decisión. Un ejemplo, podría ser la recomendación, por parte de la inteligencia artificial, de un comportamiento imprudente. Evidentemente un comportamiento imprudente, por ejemplo, conducir por encima del límite permitido, no debiera ser recomendado por una inteligencia artificial. Sin embargo, ¿y si esa recomendación se realiza porque en el vehículo va otra persona que necesita una intervención médica urgente? En este caso, existen dos daños al mismo tiempo, el del pasajero que necesita atención médica y los posibles daños a terceras personas como consecuencia de una conducción imprudente.

Idéntico razonamiento me produce el punto cuarto donde se habla de malestar y bienestar. Malestar es un término demasiado genérico y abstracto. Definirlo no es fácil... podemos decir que es cualquier situación que altera la tranquilidad o el confort de una persona desde un punto de vista negativo. Por tanto, el bienestar también afecta a la tranquilidad o el confort, pero desde un punto de vista positivo.

Mi primera impresión es que una inteligencia artificial, al igual que una persona, tiene que valorar, equilibrar, cuando va a tomar una decisión que va a generar un malestar, a pesar de que con esa decisión se va a alcanzar un bienestar superior que no se podría lograr de otra forma. Dado que estamos en presencia de conceptos demasiado genéricos, se me hace difícil tanto para las personas como para las inteligencias artificiales el calibrar la presente situación. Máxime cuando cada persona es singular. Lo que a mí me genera malestar no tiene porqué generárselo a otra persona. ¿Cómo puede una inteligencia artificial tomar una decisión sin conocer a la persona? No sería capaz de determinar si la situación le va a provocar malestar a esa persona. Incluso, algo muy común en sociedad. Un comentario inapropiado puede generar malestar en otra persona y muchas veces, ni siquiera nos hace saber dicho malestar.

El presente principio me parece poco práctico. Es incluso poco práctico para el propio ser humano. Por supuesto, desde un punto de vista bienintencionado o incluso optimista, pensamos que todas las personas buscamos cultivar el bienestar de otras personas o de todos los seres conscientes. Sin embargo, el ser

humano sigue entrando en guerra y causando daño y dolor e incluso, causamos daño y malestar a otros seres conscientes. Por ejemplo, seguimos matando animales para poder nutrirnos.

En definitiva, no creo que sea un principio que pueda ser aplicado a todos los sistemas de inteligencia artificial. Existen cierto tipo de recomendaciones o de decisiones que, de forma inevitable, van a causar daño o malestar en terceras personas. Al igual que sucede con el ser humano, en muchas ocasiones, la inteligencia artificial deberá asumir la toma de decisiones aún a sabiendas de que se va a causar un cierto daño o malestar. Lo importante será, en este caso, el equilibrio o *beneficio* logrado.

3.2. Principio de respeto a la autonomía

El presente principio establece que las inteligencias artificiales “se deben desarrollar y utilizar respetando la autonomía de las personas, a fin de que tengan un mayor control sobre sus vidas y lo que las rodea”.

Sobre el presente principio me interesa entrar a analizar sus apartados segundo, quinto y sexto.

El segundo apartado es tremendamente categórico:

No se deben desarrollar ni utilizar las inteligencias artificiales para imponer un estilo de vida específico a las personas, ya sea de forma directa o indirecta, mediante la implementación de mecanismos opresivos de vigilancia y evaluación o de incentivos.

Tengo serias dudas sobre la aplicación práctica del presente apartado en el día a día de cualquier sociedad. Puedo lograr entender el sentido e incluso la dirección que se quiere tomar con el presente apartado y, sin embargo, creo que carece de sentido. ¿Al fin y al cabo, no se impone un estilo de vida específico a las personas en cualquier sociedad? Incluso en las sociedades más democráticas y tolerantes existe un estilo de vida perfectamente marcado y tasado. Aquel estilo de vida que marca el Derecho.

La sociedad implica comportamientos: colectivos e individuales y dichos comportamientos deben ser sometidos a unas reglas del juego, a unas normas reguladoras, en definitiva, a

unas mecánicas que nos permitan resolver las posibles controversias que surgen, de forma inevitable, en la sociedad²¹.

Entiendo que este apartado se refiere a la posibilidad de que un régimen autoritario utilice sistemas de inteligencia artificial para oprimir, vigilar o evaluar a sus ciudadanos, pero seamos sinceros, los regímenes autoritarios han existido a lo largo de la historia y no han necesitado de la inteligencia artificial para lograr sus fines.

Además, en el supuesto de que eso pasase... ¿Quién va a hacer frente a esa situación? Baste ver cómo los diversos Estados giran la cabeza ante una guerra o un ataque frontal a un país que no sea el suyo. Sí, surgen sanciones internacionales, discursos de buenas intenciones y un largo etcétera que, luego, a la postre, no acaba con la guerra... Dudo mucho que ningún Estado haga poco más que criticar el uso de la inteligencia artificial con fines autoritarios...

El quinto apartado también tiene un buen sentido y dirección, máxime si nos atenemos a las propias características del siglo XXI: “No se deben desarrollar inteligencias artificiales para difundir información no confiable, mentiras o propaganda; se los debe diseñar con el objetivo de contener su difusión”.

Las noticias falsas o *fakes news* son una de las lacras más importantes del siglo XXI. La llegada de la inteligencia artificial no hace sino agravar la cuestión. La presente cuestión es tan importante que la Unión Europea presentó el 16 de junio de 2022 un Código de buenas prácticas en materia de desinformación²². El presente Código suscribe cuestiones relevantes como la desmonetización de la desinformación, la transparencia de la publicidad política, el empoderamiento de los usuarios, las mecánicas de cooperación con los verificadores de datos, etc.

Sin embargo, considero que este apartado también carece de una utilidad práctica. Estoy de acuerdo y comparto el peligro que supone la desinformación en la sociedad globalizada del siglo XXI. Sin embargo, y aunque busquemos contextos de autorre-

²¹ F.J. Santamaría Ramos, *Introducción al sistema jurídico español*, Tirant Lo Blanch, Valencia, 2022.

²² Unión Europea (2022), *Código de buenas prácticas en materia de desinformación*. Recuperado de <https://digital-strategy.ec.europa.eu/en/library/2022-strengthened-code-practice-disinformation>.

gulación como son los códigos de buenas prácticas, finalmente necesitamos recurrir al Derecho. El Reglamento Europeo de Servicios Digitales busca lograr un entorno en línea seguro y “aborda la difusión de contenidos ilícitos en línea y los riesgos para la sociedad que puede generar la difusión de desinformación y otros contenidos”²³. Como podemos observar, si desde el punto de vista social no logramos evitar las mentiras y la desinformación que generan las propias personas, difícilmente unos principios éticos ligados a la inteligencia artificial lo van a lograr.

Nos queda por analizar el sexto apartado del presente principio que establece que en el “desarrollo de la inteligencia artificial, se debe evitar la creación de dependencias mediante técnicas que capten la atención o imiten las características humanas (apariciencia, voz, etc.) que pueden causar confusión entre las inteligencias artificiales y los humanos”.

El apartado es interesante y se logra ver la finalidad de este. La sociedad tiene miedo a una pérdida de la socialización de las personas entre sí. Al fin y al cabo si una inteligencia artificial no sólo piensa y razona como el ser humano sino que además tiene rasgos similares o idénticos a los del propio ser humano, se puede caer en un alto grado de aislamiento social. Es decir, que las personas prefiramos interactuar con inteligencias artificiales antes que con otros seres humanos.

No obstante, fíjese el lector en que si bien, en otros apartados analizados, la Declaración es tremendamente tajante y categórica, no lo es en este caso. De un “no se deben crear dependencias” hemos pasado a un “se debe evitar”. Considero que no es un descuido, sino que es algo que incluso en una Declaración de estas características es evidente... En un futuro, no muy lejano, los desarrolladores dotarán de apariciencia, voz y etc. a las inteligencias artificiales para ser mucho más *amigables* para el ser humano.

²³ Parlamento y Consejo Europeo (2022), Reglamento (UE) 2022/2065 del Parlamento Europeo y del Consejo de 19 de octubre de 2022 relativo a un mercado único de servicios digitales y por el que se modifica la Directiva 2000/31/CE (Reglamento de Servicios Digitales). Recuperado de <https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:32022R2065#d1e2102-1-1>.

3.3. Principio de protección de la privacidad y la intimidad

El presente principio establece la necesidad de proteger la intimidad contra cualquier intromisión de la inteligencia artificial, así como de los sistemas de adquisición de archivos de datos²⁴.

Sin duda alguna, el presente principio me parece de lo más oportuno y quizá, uno de los principios que, desde el punto de vista jurídico, sea de los más fáciles de implementar dado que aquí estamos tratando un valor social que ya ha tenido una regulación importante a nivel mundial.

No obstante, me gustaría analizar una serie de cuestiones desarrolladas en el presente principio que pueden ser, en cierto modo, debatibles.

El presente principio se desarrolla a través de 8 apartados de los que particularmente me interesa entrar en dos de ellos.

En el primer apartado se establece la protección de la intrusión de la inteligencia artificial así como de los sistemas de adquisición de archivos de datos de aquellos “espacios personales” en los que las personas no se encuentran bajo vigilancia e evaluación digital. Me surgen varias inquietudes al respecto.

La primera inquietud que me surge es qué debemos entender por “*espacio personal*”. Y digo esto, porque considero que es un concepto que podemos denominar generacional. Es decir, la percepción de dicho espacio va a depender en gran medida, de la época en la que las personas hallamos nacido y madurado.

Bajo mi punto de vista, quizá *clásico*, en gran medida porque nací y me crié en el siglo XX, no tengo más remedio que considerar que ese “espacio personal” no es ni más ni menos que la propia vida privada en su conjunto, es decir, aquel lugar reservado sólo “a unos pocos, familiares y amigos”²⁵.

No obstante, parece que este no es el sentir de los redactores de la Declaración de Montreal ya que en su texto dan a entender que es posible diferenciar distintos “espacios personales” y que,

²⁴ Un sistema de adquisición y archivo de datos se puede definir como “todo sistema informático que pueda recopilar y registrar datos. En última instancia, estos datos se utilizan para entrenar los sistemas de IA como parámetros para la toma de decisiones”.

²⁵ J.A. Soler Martínez, “Protección constitucional de la intimidad y de los datos de carácter personal frente a las nuevas tecnologías”, in *Anuario de Derecho Canónico*, 11 (2022), pp. 93-126.

por tanto, los que deben protegerse son los que no se encuentran bajo vigilancia o evaluación digital. Y yo me pregunto, ¿existe algún espacio donde el ser humano ya no se encuentre sometido a vigilancia o evaluación? Sinceramente, considero que no. Alexa, Cortana o Siri ya son asistentes artificiales que están las 24 horas del día “pendientes de nuestros movimientos”²⁶ y que ya son “uno más de la familia, uno más del hogar familiar”.

Nótese que incluso la Ley sobre Inteligencia Artificial de la Unión Europea considera que existen una serie de sistemas de inteligencia artificial que “son componentes de seguridad de productos, o que son en sí mismos productos”²⁷, que son clasificados como de “*alto riesgo*”. Alguno de estos productos son los equipos a presión, aparatos de gas, etc. Sorprende sobremedida que dentro de esta clasificación de alto riesgo se encuentren también los juguetes. O quizá no sorprende tanto, dado que es posible, que el legislador europeo ya esté previendo las posibles vulneraciones sobre la intimidad que pueden tener los sistemas de inteligencia artificial aplicados a los juguetes. Sin duda, me parecen el producto perfecto, para *colarse* en la intimidad de todos los hogares.

Otro de los puntos que me interesa analizar es el segundo punto que establece la protección de los “pensamientos y las emociones contra aquellos usos que puedan ocasionar daños, en especial aquellos usos que imponen juicios morales sobre las personas o sus elecciones de vida”.

Comparto plenamente este posicionamiento. Existe un peligro potencial a que los sistemas de inteligencia artificial puedan ser utilizados como *agentes de la moral*. No obstante, existe otro riesgo, que, bajo mi punto de vista, es aún más peligroso. Y es

²⁶ Diario “El País”, “Empleados de Amazon escuchan a diario conversaciones que mantienen los usuarios con Alexa”. Recuperado de https://elpais.com/tecnologia/2019/04/11/actualidad/1554992401_521050.html.

²⁷ Parlamento y Consejo de la Unión Europea (2024), Reglamento (UE) 2024/1689 del Parlamento Europeo y del consejo de 13 de junio de 2024 por el que se establecen normas armonizadas en materia de inteligencia artificial y por el que se modifican los Reglamentos (CE) n.o 300/2008, (UE) n.o 167/2013, (UE) n.o 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828 (Reglamento de Inteligencia Artificial)

convertirnos en estados sometidos a la “*dictadura tecnológica*”. Soy consciente de la dureza de mis palabras pero me sorprende, por ejemplo, que en Estados democráticos, como son los Estados que conforman la Unión Europea nos planteemos cuestiones como las que aparecen en el apartado tercero del artículo 50 de la Ley de Inteligencia Artificial donde se deja entrever que va a ser posible utilizar sistemas de inteligencia artificial enfocados a la “categorización biométrica y el reconocimiento de emociones que hayan sido autorizados por ley para detectar, prevenir e investigar delitos”, eso sí, haciendo referencia a que dichos sistemas se sujetarán garantías adecuadas para los derechos y libertades de terceros y de conformidad con el Derecho de la Unión.. Pero fíjense en el matiz, no sólo se habla de la investigación propiamente dicha, hablamos de la detección y prevención, dos conceptos que implican anticipación, es decir, adelantarse a la propia comisión del delito.

Me asaltan las dudas... ¿Dónde queda el aforismo romano “*Cogitationis poenam nemo patitur*”? ¿Acaso una conquista propia de la revolución francesa²⁸ como el establecimiento de que el pensamiento no delinque ya no tienen vigor en la era del conocimiento? Considero que cometemos un grave error como sociedad si utilizamos los sistemas de inteligencia artificial no sólo como *agentes de la moralidad* sino también como sistemas que indaguen en nuestro pensamiento y en nuestras emociones para la prevención de delitos. Una persona determinada, en un momento concreto puede tener unos sentimientos y emociones muy intensos como pueden ser el odio y la ira y eso no significa que vayan a cometer un delito. Prejuizar de esta forma a las personas no es propio de ninguna sociedad democrática del siglo XXI.

3.4. Principio de solidaridad

El presente principio establece que “el desarrollo de los sistemas de inteligencia artificial debe intentar preservar los lazos de solidaridad entre las personas y las generaciones”.

²⁸ C.A. Gómez Pavajeau, “Injusto y culpabilidad en los grandes filósofos griegos”, in *Revista de Derecho Penal y Criminología*, 31 (2010), n. 91, pp. 137-151.

Sobre el presente principio únicamente me interesa entrar a analizar el apartado tercero que establece que “no se deben implementar los sistemas de inteligencia artificial para remplazar a las personas en tareas que requieren relaciones humanas de calidad, sino que se los debe desarrollar para facilitar dichas relaciones”.

Lo primero que me llama la atención es que hablar de relaciones humanas de calidad es un concepto bastante genérico. Quiero entender que se están refiriendo a un tipo específico de interacción que se da entre las personas basado en la empatía, la comunicación efectiva, el respeto mutuo y la confianza y que derivan, por tanto, en un estado saludable y satisfactorio para aquellos que las practican.

Entiendo el sentido y la dirección que implica la presente cuestión. Se trata de minimizar el impacto que los sistemas de inteligencia artificial puedan tener en un futuro a medio plazo, el remplazo de la socialización entre las personas por la socialización entre personas y agentes artificiales. Y no es descabellado pensar que esto pueda lograr. Si es posible que dos personas se enamoren simplemente chateando, no es descabellado pensar que las personas podamos tener afinidades con agentes artificiales y decidamos cultivar esas relaciones antes que cultivar relaciones con otras personas.

No obstante, hemos de pensar que, en algún momento, vamos a tener que implementar sistemas de inteligencia artificial que desarrollen relaciones de calidad, sobre todo en Estados donde la natalidad es realmente baja y no va a ser sencillo encontrar personas para desarrollar determinadas tareas como pueden ser el cuidado de personas mayores. Soy consciente de que mis palabras van contra el antropocentrismo propio de nuestra especie pero, también soy una persona realista y considero que a medio plazo, no existirán tantas facilidades para cultivar las relaciones de calidad en determinados contextos socio-laborales.

3.5. Principio de participación democrática

El presente principio establece que “los sistemas de inteligencia artificial deben cumplir con los criterios de inteligibilidad,

capacidad de justificación y accesibilidad, y se los debe someter a un análisis, un debate y control democráticos”.

Sobre este punto poco tengo que objetar, simplemente realizar dos matices de cuestiones que considero que deben ser siempre así, con independencia de encontrarnos o no en un contexto de *participación democrática*.

Los apartados 7 y 9 del presente principio estableces que “en todo momento debemos ser capaces de verificar que los sistemas de inteligencia artificial estén haciendo aquello para lo que se los programó y para lo que se los utiliza” y que, “todo usuario de un servicio que utiliza un bot de chat²⁹ debe poder identificar con facilidad si está interactuando con un sistema de inteligencia artificial o una persona real”.

Me parecen dos cuestiones de lo más razonables. Considero que, al menos con el estado actual de la técnica, es necesario que el ser humano sea capaz de constatar que los agentes artificiales realmente están realizando las tareas para las cuales se han creado. No me refiero a analizar todos y cada uno de sus procesos pero sí, ser conscientes de que se están dedicando a cumplir las finalidades para las cuales queremos utilizarlas y que, por tanto, no están desarrollando tareas secundarias o alternativas o, si lo están haciendo, que sepamos el porqué de dicha acción.

Asimismo, me parece totalmente pertinente que los seres humanos sepamos identificar si nos encontramos en presencia de otro ser humano o si, por el contrario, nuestra interacción se realiza con un agente artificial.

3.6. Principio de equidad

En mi opinión estamos en presencia de uno de los principios de la Declaración más subjetivos y, sobre todo, más ambiciosos. De una forma breve, sus siete apartados se resumen en establecer que “el desarrollo y uso de los sistemas de inteligencia artificial deben contribuir a crear una sociedad justa y equitativa”.

Creo que el presente supuesto es demasiado ambicioso y parte de un razonamiento que, a mi juicio, es erróneo. Parte de la base de

²⁹ Hablamos de “un sistema de inteligencia artificial que pueden conversar con una persona en lenguaje natural”.

que la sociedad es local, pero, la sociedad del siglo XXI es global. Voy a tratar de explicarme lo mejor posible, dado que este principio necesita de un desarrollo que no puede abarcar el presente estudio. Y aunque mi análisis será breve no carecerá de rigor.

Todo el presente principio pivota en esperanzas muy loables:

- Evitar la creación, el refuerzo o la reproducción de patrones de discriminación (social, sexual, étnica, cultural, religiosa ...).
- Tratar de eliminar las relaciones de dominación entre grupos y personas en función de diferencias de poder, riqueza o conocimiento.
- La reducción de las desigualdades y las vulnerabilidades.
- Un acceso igualitario a los recursos, conocimientos así como a las herramientas digitales esenciales.
- La equidad social.

Sin embargo, “el mundo no es igualitario ni puede serlo”³⁰. En este sentido coincido plenamente con la siguiente afirmación:

La equidad social suele referirse a diferentes ámbitos: justicia social y cohesión; de igualdad de oportunidades, derechos, libertades, de trato y de logro/resultados; recibir prestaciones básicas del estado y una distribución de la renta/riqueza “razonable”; evitar situaciones de pobreza o discriminaciones (por etnia, lugar de nacimiento o residencia, capacidad, género, edad, etc.). Es por tanto, un fenómeno histórico y cultural que se da en todas las sociedades, si bien el grado, ámbito (o bien valorado), el sentido... para establecer la desigualdad variará según el país, la cultura, el sistema de creencia o las condiciones del entorno, entre otras³¹.

Esta es la cuestión importante, el punto nuclear, el concepto de equidad social es un concepto que se da en todas las socieda-

³⁰ C. Pacella, “Aparatos estéticos”, in *Revista del Centro de Investigaciones de la Facultad de Filosofía y Humanidades, Áreas Letras*, 11 (2020), n. 18, p. 8.

³¹ R. Calvo Palomares, J.J. Portet Tiebas, M. Bou Pastor, “La equidad social como elemento esencial para el desarrollo local”, in J.A. Márquez Domínguez, J.M. Jurado Almonte, F.J. Pazos García (coordinadores), *Desarrollo local en territorios de fronteras*, Servicio de publicaciones de la Universidad de Málaga, Málaga, 2016.

des pero que varía según el país, es decir, es un fenómeno *local*. Los agentes artificiales tienen una dimensión de carácter global. Una persona en España puede utilizar un sistema de inteligencia artificial desarrollado en China. ¿Qué el ser humano trata de evitar las discriminaciones? Podemos decir que, a nivel general, es uno de los objetivos de la humanidad. ¿Realmente es una realidad? Todos sabemos que no lo es.

Algunos ejemplos:

(i) Según las Naciones Unidas, en 2023 casi 700 millones de personas en todo el mundo subsistirán con menos de 2,15 dólares. Es más, se prevé que no se alcanzará el objetivo mundial de acabar con la pobreza extrema para 2030, ya que casi 600 millones de personas siguen viviendo en la pobreza extrema³².

(ii) Amnistía Internacional establece que, durante el año 2022, los conflictos globales provocaron graves violaciones de los derechos humanos y del derecho internacional humanitario, con una respuesta internacional insuficiente. La violencia de género persistió a pesar de algunos avances legislativos, y el derecho al aborto experimentó altibajos. La pandemia de COVID-19 exacerbó las crisis económicas, alimentarias y climáticas, afectando desproporcionadamente a los grupos marginados y aumentando la desigualdad. La represión de la disidencia política alimentó los conflictos armados y los golpes de Estado, mientras que la violencia sexual en estos contextos se mantuvo como una cruel realidad³³.

(iii) Naciones Unidas, en 2022 llegó a establecer que,

En los Estados Unidos de América se habían aprobado o estaban en curso de examen leyes que limitaban los debates en las escuelas sobre la raza y otras características personales,

³² Naciones Unidas (2023), “Desafíos globales: Acabar con la pobreza”. Recuperado de <https://www.un.org/es/global-issues/ending-poverty#:~:text=Seg%C3%BAAn%20estimaciones%20recientes%2C%20en%202023,menos%20de%20%2C15%20d%C3%B3lares.&text=Se%20prev%C3%A9%20que%20no%20se,viviendo%20en%20la%20pobreza%20extrema>.

³³ Amnistía Internacional, *Informe 2022/2023 Amnistía Internacional: La situación de los derechos humanos en el mundo*, Amnesty Internacional Ltd, Reino Unido, 2023.

y también iniciativas para prohibir libros en las bibliotecas escolares y públicas, en especial los que trataban cuestiones relativas a la raza, así como a otras identidades. En algunos estados se habían aprobado leyes y medidas restrictivas del voto, que podían afectar de forma desproporcionada a los afrodescendientes³⁴.

A la humanidad le queda por recorrer un largo camino en cuanto a la equidad social se refiere y, sin embargo, el temor a los sistemas de inteligencia artificial es tal, que buscamos que un producto creado por el ser humano sea un espejo en el que desearíamos vernos reflejados. La realidad, en mi opinión, es que al igual que el ser humano no es equitativo socialmente, sus productos, las inteligencias artificiales, tampoco lo serán. Por mucho empeño que pongamos en ello.

3.7. Principio de inclusión de la diversidad

El presente principio se puede resumir estableciendo que “el desarrollo y el uso de los sistemas de inteligencia artificial deben contribuir a preservar la diversidad social y cultural y no debe restringir las posibilidades en lo que respecta a las elecciones de vida o las experiencias personales”.

Realmente, al igual que me sucede con el principio anteriormente expuesto, considero que el presente principio necesita de un desarrollo más exhaustivo del que se le puede dar en el presente trabajo. No obstante, trataré de centrarme en alguna de sus cuestiones más importantes.

En su primer apartado establece que los sistemas de inteligencia artificial “no deben conducir a una homogeneización de la sociedad mediante la normalización de los comportamientos y opiniones”.

Sin embargo, la homogeneización ya es una realidad, con independencia de la penetración en la sociedad de los agentes artificiales. De hecho:

³⁴ Asamblea General de Naciones Unidas, *Promoción y protección de los derechos humanos y las libertades fundamentales de los africanos y los afrodescendientes frente al uso excesivo de la fuerza y otras violaciones de los derechos humanos por agentes del orden, mediante un cambio transformador en favor de la justicia y la igualdad raciales*, Nueva York, 2022.

La estandarización cultural, fruto de la modernización y la creciente homogeneización de la sociedad, la mundialización de las problemáticas, la aldea global, no ha conseguido que desaparezcan los desequilibrios geopolíticos y culturales, que se reflejan en la estructura jerarquizada de las culturas y en las nuevas formas de estratificación social, en consecuencia, nuevas formas de exclusión social, educativa y cultural³⁵.

No puedo sino suscribir estas palabras, ya estamos asistiendo a un fenómeno complejo y de consecuencias ciertamente imprevisibles para la sociedad, nuestra homogeneización cultural.

Esta cuestión, también se hace patente en el apartado cuarto del presente principio cuando se habla de la posibilidad de que los agentes artificiales puedan “encerrar a las personas en un perfil de usuario o confinarlas en un filtro burbuja”³⁶.

Es decir, determinados servicios de internet pueden, en cierta forma limitar o censurar nuestra libertad de información. O mejor dicho, nuestras propias experiencias como usuarios pueden hacer que nos encerremos en nosotros mismos y no seamos capaces de localizar otros puntos de vista. Sin duda, como muy bien se dice en la Declaración, esto “limita las posibilidades de desarrollo personal”.

Por último, en su último apartado, se establece que “se debe diversificar la oferta de los sistemas de inteligencia artificial para evitar que se formen monopolios de facto y se socaven las libertades individuales”.

Me sorprende sobremanera el presente principio cuando desde hace varios años, antes incluso de la llegada de la inteligencia artificial a nuestras vidas, éramos conscientes de que: “Una característica de las plataformas digitales que más éxito tienen es la rapidez con que consiguen una cuota de mercado considera-

³⁵ B. Mulet Trobat, “Apuntes para la interpretación sociológica de las consecuencias socioculturales y educativas de la Cibercultura en la era de la globalización”, in *Estudios sobre las culturas contemporáneas*, 28 (2023), n. 56, pp.169-206.

³⁶ Hablamos de “una información filtrada que llega a una persona en internet. Diferentes servicios, ofrecen resultados personalizados para los usuarios. Esto puede aislar a las personas porque ya no pueden acceder a la información común”.

ble, dando lugar a que dichas plataformas tiendan a convertirse en monopolios”³⁷.

En este caso, suscribo totalmente la presente afirmación, aunque yo diría aún más. Debemos empezar a ser conscientes de que el gran monopolio no es ni más ni menos que la propia sociedad digital.

3.8. Principio de prudencia

De una forma resumida el presente principio establece que “toda persona que esté involucrada en el desarrollo de la inteligencia artificial debe actuar con prudencia y anticipar, en la medida de lo posible, los efectos adversos del uso de los sistemas de inteligencia artificial y tomar las medidas adecuadas para evitarlos.

Sin duda alguna, este principio me parece de lo más útil, oportuno y, además, coherente. Considero que es un principio que sí debe ser esencial en el desarrollo de los agentes artificiales. Sobre todo, en dos cuestiones importantes que interesa reseñar.

Por un lado, su primer apartado establece que

es necesario desarrollar mecanismos que tengan en cuenta las posibilidades de uso doble (beneficioso o perjudicial) de la investigación en inteligencia artificial y el desarrollo de los sistemas de inteligencia artificial (ya sean públicos o privados) a fin de limitar los usos perjudiciales.

No puedo estar más de acuerdo, con independencia de que el desarrollo sea público o privado debemos desarrollar protocolos que nos permitan observar aquello que quizá no queramos o no sepamos ver. Cuando se desarrolla una tecnología, sin duda alguna, siempre se ve el beneficio, pero como regla general, nunca observamos el peligro que su uso puede comportar. Y sino, que se lo pregunten a Alfred Nobel que inventó la dinamita para agilizar los trabajos en la construcción... Quizá si hubiese visto el peligro potencial de su invención, nunca se hubiese animado a hacerla pública...

³⁷ UNCTAD, *Informe sobre la economía digital 2019*, Servicio de publicaciones de las Naciones Unidas, New York, 2019.

Por otro lado, otra cuestión que me parece de lo más oportuno y coherente reside en el tercer punto de este principio:

Antes de lanzarlos al mercado, sin importar si son gratuitos o pagos, los sistemas de inteligencia artificial deben cumplir requisitos rigurosos de fiabilidad, seguridad e integridad, y se los debe someter a pruebas que no pongan en peligro la vida de las personas, perjudiquen su calidad de vida o afecten de forma negativa su reputación o integridad psicológica.

No obstante, y a pesar de que me parece algo muy oportuno, considero que la redacción podría mejorarse. Yo no hablaría de su lanzamiento al mercado, sino que directamente establecería el lanzamiento a la propia sociedad. Hablar de mercado, bajo mi punto de vista, podría dar a entender que sólo hablamos de aquellos agentes artificiales elaborados por el sector privado, dejando fuera de la ecuación a los agentes artificiales elaborados por el sector público... Creo que eso es un error. Cuando una entidad pública proceda a lanzar a la sociedad un agente artificial debe garantizar todas las cuestiones que acabamos de analizar: fiabilidad, integridad, seguridad, que no pongan en peligro a las personas o perjudiquen su vida.

3.9. Principio de responsabilidad

El presente principio establece, de una forma resumida que “el desarrollo y el uso de los sistemas de inteligencia artificial no deben contribuir a reducir la responsabilidad de los seres humanos a la hora de tomar decisiones”.

La idea clave que reside sobre el presente principio es que única y exclusivamente el ser humano es el responsable de las recomendaciones o acciones de un agente artificial.

Si bien entiendo el presente principio dado que, de lo que se trata es de que no existan lagunas desde el punto de vista de la responsabilidad, ya sea de carácter civil, penal, administrativa, etc.

De hecho, existen en este principio una serie de *líneas rojas*³⁸, fruto del potencial que tienen las inteligencias artificiales. Un agente artificial, por sí sólo nunca puede tomar decisiones que:

³⁸ Es decir, la decisión final debe ser tomada siempre por el ser humano.

(i) Tengan que ver con la decisión de matar.

(ii) Impliquen la comisión de un crimen o delito. Es decir, sí que podrán tomar este tipo de decisiones, pero la responsabilidad de dichos actos va a recaer directamente en la persona o personas que autorizan a que el agente artificial lo cometa.

Aun así, el presente principio también establece que “cuando un sistema de inteligencia artificial provoca daños o perjuicios y se demuestra que dicho agente es responsable pero se utilizó según lo previsto, no es razonable culpar a las personas involucradas en su desarrollo o uso”. Es decir, asistimos a una exoneración de responsabilidad que, a mi juicio, es de lo más coherente.

3.10. Principio de desarrollo sostenible

El presente principio establece que “el desarrollo y el uso de los sistemas de inteligencia artificial se deben llevar a cabo con el objetivo de garantizar una sostenibilidad fuerte del medio ambiente”

Quizá uno de los temas menos comentados cuando se habla de inteligencia artificial. Su desarrollo puede implicar un impacto ambiental importante. De hecho, su impacto ambiental “deriva principalmente de la cantidad ingente de energía requerida por los sistemas de IA para su eficaz desarrollo, entrenamiento y uso”³⁹.

En algún momento, la presente cuestión deberá ser debatida en foros internacionales ya que se deberán tomar decisiones en tanto en cuanto, a mayor desarrollo de los agentes artificiales, mayo impacto ambiental sufrirá nuestro planeta.

4. UNESCO: Recomendación sobre la ética de la inteligencia artificial

4.1. Introducción

A finales de 2021, la Conferencia General de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultu-

³⁹ D.E. Araiz Huarte, “La inteligencia artificial como agente contaminante: Concepto jurídico, impacto ambiental y futura regulación”, in *Actualidad Jurídica Ambiental*, (2023), n. 130.

ra (UNESCO) en su 41ª reunión adoptó la Recomendación sobre la ética de la inteligencia artificial.

Dentro de su preámbulo, la UNESCO reconoce que “el uso de los agentes artificiales puede llegar a influir en el pensamiento, las interacciones, así como en la adopción de decisiones, por parte de los seres humanos”.

Asimismo, la UNESCO, es consciente de la preocupación que existe, a nivel mundial, en relación con las preocupaciones éticas, en particular con cuestiones relativas a la “discriminación, desigualdad, brechas digitales, exclusión social, entre otras”. Por supuesto, la presente preocupación lleva asociado la consideración del impacto que estas tecnologías pueden tener sobre “la dignidad humana, los derechos humanos, las libertades fundamentales, la igualdad de género, la democracia, los procesos sociales, económicos, políticos y culturales, las prácticas científicas de ingeniería, el bienestar animal y el medio ambiente y los ecosistemas”.

Es por ello que la presente recomendación pretender ser una brújula que los Estados Miembros de la UNESCO, pueden aplicar, de forma voluntaria, a través de medidas, fundamentalmente, de carácter legislativo.

Tal y como se establece en el ámbito de aplicación de la presente Recomendación:

Aborda la ética de la inteligencia artificial como una reflexión normativa sistemática, basada en un marco integral, global, multicultural y evolutivo de valores, principios y acciones interdependientes, que puede guiar a las sociedades a la hora de afrontar de manera responsable los efectos conocidos y desconocidos de las tecnologías de la inteligencia artificial en los seres humanos, las sociedades y el medio ambiente y los ecosistemas, y les ofrece una base para aceptar o rechazar las tecnologías de la inteligencia artificial. Considera la ética como una base dinámica para la evaluación y la orientación normativas de las tecnologías de la inteligencia artificial, tomando como referencia la dignidad humana, el bienestar y la prevención de daños apoyándose en la ética de la ciencia y la tecnología.

La Recomendación de la UNESCO se basa en una serie de principios y aunque la propia organización reconoce que todos

y cada uno de ellos son deseables y pertinentes, existe la posibilidad de que, en algunas ocasiones, los mismos puedan entrar en contradicción y, por tanto, los Estados Miembros deberán ser capaces de gestionar dichas tensiones basándose en el principio de proporcionalidad, así como en el respeto a los derechos humanos y a las libertades fundamentales. No obstante, y en todo caso “las tecnologías de la inteligencia artificial deben estar al servicio del bien de la humanidad”. La UNESCO sin duda, aglutina en sus palabras los deseos de la humanidad. No obstante, me pregunto si todos y cada uno de sus Estados Miembros desean y quieren hacer realidad dichos deseos, máxime cuando países como Sudán, Rusia y Ucrania, Países Miembros, no lo olvidemos, han cogido las armas para dirimir sus disputas.

En todo caso, el texto ofrece una serie de valores y principios que merece la pena enumerar. Los valores de la presente Recomendación son los siguientes:

- Respeto, protección y promoción de los derechos humanos, las libertades fundamentales y la dignidad humana.
- Prosperidad del medio ambiente y los ecosistemas.
- Garantía de la diversidad y la inclusión.
- Vida en sociedades pacíficas, justas e interconectadas.

En relación con los principios, son los siguientes:

- Proporcionalidad e inocuidad.
- Seguridad y protección.
- Equidad y no discriminación.
- Sostenibilidad.
- Derecho de la intimidad y protección de datos.
- Supervisión y decisión humanas.
- Transparencia y explicabilidad.
- Responsabilidad y rendición de cuentas.
- Sensibilización y educación.
- Gobernanza y colaboración adaptativas y de múltiples partes interesadas.

No obstante, quizá la parte más interesante de la presente Recomendación son los ámbitos de actuación que plantea. Vamos a

proceder a su enumeración para, de forma posterior, analizarlos detenidamente:

- Evaluación del impacto ético.
- Gobernanza y administración éticas.
- Política de datos.
- Desarrollo y cooperación internacional.
- Medio ambiente y ecosistemas.
- Género.
- Cultura.
- Educación e investigación.
- Comunicación e información.
- Economía y trabajo.
- Salud y bienestar social.

4.2. Evaluación del impacto ético

El presente ámbito de actuación me parece de sumo interés. En él se plantean varias cuestiones que considero que deben ser necesarias en cualquier regulación sobre inteligencia artificial:

- Evaluaciones de impacto ético.
- Mecanismos de diligencia y supervisión.
- Investigación de los efectos sociológicos y psicológicos de las recomendaciones basadas en la inteligencia artificial sobre los seres humanos en lo relativo a su autonomía de decisión.

Considero totalmente necesarios los presentes mecanismos. Ante una tecnología como la presente es recomendable detenerse a analizar los beneficios, así como los perjuicios y riesgos de los agentes artificiales. Y dicho análisis debe ser realizado de forma previa a su puesta en funcionamiento.

El análisis no sólo debe ser previo sino también durante el funcionamiento de las presentes tecnologías. Debemos implementar mecánicas capaces de “determinar, prevenir y atenuar los riesgos”, en caso de que éstos se materialicen. Incluso, como plantea la Recomendación los agentes artificiales deben ser probados, incluso en escenarios reales, antes de su puesta en funcionamiento. Bajo mi punto de vista, esta es la única opción de poder saber, con seguridad, que los agentes artificiales no suponen ningún riesgo para el ser humano.

En este sentido, la Ley de Inteligencia Artificial, elaborado por la Unión Europea⁴⁰, contempla la presente cuestión es su artículo 27 que lleva por título “Evaluación del impacto sobre los derechos fundamentales de los sistemas de inteligencia artificial de alto riesgo”. En este sentido, se alaba la intención del legislador europeo, aunque también es criticable que la presente evaluación únicamente se plantea para los sistemas de inteligencia artificial catalogados por la Unión Europea como de alto riesgo. No obstante, la Unión Europea es pionera en este sentido y cualquier norma, tanto nacional como internacional debiera incluir regulación en el presente sentido.

4.3. Gobernanza y administración éticas

Considero que el presente ámbito de aplicación tiene una derivada jurídica de interés, en tanto en cuanto, es un ámbito que se puede materializar, siempre y cuando haya interés y voluntad política.

A grandes rasgos, el presente ámbito se centra en las siguientes cuestiones:

- Velar por que los mecanismos de gobernanza de la inteligencia artificial sean inclusivos, transparentes, multidisciplinares y multilaterales.

- Velar por que se investiguen y reparen los daños causados mediante sistemas de inteligencia artificial, estableciendo mecanismos de aplicación estrictos y medidas correctivas.

- Elaborar estrategias, tanto de carácter nacional como regional sobre inteligencia artificial.

- Elaborar mecanismos de certificación, así como sistemas de reconocimiento mutuo de dichas certificaciones.

⁴⁰ Parlamento y Consejo de la Unión Europea (2024), Reglamento (UE) 2024/1689 del Parlamento Europeo y del consejo de 13 de junio de 2024 por el que se establecen normas armonizadas en materia de inteligencia artificial y por el que se modifican los Reglamentos (CE) n.o 300/2008, (UE) n.o 167/2013, (UE) n.o 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828 (Reglamento de Inteligencia Artificial).

- Realizar autoevaluaciones transparentes de los sistemas de inteligencia artificial, tanto de los ya existentes, como de aquellos que se puedan ir proponiendo.

- Diseñar e implementar una figura de responsable independiente de la ética de la inteligencia artificial, así como crear una red de responsables independientes en los planos nacional, regional e internacional.

- Elaborar legislaciones nacionales respetuosas con los derechos humanos y las libertades fundamentales.

- Vigilar el impacto social y económico que puede tener el uso de inteligencia artificial en ámbitos de carácter sensible tales como los derechos humanos, la aplicación de la ley, la asistencia social, el empleo o la salud, por citar algunos ejemplos.

- Elaborar normas internacionales que describan los niveles de seguridad y transparencia mensurables y comprobables, de modo que se puedan evaluar objetivamente los sistemas y determinar los niveles de cumplimiento.

- Establecer marcos reguladores para alcanzar la rendición de cuentas y la responsabilidad por el contenido y los resultados de los sistemas de inteligencia artificial en las diferentes etapas de su ciclo de vida.

- Establecer requisitos claros de transparencia y explicabilidad de los sistemas de inteligencia artificial para ayudar a garantizar la fiabilidad de dichos sistemas durante todo su ciclo de vida.

Si bien existen algunas cuestiones que considero que pueden ser difíciles de materializar, como puede ser establecer un consenso internacional sobre qué podemos considerar inclusivo y transparente, por ejemplo, sí que coincido en la mayoría de los objetivos aquí analizados.

Se hace totalmente necesario elaborar una legislación homogénea en materia de inteligencia artificial. Es evidente, que un primer paso, debe centrarse en la elaboración de normas nacionales o incluso regionales, como es el caso de la Ley de Inteligencia artificial que se está elaborando por parte de la Unión Europea. No obstante, el paso definitivo y totalmente necesario reside en la elaboración de una norma internacional, que debiera ser firmada y ratificada por todos los Estados, es decir, un texto

con aplicabilidad y, por tanto, con carácter sancionador y coercitivo a nivel mundial. En esta materia, no basta con que algunos estados ratifiquen, la ratificación de un texto sobre la presente materia debe ser global, ya que lo que está en juego es la salvaguarda de la propia humanidad como especie.

En relación a los mecanismos de certificación, por ejemplo, la Unión Europea ya está trabajando en su regulación. La Ley de Inteligencia artificial dedica una sección completa a esta cuestión que lleva por título “Normas, evaluación de la conformidad, certificados y registro”. Asimismo, la presente norma tiene todo un capítulo dedicado a la Gobernanza donde se establece la creación del Consejo Europeo de Inteligencia Artificial, así como la creación de un Grupo de expertos científicos independientes y la designación de autoridades nacionales competentes en la materia.

4.4. Política de datos

El presente ámbito de aplicación se centra en una de las cuestiones que más preocupa en relación con los agentes artificiales. Su impacto en nuestro derecho a la intimidad y a la protección de nuestros datos personales.

En este sentido, este ámbito de actuación se centra en las presentes cuestiones:

- Elaborar estrategias que garanticen la evaluación continua de la calidad de los datos de entrenamiento para los sistemas de inteligencia artificial.

- Establecer mecanismos de retroalimentación para aprender de los errores y compartir las mejores prácticas entre todos los actores de la inteligencia artificial.

- Establecer salvaguardias para proteger el derecho a la intimidad, así como a la protección de datos de carácter personal.

- Promover repositorios abiertos de datos y códigos fuente públicos o de financiación pública y fideicomisos de datos, a fin de apoyar el intercambio seguro, equitativo, legal y ético de datos, entre otros.

- Promover y facilitar el uso de conjunto de datos sólidos y de calidad para el entrenamiento, el desarrollo y la utilización

de los sistemas de inteligencia artificial y ejercer vigilancia al supervisar su recopilación y utilización.

- Alentar a las empresas del sector privado a que compartan con todas las partes interesadas, los datos que recopilan, en beneficio de la investigación, la innovación o el interés público.

Sin duda, a medida que avanzamos en los ámbitos de actuación propuestos por la UNESCO, nos damos cuenta de la necesidad de establecer un marco internacional global en la presente materia. Un claro ejemplo es la protección de datos de carácter personal. Si bien, la presente materia está empezando a tener recorrido en las legislaciones nacionales, es evidente que, a nivel mundial, la situación es muy dispar. Aún existen muchos países que o bien no disponen de normativa relativa a la protección de datos, o bien disponen de normativa sectorial sobre la materia.

Asimismo, nótese, que en relación a los datos abiertos (*open data*) pero aplicados al campo de la inteligencia artificial, la propia UNESCO establece que se debe alentar, es decir, animar o estimular a que el sector privado comparta información. Sin duda, la UNESCO está cargada de buenas intenciones pero dudo mucho que el sector privado comparta información sino es a través de una normativa que implique la realización de dicho comportamiento.

4.5. Desarrollo y cooperación internacional

Una vez más, en el presente ámbito de actuación, me toca repetirme. No basta con las buenas intenciones en una materia como la presente. Se hace necesario una norma internacional mundial ratificada por todos los Estados y en donde se contemple, entre otras muchas cosas, cómo vamos a lograr el presente ámbito de actuación: el desarrollo y la cooperación internacional que se sustenta sobre las siguientes cuestiones:

- A nivel mundial se debe dar prioridad a la ética de la inteligencia artificial.
- Establecer plataformas de cooperación internacional.
- Procurar promover la colaboración internacional en materia de investigación e innovación en inteligencia artificial.

- Alentar la cooperación y colaboración internacionales en el ámbito de la inteligencia artificial para salvar las divisiones geotecnológicas.

4.6. Medio ambiente y ecosistemas

Otra cuestión de interés es el desarrollo sostenible y la protección del medio ambiente y los ecosistemas. Este tema, quizá no es el más tratado por la doctrina pero considero que todos los Estados, en su conjunto, debieran ir evaluando el impacto ambiental, tanto de corte directo como de corte indirecto que implica el desarrollo y puesta en funcionamiento de los agentes artificiales. Está claro que cuestiones como la huella de carbono, el consumo de energía y, por supuesto, el impacto ambiental debe ser analizado convenientemente. Un análisis exhaustivo de la presente cuestión incluso podría implicar que se desarrollen normas jurídicas que de alguna forma limiten el desarrollo de estos sistemas

4.7. Género

La igualdad de género es una cuestión que tiene un impacto transcendental en la humanidad y por la cual, a nivel nacional, regional e internacional se están dando grandes pasos que aún así, son insuficientes. Siendo conscientes de la importancia y la necesidad de reducir la brecha de género, la UNESCO busca los siguientes objetivos, en lo que a la inteligencia artificial se refiere:

- Lograr la igualdad de género y garantizar que no se conculcan los derechos humanos y las libertades fundamentales de las niñas y las mujeres, ni su seguridad e integridad, en ninguna etapa del ciclo de vida de los sistemas de inteligencia artificial.

- Velar por que las políticas digitales nacionales incluyan un plan de acción en materia de género destinado a apoyar a las niñas y las mujeres para que no queden excluidas de la economía digital impulsada por la Inteligencia artificial.

- Asegurarse de que estas tecnologías no exacerbén las ya amplias brechas que existen entre los géneros en varios ámbitos del mundo analógico.

- Velar porque los estereotipos de género y los sesgos discriminatorios no se trasladen a los sistemas de inteligencia artificial.

- Alentar la iniciativa empresarial, la participación y el compromiso de las mujeres en todas las etapas del ciclo de vida de los sistemas de inteligencia artificial.

- Promover la diversidad de género en la investigación en IA en el mundo universitario y la industria.

No me cabe duda de que en esta materia hay mucho trabajo por hacer. Un trabajo que es totalmente necesario. En el siglo XXI la brecha de género debiera ser algo minoritario y en clara recesión. No obstante, permítame el lector que en la presente materia no sea tan optimista. En primer lugar, dada las claras diferencias culturales que existen en el mundo. Tal y como establece el observatorio de la islamofobia en los medios: “[...] es importante reconocer que la brecha de la desigualdad de género es más difícil de reducir en las sociedades donde la religión sigue siendo un referencial fundamental – como es el caso de la mayoría de los países arabo-musulmanes”⁴¹.

La religión ciertamente puede ser una de las cuestiones que impidan reducir la brecha pero que no se me malinterprete, no es la única. Cuestiones de carácter político, ideológico o incluso social pueden ayudar a que la brecha siga estando presente. ONU Mujeres, sin entrar en la cuestión de la inteligencia artificial propiamente dicha, hace argumentaciones ciertamente desalentadoras: “La diferencia salarial de género permanece en el 23% a nivel mundial y, si no se toman medidas contundentes, pasarán otros 68 años antes de lograr la igualdad en esta esfera”⁴².

Creo que el punto clave es la *contundencia*. Si queremos lograr avances en la reducción de la brecha de género hay que ser

⁴¹ Observatorio de la islamofobia en los medios (2019), *La igualdad de género en el mundo musulmán: se necesita una visión reformista*. Recuperado de <https://www.observatorioislamofobia.org/2019/09/30/la-igualdad-genero-mundo-musulman-se-necesita-una-vision-reformista/>.

⁴² ONU Mujeres (2018), *Hacer las promesas realidad: La igualdad de género en la agenda 2030 para el desarrollo sostenible*. Recuperado de <https://www.unwomen.org/sites/default/files/Headquarters/Attachments/Sections/Library/Publications/2018/SDG-report-Gender-equality-in-the-2030-Agenda-for-Sustainable-Development-2018-es.pdf>.

contundentes. En lo que a la inteligencia artificial y esta materia respecta, vuelvo a repetirme, necesitamos una regulación internacional de escala mundial.

4.8. Cultura

Un ámbito de actuación de “*creciente actualidad*” en tanto en cuanto, hoy en día, se están viendo grandes avances en los agentes artificiales enfocados al presente sector. En este sentido, la UNESCO establece las siguientes cuestiones:

- Incorporar agentes artificiales a la preservación, el enriquecimiento, la comprensión, la promoción, la gestión y la accesibilidad del patrimonio cultural, material, documental e inmaterial.

- Se debe examinar y abordar las repercusiones generadas en la cultura por los sistemas de inteligencia artificial.

- Se debe promover la educación en inteligencia artificial de los artistas y profesionales creativos.

- Se debe promover investigaciones que analicen la intersección entre la inteligencia artificial y la propiedad intelectual.

- Alentar a los museos, las galerías, las bibliotecas y los archivos de ámbito nacional a que utilicen sistemas de inteligencia artificial para realizar sus colecciones y enriquecer sus bibliotecas, bases de datos y bases de conocimientos, ofreciendo al mismo tiempo acceso a sus usuarios.

Más allá de los claros beneficios que pueden tener la inteligencia artificial en el desarrollo y progreso de la cultura sí que es cierto que necesitamos resolver algunas cuestiones esenciales. La más importante, en mi opinión es clarificar qué sucede con las creaciones de carácter intelectual desarrolladas por agentes artificiales. ¿Se pueden proteger obras generadas por agentes artificiales en base a las normativas de propiedad intelectual existentes en la actualidad? ¿Se hace necesario replantear la propiedad intelectual y su normativa? Considero que es la primera cuestión que debe ser resuelta.

4.9. Educación e investigación

En el presente ámbito de actuación la UNESCO plantea las siguientes cuestiones:

- Impartición para todo el mundo de conocimientos adecuados en materia de inteligencia artificial.
- Promover programas generales de sensibilización sobre los avances en inteligencia artificial.
- Alentar las iniciativas de investigación sobre la utilización responsable y ética de las tecnologías de la inteligencia artificial en la enseñanza, la formación de docentes y el aprendizaje electrónico.
- Velar porque las tecnologías de la inteligencia artificial empoderen a los educandos y los docentes y mejoren su experiencia, teniendo presente que los aspectos relacionales y sociales y el valor de las formas tradicionales de educación son fundamentales.
- Velar porque los sistemas de inteligencia artificial apoyen el proceso de aprendizaje sin reducir las capacidades cognitivas y sin recabar información sensible.
- Promover la inclusión digital en los programas de educación en materia de inteligencia artificial en todos los niveles.
- Elaborar planes de estudios sobre la ética de la inteligencia artificial para todos los niveles y promover la colaboración cruzada entre la enseñanza de competencias técnicas de inteligencia artificial y los aspectos humanísticos, éticos y sociales de educación en inteligencia artificial.
- Promover y apoyar investigaciones sobre la ética de la inteligencia artificial.
- Promover públicamente las mejores prácticas de los investigadores y las empresas que desarrollan la inteligencia artificial de forma ética y cooperar con ellos.
- Velar por que los investigadores en inteligencia artificial reciban formación en ética de la investigación y exigirles que tengan en cuenta consideraciones éticas en sus concepciones, productos y publicaciones.
- Alentar a las empresas a que faciliten el acceso de la comunidad científica a sus datos para la investigación.
- Velar por que cualquier evolución futura relacionada con las tecnologías de la inteligencia artificial se base en investigaciones científicas rigurosas e independientes.

- Promover la investigación interdisciplinaria en inteligencia artificial mediante la inclusión de disciplinas distintas de la ciencia, la tecnología, la ingeniería y las matemáticas, como pueden ser los estudios culturales, la educación, la ética, las relaciones internacionales, el derecho, la lingüística, la filosofía, las ciencias políticas, la sociología y la psicología.

No me cabe duda de que nos encontramos en un punto de inflexión. El desarrollo de los sistemas de inteligencia artificial puede generar una revolución sin precedentes en la historia del ser humano, mayor si cabe, que la generada con el desarrollo de las tecnologías de la información y de las comunicaciones. Por este motivo tanto la formación como la investigación me parecen dos cuestiones esenciales que deben ser potenciadas.

Por un lado, considero que toda persona debe tener unos conocimientos, al menos esenciales, sobre inteligencia artificial y, por tanto, en la educación reglada, incluso a niveles de primaria, se debería acercar al estudiante al conocimiento relacionado con la inteligencia artificial. Evidentemente, no sólo la educación reglada debe encargarse de hacer llegar el conocimiento. Por este motivo, los Estados creo que tienen no sólo el deber sino también la obligación de generar planes de sensibilización para poder llegar al mayor número de personas.

Por otro lado, aquellas personas que investigan, diseñan y desarrollan sistemas de inteligencia artificial debiera tener una formación multidisciplinar donde se aúnen tanto la ciencia como las humanidades. Los sistemas de inteligencia artificial, por su propia definición y características deben llevar *incorporados* una serie de valores y una ética para el desarrollo de sus actividades. Por tanto, sus creadores, deben ser los encargados de dar esos primeros valores éticos a los sistemas de inteligencia artificial y, por este motivo, su pensamiento debe ser completamente multidisciplinar.

4.10. Comunicación e información

Otra de las cuestiones sobre las que pueden impactar los sistemas de inteligencia artificial es en la libertad de expresión y el acceso a la información. Por este motivo, la UNESCO considera que sobre el presente ámbito de actuación es necesario:

- Utilizar los sistemas de inteligencia artificial para mejorar el acceso a la información y el conocimiento.

- Garantizar que los sistemas de inteligencia artificial respetan y promueven la libertad de expresión y el acceso a la información.

- Lograr que los Estados miembros inviertan en competencias digitales de alfabetización mediática e informacional y promoverlas.

- Que se creen entornos propicios para que los medios de comunicación tengan los derechos y recursos necesarios para informar eficazmente sobre las ventajas e inconvenientes de los sistemas de inteligencia artificial.

- Alentar a los medios de comunicación a que hagan un uso ético de los sistemas de inteligencia artificial en su trabajo.

Sin duda alguna, “las fake news o noticias falsas se han convertido en una constante dentro del Periodismo y en un verdadero problema a la hora de discernir la verisimilitud de una información”⁴³. El avance de los sistemas de inteligencia artificial no hace sino agravar la presente cuestión.

4.11. Economía y trabajo

Bajo mi punto de vista, la presente es una de las cuestiones que más nos debe preocupar como sociedad. ¿Qué impacto va a tener el desarrollo de los agentes artificiales sobre el empleo? En este sentido, el presente ámbito de actuación se centra en las presentes cuestiones:

- Se debe evaluar y abordar el impacto de los sistemas de inteligencia artificial en los mercados de trabajo.

- Se debe garantizar una transición equitativa a los empleados en situación de riesgo.

- Se deben explorar programas de protección social para aquellos que no puedan reconvertirse.

⁴³ M. Alonso González, “Las *fake news* o noticias falsas se han convertido en una constante dentro del Periodismo y en un verdadero problema a la hora de discernir la verisimilitud de una información”, in *Revista internacional de comunicación*, (2019), n. 45, pp. 29-52.

- Se deben adoptar medidas adecuadas para garantizar la competitividad de los mercados y la protección de los consumidores con la finalidad de evitar los abusos de posición dominante.

Considero que estos cuatro puntos son imprescindibles y deben ser abordados por todos los Estados. Debemos analizar cuidadosamente cómo se va a ver modificado el mercado de trabajo con la irrupción de los sistemas de inteligencia artificial.

No debemos olvidar que:

En numerosos estudios relacionados con el potencial impacto de la automatización se encontró que casi la mitad de los empleos en Estados Unidos y el 60% de los trabajos de los países de la Asociación de Naciones del Sudeste Asiático (ASEAN) están en riesgo de ser automatizados⁴⁴.

La automatización supondrá la pérdida de muchos puestos de trabajo. Por tanto, la sociedad debe estar preparada para activar programas de reciclaje de los trabajadores, así como sistemas de protección social para aquellas personas que no logren un correcto acomodo en el mercado laboral. De hecho, la Unión Europea, en 2015, ya estableció la posibilidad “de que debería considerarse seriamente la posibilidad de introducir una renta básica universal”⁴⁵.

Considero que los Estados deben comenzar a trabajar sobre el presente asunto antes de que sea demasiado tarde. Es totalmente necesario anticiparse, en lo que a la presente cuestión se refiere.

4.12. Salud y bienestar social

El último ámbito de actuación planteado por la UNESCO tiene que ver con la salud y el bienestar social. En este sentido, me parece de interés matizar las siguientes cuestiones:

- Se deben emplear sistemas de inteligencia artificial para mejorar la salud humana y proteger el derecho a la vida.

⁴⁴ A.M. López Estupiñán, L. Peña Mesa, “Inteligencia Artificial: El futuro del empleo”, in *Revista Lecciones Vitales*, 1 (2023), pp. 1-7.

⁴⁵ Parlamento Europeo (2015), *Recomendaciones destinadas a la Comisión sobre normas de Derecho Civil sobre robótica*. Recuperado de https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_ES.html.

- El desarrollo y despliegue de los sistemas de inteligencia artificial relacionados con la salud deben ser seguros, eficaces, eficientes y probados desde el punto de vista científico y médico. Asimismo, deben facilitar la innovación y el progreso médico con base empírica.

- Se debe prestar especial atención a la regulación de las soluciones de predicción, detección y tratamiento médicos en los agentes artificiales, con especial énfasis, entre otras cuestiones, de los sesgos, la privacidad y, sobre todo, el que la decisión final sobre el diagnóstico médico y el tratamiento correspondan siempre a seres humanos.

- Desarrollo de investigaciones sobre los efectos y la regulación de los posibles daños de los sistemas de inteligencia artificial para la salud mental.

- Elaborar directrices sobre las interacciones entre seres humanos y los sistemas de inteligencia artificial y sus repercusiones en las relaciones entre los seres humanos, con especial énfasis, en aquellos colectivos más vulnerables como pueden ser las personas de avanzada edad, las personas con discapacidad o los menores de edad.

- Las cuestiones éticas relativas a los sistemas basados en inteligencia artificial utilizados en las neurotecnologías y las interfaces cerebro-ordenador deberían tenerse en cuenta a fin de preservar la dignidad y la autonomía humanas.

- Velar por que los usuarios puedan determinar fácilmente si interactúan con un ser vivo o con un sistema de inteligencia artificial. Asimismo, deben ser capaces de rechazar eficazmente dicha interacción y solicitar la intervención humana.

- Alentar y promover la investigación colaborativa sobre los efectos de la interacción a largo plazo de las personas con los sistemas de inteligencia artificial.

Quizá una de las mayores preocupaciones, a nivel social tiene que ver con la posibilidad de que se use la inteligencia artificial como un sistema o mecanismo que vaya contra nuestra propia existencia. De ahí la preocupación y especial sensibilidad por el respeto a la salud y a la vida humana. Si bien, es algo que nos debe preocupar, considero que se pueden dar otros usos, que nos pueden pasar más desapercibidos y que puedan tener una mayor

repercusión para el ser humano. Por ejemplo, cuestiones relativas a cómo van a interactuar las inteligencias artificiales con los colectivos más vulnerables, ya que pueden ser mecanismos que pueden utilizarse para modificar comportamientos, actitudes o incluso pensamientos en los presentes colectivos.

Asimismo, también considero que es necesario poner el foco en cómo la interacción con los sistemas de inteligencia artificial nos puede afectar en nuestras relaciones con el resto de los seres humanos. Es posible que pueda darse un efecto de asilamiento social en personas que sólo tengan interés por interactuar con sistemas de inteligencia artificial. Si bien, “la literatura revela que las TIC pueden ser un medio para aliviar el sentimiento de soledad y la percepción de aislamiento social en adultos mayores durante los contextos de crisis”⁴⁶, es posible que este efecto pueda darse a la inversa en las interacciones con agentes artificiales, tanto en situaciones de normalidad como en situaciones de crisis.

Por último, también creo que las personas tenemos derecho a decidir. Decidir si deseamos entablar una interacción con un agente artificial o si, por el contrario, deseamos que nuestra interacción se realice con un ser vivo, es decir, con un ser humano.

5. Conclusiones

Nos encontramos inmersos en una revolución tecnológica de dimensiones titánicas y que ya está transformando nuestra forma de vida, nuestro empleo y cómo nos relacionamos. Sin duda uno de los avances más prometedores y que, al mismo tiempo más preocupación nos aporta es el desarrollo de los sistemas de inteligencia artificial.

Para el desarrollo de los sistemas de inteligencia artificial se hace necesario reconocer a la ética como una disciplina de relevancia en la presente materia. Es totalmente necesario, para el correcto desarrollo de la inteligencia artificial, la aplicación de una serie de consideraciones éticas y morales que son esenciales

⁴⁶ A. Urbina Carreño, “La relación entre el uso de las TIC y el sentimiento de soledad: una mirada desde los Determinantes Sociales de la Salud de las personas mayores en Chile”, in *Revista española de comunicación en salud*, suplemento, 1 (2023), pp. 74-80.

si queremos garantizar un desarrollo responsable y respetuoso de la presente tecnología.

Como hemos tenido oportunidad de ver, definir a la presente tecnología no es sencillo. Es una cuestión compleja y en constante evolución, debido al rápido desarrollo de la materia y a los constantes cambios sobre la percepción de esta disciplina. Aunque no existe un consenso en torno a la definición sí que podemos observar algunas características de los sistemas de inteligencia artificial tales como la capacidad de autonomía, su aprendizaje, su adaptabilidad y su capacidad para la toma de decisiones. También hemos tenido oportunidad de ver que distintas instituciones y organizaciones han realizado importantes esfuerzos para tratar de llegar a una definición de inteligencia artificial e incluso hemos analizado enfoques doctrinales que llegan a considerar aspectos, dentro de los agentes artificiales, tales como el comportamiento inteligente y la conciencia.

A grandes rasgos podemos concluir que los sistemas de inteligencia artificial son sistemas capaces de comportarse de manera inteligente, tienen capacidad para aprender, adaptarse y tomar decisiones y los podemos dividir en dos categorías: agentes virtuales y agentes físicos.

No obstante, creo que debemos tener muy presente que debemos seguir avanzando en la definición de inteligencia artificial. Para ello es necesario seguir revisando y actualizando el concepto de inteligencia artificial a medida que avance la tecnología y vayamos profundizando en el entendimiento de la presente disciplina.

A lo largo del presente trabajo hemos analizado la Declaración de Montreal y hemos tratado de desarrollar una opinión crítica sobre los principios éticos en ella contenidos. Dichos principios tienen una clara misión: guiar el desarrollo y la implementación de la inteligencia artificial de manera que se beneficie a la humanidad. Hemos sido críticos en nuestras reflexiones. Muchos de los principios planteados por la presente Declaración adolecen de definiciones claras y concisas lo cual no beneficia a su aplicación práctica o a su implementación efectiva.

Asimismo, algunos de sus principios plantean desafíos en lo relativo a la prevención del uso de inteligencias artificiales para

incluir en las decisiones de las personas o en torno a que a las personas se les puede imponer un estilo de vida específico gracias a los agentes artificiales.

Surgen muchos interrogantes sobre cómo proteger la intimidad y los datos personales en un mundo cada vez más digitalizado y fiscalizado. Asimismo, debemos enfrentar claros desafíos en lo relativo a cómo van a interactuar los agentes artificiales con los seres humanos y sobre todo, si estamos dispuestos a que las relaciones humanas de calidad puedan ser sustituidas por relaciones con agentes artificiales en un futuro próximo.

Enfrentamos un desafío de importancia en relación a la equidad social dado que el uso y desarrollo de la inteligencia artificial se va a ver claramente condicionado por unas condiciones locales y culturales muy variadas a lo largo del mundo. Nuestras sociedades no tienen las mismas condiciones locales y culturales y, por tanto, esto afecta a una materia como la inteligencia artificial.

Los Estados deben armarse de mecanismos que aborden los posibles usos perjudiciales de la inteligencia artificial, así como de mecánicas para mitigar los riesgos en el uso de la inteligencia artificial y garantizar un uso y desarrollo responsables.

Por otro lado, la responsabilidad humana en la toma de decisiones relacionadas con la inteligencia artificial es fundamental. Por tanto, las personas debemos asumir la responsabilidad final por las acciones de los agentes artificiales de cara a evitar lagunas de responsabilidad y garantizar la rendición de cuentas.

También debemos destacar la importancia de considerar el impacto ambiental en el desarrollo y uso de la inteligencia artificial. Cuestiones como el consumo masivo de energía y el impacto ambiental de estas tecnologías no pueden dejarse en un segundo plano.

Por otro lado, hemos tenido oportunidad de analizar la Recomendación sobre la ética de la inteligencia artificial adoptada por la UNESCO. En el presente texto se reconoce la influencia potencial de la inteligencia artificial en el pensamiento, las interacciones y las decisiones humanas, así como la necesidad de abordar preocupaciones éticas, como la discriminación y la

exclusión social, y sus impactos en la dignidad humana, los derechos humanos y el medio ambiente.

La Recomendación establece una serie de valores y principios que deben guiar el desarrollo y uso de la inteligencia artificial, incluyendo el respeto a los derechos humanos, la promoción de la diversidad y la inclusión, y la garantía de la seguridad y la transparencia. Además, propone áreas de acción específicas, como la evaluación del impacto ético, la gobernanza y administración éticas, y la política de datos, entre otras.

Sin embargo, para que estos principios se conviertan en realidad, es necesario un compromiso global y una cooperación internacional efectiva.

En resumen, el desarrollo y la implementación de la inteligencia artificial plantean una serie de desafíos éticos, sociales y legales que requieren una respuesta integral y colaborativa a nivel internacional. Es necesario garantizar que la inteligencia artificial se utilice de manera responsable y ética para promover el bienestar humano y proteger los derechos fundamentales en todos los ámbitos de la sociedad.

Podemos concluir que los textos analizados en el presente trabajo establecen una base ética para el desarrollo de la inteligencia artificial aunque es evidente que aún quedan muchos aspectos por definir y abordar para garantizar que el desarrollo y uso de la inteligencia artificial sea responsable y en beneficio de la humanidad.

Todos los principios analizados en este trabajo requieren de una atención continua así como de una acción coordinada a nivel global si queremos garantizar un futuro sostenible y equitativo para la humanidad.