

ALGOR-ÉTICA Y UNA MIRADA BIOÉTICA A LA INTELIGENCIA ARTIFICIAL

ALGOR-ETHICS AND A BIOETHICAL LOOK AT ARTIFICIAL INTELLIGENCE

ALGOR-ÉTICA E UMA VISÃO BIOÉTICA DA INTELIGÊNCIA ARTIFICIAL

Gilberto A. Gamboa-Bernal¹

DOI: 10.5294/pebi.2024.28.2.1

PARA CITAR ESTE EDITORIAL / TO REFERENCE THIS EDITORIAL / PARA CITAR ESTE EDITORIAL

Gamboa-Bernal GA. Algor-ética y una mirada bioética a la inteligencia artificial. *Pers Bioet.* 2024;28(2):e2821. DOI: <https://doi.org/10.5294/pebi.2024.28.2.1>

PALABRAS CLAVE (FUENTE: DeCS): Algor-ética; algoritmos; antropología; dignidad; era digital; inteligencia artificial.

KEYWORDS (SOURCE: DeCS): Algor-ethics; algorithms; anthropology; dignity; digital age; artificial intelligence.

PALAVRAS-CHAVE (FONTE: DeCS): Algor-ética; algoritmos; antropologia; dignidade; era digital; inteligência artificial.

¹ <https://orcid.org/0000-0002-1857-9335>. Universidad de La Sabana, Colombia. gilberto.gamboa@unisabana.edu.co

INTRODUCCIÓN

Los días 14 y 15 de noviembre del 2023, la *MIT Technology Review* organizó un encuentro en el auditorio Kresge del Instituto Tecnológico de Massachusetts (MIT) en el que investigadores, tecnólogos, industriales y empresarios se reunieron para dialogar sobre tres campos que están remodelando el mundo: la biotecnología, la inteligencia artificial y el cambio climático (1). En *Persona y Bioética* se han hecho reflexiones sobre el primer y el tercer tema, pero es imperioso empezar a hablar sobre el segundo. En primer lugar, es necesario considerar esas tres fuerzas no de manera aislada sino interactuante, pues funcionan siguiendo un mismo paradigma que las armoniza: lo que se puede y lo que se debe hacer.

El auge de la inteligencia artificial (IA), con el impulso que le dan las Big Tech, lleva a preguntarse no solo por las ventajas, beneficios y potencialidades de estos desarrollos tecnológicos, sino también a plantear reflexiones desde la ética y la bioética; por ejemplo, el término *algor-ética* fue acuñado en 2018 para hacer referencia a las discusiones éticas que giran en torno a desarrollos tecnológicos que incluyan la IA y los algoritmos relacionados con ella (2).

Para la Unión Europea (UE), la IA es “[...] la habilidad de una máquina de mostrar capacidades similares a las humanas, como el razonamiento, el aprendizaje, la planificación y la creatividad” (3). En su origen, la IA está relacionada con la inteligencia humana en dos sentidos, y esto no se debe perder de vista: i) la IA es pensada por seres humanos y ii) se planteó desde el principio como una actividad inspirada en la forma en que funciona el cerebro humano, es decir, con base en los sistemas neuronales, a diferencia de como operaban los primeros computadores. Partiendo del llamado *perceptrón de Rosenblatt*

como unidad fundamental y básica de arquitecturas neuronales complejas y de algoritmos de aprendizaje, se empezó a hablar de una IA capaz de aprendizaje automático, procesamiento de la experiencia, toma de decisiones, reconocimiento de emociones y adaptación a datos invisibles, entre muchas otras funciones (4).

La IA es un instrumento muy poderoso, cuyos efectos dependen del tipo de datos con los que se alimente, pero también de las finalidades a las que se oriente, en última instancia, de su utilización. Además, la IA genera en el ser humano una percepción dual: por un lado, está el entusiasmo que despiertan sus aplicaciones, pero, por otro, está el miedo que generan sus riesgos. En todo caso, no se puede perder de vista que la IA es sobre todo una herramienta, y que los beneficios o perjuicios dependerán del uso que se haga de este instrumento (5).

Hay dos tipos de IA: la clásica, que consiste fundamentalmente en sistemas supervisados, con la capacidad de tomar decisiones, a través de las llamadas máquinas de soporte vectorial, y la IA generativa, aquella que toma datos y características suministrados y hace modelos funcionales a partir del procesamiento del lenguaje. En este texto se hablará de esta segunda forma de IA.

Uno de los centros pioneros en investigación informática y de IA es el Laboratorio de Ciencias de la Computación e Inteligencia Artificial (CSAIL), perteneciente al MIT. Consta de un ecosistema conformado por profesores e investigadores que exploran las diversas alternativas para probar, diseñar y aplicar la IA generativa en una gran cantidad de situaciones, desde la automatización de procesos de codificación hasta problemáticas derivadas del cambio climático.

PRESENCIA ACTUAL Y FUTURA DE LA IA

Algunas aplicaciones de IA se usan desde hace 70 años, pero los vertiginosos avances de las ciencias informáticas (la potencia creciente de los computadores, la disponibilidad y manejo de datos, el diseño de algoritmos cada vez más complejos, etc.) ha permitido que la IA se haga presente en la cotidianidad de muchas más personas. Algunos ejemplos de la presencia actual de la IA son las búsquedas en internet, las compras *online*, los asistentes digitales personales en teléfonos inteligentes, las traducciones automáticas, los asistentes y los sensores para la conducción segura de vehículos, entre otros. En la práctica médica y en la investigación biomédica, la IA analiza grandes cantidades de datos para mejorar el diagnóstico de enfermedades, sugerir pautas para prevenirlas o hacer pronósticos de morbilidad (6).

En el transporte, la IA sirve para optimizar el rendimiento, mejorar la seguridad, dar eficiencia al tráfico o posibilitar la conducción autónoma, entre otras funciones. Sus aplicaciones en la industria son múltiples y cada vez están mejor desarrolladas: en la ciberseguridad, es útil para detectar noticias falsas o combatir la desinformación; en sistemas de alerta, para evitar desastres, incluidos los naturales, y un largo etcétera.

Se puede afirmar que la IA está acelerándolo todo y que además brinda a las personas capacidades creativas que antes se consideraban imposibles de automatizar. En biotecnología el avance también es patente: se mejora y se extiende la vida y se puede alterar tanto el ADN como la misma demografía. Por otra parte, el cambio climático obliga a examinar el impacto ambiental de cada decisión empresarial y las formas para conjurarlo a través de desarrollos tecnológicos.

Como se expuso anteriormente, estas tres fuerzas del cambio funcionan de manera convergente: la IA interacciona con la biotecnología y con las innovaciones que pueden impactar el cambio climático. Esta interacción también tiende a redefinir los modelos y estrategias de negocio, pues configura un futuro donde se hacen presentes la disrupción, la sostenibilidad y una economía verde. El desarrollo de algunas de las tecnologías limpias, que son muy prometedoras, incluye la utilización de la IA y promueve la migración a una economía más sostenible.

La confluencia de la biología y la tecnología, la biotecnología, ha impulsado una profusión de procedimientos y tratamientos médicos que hacen posibles vidas más saludables y prolongadas. La era digital de la atención médica se caracteriza por las terapias genéticas, la medicina programable y personalizada, así como también por nuevos dispositivos médicos de diverso orden. Ahora se habla de la medicina 4P, que, con la ayuda de la IA, es personalizada, preventiva, predictiva y proactiva (7).

La intersección entre la biología y la tecnología ha generado una disminución en la brecha entre las aplicaciones de la IA y la medicina, uno de los campos donde la incursión de la IA está más retrasada. La gran cantidad de procedimientos médicos, diagnósticos y tratamientos que se están desarrollando hacen previsible vidas más prolongadas y saludables a la vez. La terapia génica, la medicina programable y personalizada y los dispositivos médicos con IA ofrecen posibilidades muy amplias, en beneficio de la atención médica para los seres humanos.

Uno de los principales problemas que enfrenta la aplicación de la IA en la fabricación de medicamentos y en el desarrollo de terapias es la gran complejidad para modelar el sistema biológico, lo que lleva a la casi imposibilidad

de predecir los efectos secundarios o adversos de las nuevas moléculas sintéticas obtenidas mediante la IA. Desarrollar la capacidad de adelantar patrones, modelar el sistema biológico y desbloquear la capacidad de programación celular con el fin de resolver el problema de la previsibilidad de los efectos de esos fármacos y de las nuevas terapias es un reto más para la IA.

Los secretos de la biología generativa serán más fácilmente abordables cuando, mediante la IA, el estudio de millones de proteínas lleve a entender las reglas mediante las cuales la naturaleza codifica sus funciones. Este descubrimiento permitirá, entre otras cosas, producir nuevos fármacos con menos eventos secundarios adversos, así como acortar los tiempos de las distintas fases de su producción.

Algunas tecnologías limpias son prometedoras, ya que impulsan la transición hacia una economía verde, así como el uso de energía que no deje huella de carbono por parte de la industria. También se busca solucionar la problemática de las baterías recargables, dado que hoy en día sus residuos son incluso más contaminantes que las fuentes de energía tradicionales. En este sentido, el futuro se proyecta hacia las reacciones de fusión, pues en la actualidad es posible producir más energía por este medio que la que se consume en el proceso, lo que supone un paso adelante en las perspectivas de la obtención de energía abundante y limpia.

Con seguridad, los avances de la IA generativa cambiarán también la manera en que se hacen los negocios y en que se transforman muchas industrias. Todo esto tendrá su impacto en la forma como se trabajará, como se harán los nuevos contratos laborales y en cómo será el régimen de contratación, entre otros asuntos.

La inserción de la IA en casi todos los ámbitos obliga a que se piensen muchos temas para que esa integración se haga de manera responsable y genere confianza, de modo que, en su utilización, las preocupaciones iniciales se tornen en seguridades. Esto implica pensar cómo mitigar los riesgos, propiciar la seguridad, establecer los lineamientos éticos y la preservación del bien común.

Otro frente importante es la integración de la IA con la robótica, que llegará a la producción de máquinas “inteligentes” más avanzadas que las que existen hoy. Actualmente hay robots funcionando en muchos ambientes laborales, industriales, de entretenimiento, en los hogares; esa presencia se incrementará, de modo que también se ampliará su demanda comercial en esos escenarios y en muchos más.

La disrupción que se producirá en la tecnología será la causa del cambio de muchas estrategias industriales, comerciales, empresariales, educativas, de mercadeo, entre otras. Pero también cambiarán las pautas regulatorias, que deberían incluir cómo probar, monitorear, evaluar y auditar los nuevos dispositivos, para que la gestión de la IA deje atrás el caos inicial de sus aplicaciones y se dé paso a la transparencia en su utilización. Así, se podrá hacer una gestión racional del riesgo y se aprovecharán mejor las tecnologías emergentes.

En el campo del entretenimiento, las aplicaciones de la IA llevarán a que se restructure la forma como se utiliza el tiempo libre y se administra el ocio. En el caso del séptimo arte, la creatividad humana se combinará con contenidos generados por la IA, y la puesta en escena será más versátil; también los tiempos de producción y posproducción se reducirán, haciendo menos costoso el producto final.

La IA podrá impactar positivamente en la crisis alimentaria que vive el planeta, ya que la producción de proteínas cultivadas en un laboratorio será más segura y económica que la producción de alimentos transgénicos. Así se podrían superar los primeros retos de los Objetivos de Desarrollo Sostenible y satisfacer las necesidades alimentarias de una población en crecimiento, mediante el uso racional y ético de herramientas que favorezcan al ser humano, en lugar de buscar eliminar la pobreza prescindiendo de la gente pobre (8).

Los problemas relacionados con la alimentación son otra “pandemia” del siglo XXI de la que poco se habla, pero sus efectos son alarmantes y la IA puede ayudar a combatirlos: casi 1800 millones de personas tienen sobrepeso, alrededor de 900 millones sufren de desnutrición (deficiencias de macro- y micronutrientes) y cerca de 860 millones de personas padecen de obesidad a nivel global (9).

El sector financiero también se verá beneficiado con las aplicaciones de IA: la velocidad se verá incrementada en varios frentes, por ejemplo, en las transacciones, en el estudio de solicitudes de crédito o en la creación de soluciones financieras y de renegociación de deudas.

El desarrollo de tecnologías y herramientas cognitivas emergentes se verá potenciado por la IA, con las consiguientes alertas sobre su efecto en los seres humanos, que van mucho más allá de lo que simplemente supondría un implante de dispositivos en el sistema nervioso central. Las precauciones están orientadas a que esos desarrollos no se constituyan en instrumentos para una eventual consolidación de las ideas y acciones transhumanistas.

Los sistemas generativos de IA pueden fabricar información errónea o falsa, que puede ir desde simples “alucinaciones” o ficciones que distorsionen la realidad hasta presentar esas invenciones como hechos, con las consiguientes consecuencias negativas en cualquiera de los planos del funcionamiento humano. Para que la IA sea más confiable, debe diseñarse de manera que se puedan evitar esos problemas.

La IA puede mejorar la utilización de la química en la economía circular. Uno de los principales problemas de la actualidad es la gran cantidad de residuos que el ser humano genera y que de distintas maneras impactan al medioambiente. La química, con la ayuda de la IA, tiene posibilidades reales de hacer que esos residuos se conviertan en productos útiles, en nuevos materiales o en combustibles limpios.

En el campo educativo, la IA podrá facilitar una cobertura mayor al incrementar el acceso al saber, apoyar exponencialmente la investigación científica y fortalecer la pertinencia de los contenidos para alinearlos más con la solución de problemas reales y cercanos. Este será uno de sus mayores aportes para paliar la inequidad y extender la democratización del conocimiento.

Mediante la IA es posible crear contenido hablado y escrito de diverso tipo, destinado a la comunicación, la información o el entretenimiento. El reto en este campo está en la protección legal que esos contenidos puedan tener, las regulaciones de propiedad intelectual que se deben generar y las leyes que se deben cambiar o crear para orientar la utilización de la IA por las personas, la sociedad o la industria.

NECESIDAD DE LA ALGOR-ÉTICA

En la actualidad es un lugar común afirmar que la IA cambiará el mundo, pues ya lo está haciendo: casi en cada actividad humana se encuentran desarrollos de IA, con distintos grados de complejidad y utilización. Ante este aluvión, es importante plantear preguntas sobre su utilización actual y futura, pero sobre todo dar respuestas desde una perspectiva ética y bioética que ayuden a gestionar adecuadamente esta innovación.

La ética y la bioética son determinantes para discutir sobre la IA. Como, al menos por ahora, la IA es falible (entre otras causas por la escasa calidad de los datos con los que se alimenta y la imperfección de los sensores que pueden aportarlos) y puede cometer errores, hay que elaborar un sistema ético sobre la manera de gestionarla e impedir que dañe a los seres humanos, produzca más injusticias o incremente la desigualdad y los desequilibrios globales (10).

El afán por desarrollar la IA como herramienta omnipresente puede percibirse como una apuesta fuerte para que el presente y el futuro de la era digital sean gobernados por algoritmos hasta constituir una extendida algor-cracia, que demanda imperiosamente una algor-ética que oriente la utilización de esa nueva herramienta y sus aplicaciones.

Un considerable desafío para el desarrollo de la algor-ética está en la capacidad de traducir pautas éticas en algo reconocible por una máquina: la percepción de que una acción es ética o no es una característica propiamente humana, pero será necesario intentar que las máquinas reconozcan el valor ético y lo hagan computable (11).

En este sentido, Paolo Benanti ha planteado la necesidad de unas pautas éticas que se deberían seguir al diseñar

y aplicar la IA. Estas pautas se verán enriquecidas si se tiene claro que uno de los riesgos de tomar decisiones libres está relacionado con la posibilidad de producir efectos no deseados, con la consiguiente incertidumbre; de ahí que sea importante aprender a gestionar ese riesgo. Otro elemento clave es aprender a resolver la tensión entre la búsqueda de la igualdad, motor de casi todas las guerras y revoluciones, y el ideal de felicidad: la IA puede producir muchas desigualdades en su afán de allanar el camino a la felicidad. Pero el mayor reto de todos será poder traducir las pautas éticas en directivas ejecutables por una máquina (12).

Esas pautas éticas se deben soportar en el reconocimiento de los bienes humanos que es necesario proteger y promover, y esto desde las fases iniciales de diseño, de tal manera que la responsabilidad ética se tenga en cuenta desde ese momento, sin pretender descargar dicha responsabilidad en las máquinas (13).

Pero también es importante el diseño de modelos que permitan detectar precozmente cuándo la IA pone en peligro esos bienes relacionados con los seres humanos, pues la interacción directa de la IA (*chatbots*) al sostener conversaciones genera relaciones empáticas con las máquinas pensadas para tranquilizar, acompañar, servir de soporte afectivo o simplemente para “pasar el tiempo”, y puede tener efectos inesperados o inconvenientes, ya que su diseño les permite aprender a responder, de manera personalizada, a las necesidades psicológicas o físicas de los seres humanos (14).

DECÁLOGO DE PAUTAS ÉTICAS

En la creación, el diseño y el uso de sistemas inteligentes se deben considerar elementos que permitan que lo que es técnicamente posible esté a favor de los humanos y

nunca les haga daño ni propicie condiciones adversas o lesivas para las personas humanas. Las siguientes guías éticas ayudarán a que la IA promueva la diversidad, sin fomentar la desigualdad y funcione siempre como instrumento útil para mejorar las condiciones de vida de los seres humanos.

1. Respeto por la dignidad de cada ser humano y de sus derechos fundamentales

Es la primera y principal pauta para considerar. La cultura actual parece tolerar una gran miopía sobre el sentido de lo humano y uno de sus principales atributos: la dignidad. Cualquier IA y sus aplicaciones han de tomar como fundamento la dignidad en su cuádruple dimensión: ontológica, moral, social y existencial. Solo así se podrá evitar la debilidad del *ethos* que no capta el valor y los derechos fundamentales de toda persona humana.

2. Transparencia

Es imperativo que el proceso de desarrollo y la aplicación de la IA tenga transparencia para garantizar que los algoritmos sean comprensibles y reproducibles, de modo que la producción y distribución de las utilidades prácticas de la IA no tenga mayores inconvenientes.

3. Responsabilidad

El mundo tecnológico, productores y usuarios, han de compartir la responsabilidad del diseño y uso de la IA, que parte de garantizar un futuro donde tanto la innovación digital como el progreso tecnológico pongan a la persona humana en el centro. Otro elemento relacionado con la responsabilidad debe ser la garantía de poder informar y justificar las acciones que se emprenden, así como sus consecuencias.

4. Integridad

Como la IA se nutre de datos, la integridad de la información es clave para que la transformación digital sea verdaderamente útil y los desarrollos posteriores sean seguros para el ser humano. Otro frente de la integridad se da en el marco de la gestión académica, pues el uso de la IA puede derivar en plagio; sin embargo, la misma IA puede también desarrollar instrumentos para detectarlo, lo que llevaría a disuadir su mala utilización y garantizar así la integridad académica (15).

5. Seguridad

Aparte de la ciberseguridad, que requiere el manejo de datos, este imperativo ético hace más referencia a la protección que se debe garantizar a los seres humanos. Esta pauta debería ser una refinada extensión de las reglas que Isaac Asimov propuso para las máquinas en el siglo pasado (16).

6. Control de sesgos algorítmicos

Como la IA funciona con algoritmos que se elaboran sobre los datos proporcionados por seres humanos, un sistema informático puede tomar valores o creencias de quienes lo alimentan y constituir un sesgo algorítmico, con los consecuentes peligros de transmitir estereotipos que llegan a lesionar a otros seres humanos, pues se puede impactar temas sensibles como la privacidad o la discriminación. El diseño de la IA debe establecer los mecanismos para el control de este tipo de sesgos (17).

7. Equidad

La equidad, como virtud social, se encuadra entre el derecho positivo y el derecho natural, y permite juzgar

la acción humana más allá de la ley y la justicia, de modo que lo equitativo prime sobre lo justo legal, pero dentro de lo justo natural. Así, la equidad funciona como un principio de racionalidad que permite ir más allá de lo “correcto” (lo indicado por la ley) y pensar en lo que es bueno para la persona singular y para la sociedad.

En relación con la IA, la equidad se debe garantizar al menos en tres ámbitos: en la imparcialidad, que implica no generar prejuicios ni facilitar la actuación a partir de ellos; en el acceso a los recursos de la IA, en cuanto a cobertura y precios; y en la disminución de las inequidades y la mitigación de las desigualdades sociales (18).

8. Sociabilidad

La complejidad del ser humano, que se manifiesta en sus dualidades, le lleva a interrelacionarse con sus semejantes y con su entorno a partir de una característica derivada del trascendental personal de la coexistencia. Esta sociabilidad en el ser humano es el origen de la sociedad y de una de sus actividades más propias: el diálogo. Los desarrollos de la IA han de propiciar esa característica evitando el solipsismo, el sedentarismo, la soledad o el aislamiento social. Si la persona humana necesita dialogar, pues es un ser esencialmente dialogante, los recursos de la IA pueden pensarse para favorecer la formación de la personalidad, modular el propio carácter, apropiarse la lengua, las costumbres y los valores de la sociedad en la que se nace, tener una identidad propia y una personalidad madura integrada con el entorno, con el fin de que se puedan establecer unas relaciones interpersonales adecuadas (19).

9. Solidaridad

Una consecuencia de la pauta anterior es la ayuda que los seres humanos se pueden prestar entre sí; una ayuda que parte de la responsabilidad personal, pues quien no la tiene no es capaz de pensar en el bien de los otros. La solidaridad parte de dos hechos: el Estado no puede llegar a todo y la iniciativa individual no se agota en el interés personal, pero emerge cuando la persona es capaz de captar las necesidades de los otros, lo cual se configura como virtud humana. Los diseños tecnológicos deben ser entonces procesos solidarios que se direccionen de una manera adecuada, considerando que la solidaridad se mide por obras de servicio. En este sentido, la IA tiene un panorama muy extenso, pues con sus aplicaciones puede servir a personas singulares y a la sociedad en general. Por ejemplo, en las nuevas formas de marginación, de fragilidad social o de vulnerabilidad, se pueden encontrar soluciones que van más allá de acuerdos legislativos o decisiones políticas y se cristalizan en actividades de acción social (20).

10. Subsidiariedad

La realización del bien común implica, entre otras cosas, el principio de subsidiariedad que lleva a objetivar acciones por parte de las personas individuales y de la sociedad civil, en las que los Estados no deben o no pueden llegar, y viceversa. Dentro de los múltiples campos en los que se puede aplicar el principio de subsidiariedad, a través de la IA, se puede mencionar el cambio climático. El desarrollo de aplicaciones específicas que incorporen tecnologías con IA puede ayudar a mitigar los efectos del cambio climático, a hacer prospecciones sobre el impacto real de las acciones humanas en su génesis, a

establecer políticas climáticas con objetivos realistas, a dimensionar adecuadamente los efectos del cambio climático y establecer pautas para su oportuno control, entre otros beneficios (21). La subsidiariedad también se puede aplicar en la regulación del mercado, tanto interno como externo, en las transferencias comerciales, en los modelos de negocio, en el derecho comercial y aduanero, en la producción y distribución de bienes de consumo, y un largo etcétera (22).

REGULACIÓN JURÍDICA

Además de las pautas éticas, también hay que pensar en los desafíos legales que los sistemas de IA plantean. La interrelación con los sistemas de IA ha impulsado consensos internacionales sobre estas temáticas. Una de las primeras iniciativas la tomó la Unión Europea (UE). Como la IA es clave en la transformación digital de la sociedad, la UE la ha incluido dentro de sus prioridades, pues busca convertirse en un líder mundial en la economía de datos y sus aplicaciones. Desde el 2021, el Parlamento Europeo organizó una comisión para redactar el primer marco regulatorio de la IA.

Para desarrollar su trabajo, esta comisión se propuso como objetivos establecer una definición neutral y uniforme de lo que es y será la IA, para pasar a garantizar que los sistemas que la utilizan o la incorporan (herramientas, servicios, componentes, procesos, etc.) en la UE tengan las siguientes características: que sean seguros, transparentes, rastreables, no discriminatorios y respetuosos con el medioambiente. Es llamativo que en esos objetivos se mencione expresamente el respeto por el medioambiente, pero no a la persona humana. Sin embargo, sí se deja constancia de lo importante que es la supervisión por parte de seres humanos para evitar

consecuencias perjudiciales por automatizar esta tarea. Esta reglamentación es resultado de varios estudios sobre los beneficios y los riesgos de la IA (23).

La Ley de Inteligencia Artificial fue aprobada por el Parlamento Europeo el 13 de marzo del 2024, con 523 votos a favor, 46 en contra y 49 abstenciones (24), y contiene una clasificación para los sistemas de IA con un enfoque basado en el riesgo. Algunos sistemas que presentan “riesgos inaceptables” fueron prohibidos. Aquellos que son tipificados como de “alto riesgo”, por tener un impacto perjudicial en la vida, la salud, la seguridad o sobre los derechos fundamentales de los seres humanos, son autorizados, pero sujetos a una estricta serie de requisitos y obligaciones. Los sistemas de IA que presentan “riesgos limitados” por la falta de transparencia quedaron sujetos a unos requisitos de información y aquellos otros con “riesgo mínimo” no tuvieron obligaciones adicionales.

La ley también incluye reglas específicas de uso general sobre la inteligencia artificial, límites al uso de sistemas de identificación biométrica por parte de las fuerzas de seguridad, prohibición del uso de la IA para manipular o explotar las vulnerabilidades de los seres humanos, y el derecho de los usuarios a presentar reclamaciones y recibir explicaciones (25).

REFLEXIONES FINALES

El panorama actual se caracteriza, entre otras cosas, por la presencia de una coyuntura única en la que la tecnología está presente en casi todos los ámbitos a escala global, desde estrategias empresariales y comerciales hasta políticas públicas. Por eso, además de revisar el estado actual de la investigación y la tecnología en estos

temas, es importante preguntarse por el impacto de los cambios y por las estrategias que se deben establecer para guiar, con perspectiva ética y bioética, las iniciativas aplicadas que van surgiendo, de tal manera que se respete profundamente al ser humano y al medioambiente.

El uso de la libertad no se detiene en tomar decisiones, en elegir entre el bien o el mal; la ética apunta a que el ser humano aprenda a utilizar la libertad para elegir el bien, pero sobre todo para ejercitarse en llevar una vida plenamente humana. Para esto es imprescindible una comprensión cabal de la persona, de su propósito en la vida, de su interrelación con los demás, de su capacidad de trascender y de su papel como gestor de su propia cultura.

Como la IA puede tomar decisiones independientes sin el concurso de los seres humanos, es necesario garantizar y proteger los mecanismos de control relevantes para que tales decisiones no pongan en peligro la vida, la integridad o la dignidad de ningún ser humano.

La IA generativa tiene que desarrollarse en el ámbito del código abierto, un nuevo enfoque que promueve democratizar el acceso a la información, garantizar la transparencia y mejorar la veracidad en el desarrollo investigativo, y que implica un equilibrio entre las fuentes de información, las barreras a la innovación y las regulaciones legales y éticas.

La IA está un tanto rezagada en el campo médico, tal vez porque sus implicaciones éticas y bioéticas han sido, hasta el momento, poco estudiadas y desarrolladas. La relación entre seres humanos y máquinas en el campo médico, y en otras áreas de aplicación de la IA, tendrá que moverse en el marco de la supervisión humana directa, con lo que se empieza a denominar *medicina simbiótica*: la

responsabilidad de los equipos de salud con los pacientes hace necesaria una asociación especial entre el personal de salud y los instrumentos que incorporan la IA (26).

Así como en el diseño de la IA se incluye la capacidad de captar y expresar emociones humanas, tal vez sea oportuno plantearse que las máquinas tengan una especie de sensación de incertidumbre. Esto puede ser importante para que cuando el mecanismo de IA no sepa con certeza cómo proceder para salvaguardar un valor humano (la vida en primer término), deba solicitar el concurso de un ser humano. La presencia de la incertidumbre puede ser el núcleo de la toma de decisiones, para hacer realidad que el desarrollo de la IA se forje con diseños centrados en el ser humano.

Con este fin, es imprescindible tener claridad sobre algunos conceptos antropológicos básicos, que llevan a plantear el desarrollo de la IA como una cuestión filosófica y ética, antes que puramente técnica o tecnológica. Para apuntar al bien particular y al bien común, y construir un mejor futuro para el ser humano y el ecosistema donde vive, es necesario que en las decisiones no cuenten solo los resultados, los efectos, la utilidad, sino que igualmente es perentorio tener claros los objetivos que se persiguen, las intenciones, las virtudes y valores que entran en juego, así como también las responsabilidades y deberes que se derivan de cada desarrollo de IA.

Los beneficios reales o potenciales no pueden obnubilar ni ocultar los riesgos concretos, ni facilitar la extensión de modelos familiares, culturales o socioeconómicos que puedan poner en peligro al ser humano singularmente considerado o a la misma especie humana. La personificación que hace la IA del paradigma tecnológico no puede ser concomitante con un paradigma tecnocrático, de

pensamiento único, que perpetúe las fallas estructurales de la cultura contemporánea, sino que, más bien, la IA puede ayudar a evitar la extensión de esa ideología que pretende encumbrar la tecnociencia en la vida económica y social, en aras de la utopía de un progreso falsamente sostenible y de una evolución transhumanista.

Hay que evitar el peligro de enfocarse solo en las reglas y dejar de lado al ser humano y a los objetivos primarios de su vida, como ser individual y social. Hay que pensar en la forma como la IA puede ayudar al humano a ser lo que es, por lo que resulta imprescindible humanizar a la IA sobre la base de una antropología verdadera, para garantizar su utilización al servicio de lo humano. Si la primacía es la dignidad humana, la IA podrá ayudar a configurar una cultura de la vida y a evitar que se siga estableciendo globalmente una cultura de la muerte.

REFERENCIAS

1. MIT Technology Review. EmTech. Big ideas. Big decisions. Big impact [Internet]. [consultado 2024 Mayo 31]. Disponible en: <https://event.technologyreview.com/emtech-mit-2023/detailed-agenda>
2. Benanti P. Oracoli: Tra algoretica e algocrazia. Roma: Ed. Luca Sossella; 2018.
3. European Parliament. What is artificial intelligence and how is it used? [Internet]. [consultado 2024 Mayo 31]. Disponible en: <https://www.europarl.europa.eu/news/it/headlines/society/20200827STO85804/che-cos-e-l-intelligenza-artificiale-e-come-viene-usata>
4. Boreland B, Kunze H, Levere KM. Artificial Neural Networks. En: Kunze H, De La Torre D, Riccoboni A et al., editores. Engineering mathematics and artificial intelligence: foundations, methods, and applications. Boca Ratón, FL: CRC Press; 2024. pp. 227-244. DOI: <https://doi.org/10.1201/9781003283980-10>
5. Colagrande F, Tridente G. 50 domande & risposte sull'intelligenza artificiale. Roma: EDUSC; 2024.
6. Zhu Z, Shi D, Guankai P et al. Retinal age gap as a predictive biomarker for mortality risk. *Br J Ophthalmol*. 2023;107:547-54. DOI: <https://doi.org/10.1136/bjophthalmol-2021-319807>
7. Schmidt-Bortolini V, Colombo C. Artificial intelligence in medicine: the need to see beyond. *Brazilian J of Law, Technology and Innovation*. 2024 Ene 30;2(1):71-89. DOI: <https://doi.org/10.59224/bjlti.v2i1.71-89>
8. Gamboa-Bernal G. Objetivos de Desarrollo Sostenible (ODS) y pospandemia: mirada bioética a un nuevo marco de reflexión y aplicación. En: López-González SP, coordinador. Objetivos de Desarrollo Sostenible de la Agenda 2030. Una visión iberoamericana. Ciudad de México: Tirant lo Blanch; 2022. pp. 69-91.
9. Echeverry-Raad J. A falsehood that has been repeated many times becomes true, the origin of the diabetes pandemic, the most lethal of the 21st century? *J Diab Metab Disorder*. 2024;11(1):39-50. DOI: <https://doi.org/10.15406/jdmcd.2024.11.00276>
10. Benati P. Human in the loop. Decisioni umane e intelligenze artificiali. Milán: Mondadori; 2022.
11. Benati P. Pacem in cyberspace, auspicia algoretichs. En: Ruggie F, editor. AI in the age of cyber-disorder. actors, trends, and prospects. Milán: ISPI-Brookings; 2020. pp. 134-153.
12. Benanti P. Algor-etica: l'incertezza ci salverà [Internet]. [consultado 2024 Jun 27]. Disponible en: <https://www.paolobenanti.com/post/2019/02/13/algor-etica-ricerca-universale>
13. Villalba JF. Algor-ética: la ética en la inteligencia artificial. *An Fac Cienc Jurídicas Soc Univ Nacional de la Plata*. 2020;17(50):679-698. DOI: <https://doi.org/10.24215/25916386e062>
14. Peroraro R, Curzel E. Convocatoria de Roma por la ética de la IA: el nacimiento de un movimiento. *Medicina y Ética*. 2023;34(2):315-349. DOI: <https://doi.org/10.36105/mye.2023v34n2.01>
15. Kumar R, Eaton SE, Mindzak M, et al. Academic integrity and artificial intelligence: an overview. En: Eaton SE, editor.

- Second Handbook of Academic Integrity. Cham: Springer; 2024. pp. 1583-96. DOI: https://doi.org/10.1007/978-3-031-54144-5_153
16. Morandín-Ahuerma F. Veintitrés principios de Asilomar para la inteligencia artificial y el futuro de la vida. En: Morandín F, editor. Principios normativos para una ética de la inteligencia artificial. Puebla: Concytep; 2023. pp. 5-27. DOI: <https://doi.org/10.31219/osf.io/dgnq8>
 17. Arora A, Barrett M, Lee E et al. Risk and the future of AI: algorithmic bias, data colonialism, and marginalization. *Inf Organization*. 2023;33(3): 100478. DOI: <https://doi.org/10.1016/j.infoandorg.2023.100478>
 18. Cachat-Rosset G, Klarsfeld A. Diversity, equity, and inclusion in artificial intelligence: an evaluation of guidelines. *Appl Artificial Intelligence*. 2023;37(1):2176618. DOI: <https://doi.org/10.1080/08839514.2023.2176618>
 19. Banerji R. Artificial intelligence, humanness, and nonverbal sociality. *Anthropology in Action*. 2023;30(3):49-58. <https://doi.org/10.3167/aia.2023.300306>
 20. Rudschies C. Exploring the concept of solidarity in the context of AI: an ethics in design approach. *Digital Society*. 2023;2(1):1. DOI: <https://doi.org/10.1007/s44206-022-00027-x>
 21. Cortez F. Artificial intelligence, climate change and innovative democratic governance. *European J Risk Regulation*. 2023;14(3):484-503. DOI: <https://doi.org/10.1017/err.2023.60>
 22. Pazos-Castro R. El principio de subsidiariedad en la normativa europea para un mercado único digital. *Cuad Derecho Transnacional*. 2023;15(2):801-32. DOI: <https://doi.org/10.20318/cdt.2023.8079>
 23. European Parliament. Artificial intelligence: threats and opportunities [Internet]. [consultado 2024 Jun 30]. Disponible en: <https://www.europarl.europa.eu/topics/en/article/20200918STO87404/artificial-intelligence-threats-and-opportunities>
 24. Parlamento Europeo. Ley de Inteligencia Artificial [Internet]. [consultado 2024 Jun 30]. Disponible en: https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138_ES.pdf
 25. European Parliament. Artificial Intelligence Act: MEPs adopt landmark law [Internet]. [consultado 2024 Jun 30]. Disponible en: <https://www.europarl.europa.eu/news/en/press-room/20240308IPR19015/artificial-intelligence-act-meps-adopt-landmark-law>
 26. Lee P, Goldberg C, Khane I. The AI revolution in medicine: GPT-4 and beyond. Hoboken, NJ: Pearson; 2023.