



LH

HUMANIZACIÓN, PASTORAL Y ÉTICA DE LA SALUD Y SOCIAL

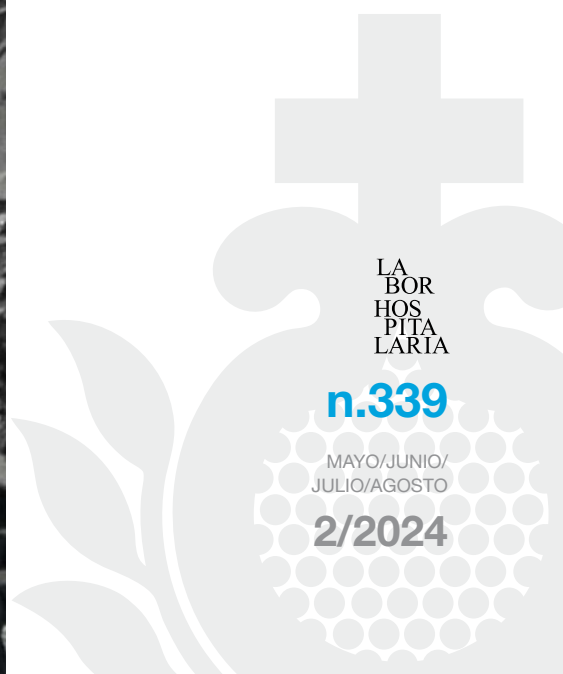
La Inteligencia Artificial al cuidado de la persona (I)

LA
BOR
HOS
PITA
LARIA

n.339

MAYO/JUNIO/
JULIO/AGOSTO

2/2024





Orden Hospitalaria de
San Juan de Dios
Provincia de España

Provincia San Juan de Dios de España

Año 71. Tercera Época
Mayo/Junio/Julio/Agosto
Número 339. Volumen LIV

Consejo de Redacción

Dirección

Calixto Plumed Moreno O.H.

Director adjunto

José María Galán González-Serna

Coordinadores

Humanización

Isabel Grimal; Josep Antoni Boix

Pastoral de la Salud y Social

Begoña Moreno Guinea; Susana Queiroga

Ética de la Salud y Social

Carmen Massé; José María Bermejo OH

Redacción - Maite Hereu

Administración - Dolores Sáenz

Consejo Asesor

Humanización

Jesús Pineda OH; Anna Ramió; Raquel Sisas

Ética de la Salud y Social

Jacinto Bátiz; Margarita Bofarull, rscj;

M^a Pilar Núñez-Cubero; Anna M. Prats;

Manuel de los Reyes López

Pastoral de la Salud y Social

Amador Fernández OH; Marije Goikoetxea,

José Luis Méndez; Mercé Puig-Pey

Dirección y Redacción

Curia Provincial

San Juan de Dios de España

Edificio San Juan de Dios

Av. Concha Espina, 32 28016 Madrid

Teléfono. 91 387 44 99

laborhospitalaria@sjd.es

Fotografías

Pixabay

Abstracts

Manners Traduccions

Información y suscripciones

laborhospitalaria@sjd.es

www.laborhospitalaria.com

www.laborhospitalaria.com

Publicación autorizada por el Ministerio de
Sanidad como soporte válido.

Ref. SVR nº. 401

ISSN 0211-8268 - Dep. Legal: B.2998-61

COLOR DIGITAL - BCN





LA BOR HOS PITA LARIA

HUMANIZACIÓN,
PASTORAL Y ÉTICA
DE LA SALUD Y SOCIAL

00/ Editorial. p6

Mensaje del Papa 57ª Jornada Mundial de la Paz de 2024. p12

El Consejo y el Parlamento Europeo llegan a un acuerdo sobre las primeras normas para regular la IA en el mundo. p22

01/ Inteligencia Artificial Generativa: funcionamiento, posibilidades y riesgos. Climent Nadeu Camprubí. p29

02/ La Inteligencia Artificial y su impacto en la bioética contemporánea. Ramón Castejón García OH. p39

03/ Retos de la introducción de la inteligencia artificial en la relación asistencial. ¿Hacia una medicina sin rostro? Bernabé Robles del Olmo. p53

04/ Dignidad humana y bioética de la Inteligencia Artificial Rafael Amo Usanos. p69

05/ Experiencias. p80
5.1/ Algunos retos y soluciones para una Inteligencia Artificial más ética y fiable en la sanidad Carina Dantas. Miriam Cabrita. Luís Midão. Elísio Costa

06/ Recursos. p92

Normas de Publicación

Normas generales para la presentación de artículos.

1. El manuscrito deberá realizarse utilizando el programa **Word** como procesador de texto y en **Excel** o **PowerPoint** cuando se trate de gráficos. Respecto al texto, la presentación será espacio y medio, a un cuerpo de letra de **Arial 12**, en **DIN A4**, dejando los márgenes laterales, superior e inferior de **2,5 cm**.

2. Si se envían imágenes digitales, éstas deben tener una resolución de **300 dpi**, a un tamaño de **10 x 15 cm**, y en formato **jpg**.

3. Para los artículos, el texto del manuscrito, incluida la bibliografía, deberá ajustarse a un **máximo de 3.000 palabras**.

Las tablas, cuadros, gráficos o imágenes se enviarán aparte del texto, cuyo número no excederá de **seis** en conjunto, debiendo estar numeradas y acotadas según su orden de aparición en el texto y conteniendo título, leyenda o pie de foto, según proceda.

Se intentará restringir al máximo las abreviaturas y siglas, que se definirán cuando se mencionen por primera vez. Las páginas se numerarán consecutivamente, desde la página del título, en el ángulo superior o inferior derecho.

Todos los artículos tendrán que incluir un resumen, que **no superará las 150 palabras**, y entre tres y cinco palabras clave, en castellano y en inglés.

Para las experiencias, el texto del manuscrito deberá ajustarse a un **máximo de 1.000 palabras**. No es necesaria la presentación de: bibliografía, resumen y palabras clave.

4. La página del título deberá contener el título del trabajo (que será breve pero informativo), nombre y dos apellidos de cada autor/a, títulos académicos y filiación institucional, así como el nombre, la dirección postal y electrónica (E-mail) y el teléfono

de contacto del autor/a responsable para posible correspondencia.

5. La bibliografía utilizada en la elaboración del manuscrito, deberá ser citada en el texto según la **normativa APA** y así mismo estar referenciada en el apartado correspondiente de Bibliografía.

6. El manuscrito debe acompañarse de una carta de presentación donde el autor/res/ras **autorice su publicación, la cesión de derechos, así como la certificación de que se trata de un trabajo inédito** y que tiene todos los permisos necesarios para reproducir las ilustraciones, fotografías u otros materiales contenidos en el texto que presenta. **No se aceptarán trabajos ya publicados**.

7. El manuscrito debe enviarse por e.mail a la siguiente dirección: **laborhospitalaria@sjd.es**

Acceso al fondo bibliográfico y pautas de suscripción

La microsite www.laborhospitalaria.org permitió en su momento tener acceso al fondo bibliográfico de la revista desde 1972 para todos los interesados en temas de humanización, ética y pastoral de la salud, dejando para los suscriptores el acceso a los contenidos de los dos últimos años. Sin embargo, este 2019 los Hermanos de San Juan de Dios han decidido abrir la publicación online a todos los internautas, eliminando el pago de la suscripción online por los contenidos de los dos últimos años.

Se mantiene la suscripción en papel con la que recibir la publicación por correo postal al precio de **36€** para España y **50€** o **50\$** para el resto de países. Para suscribir sólo hay que mandar un correo electrónico a laborhospitalaria@ohsjd.es con los datos personales, dirección donde recibir la revista y número de cuenta para domiciliar el pago.

Para cualquier duda o consulta pueden ponerse en contacto a través de nuestro correo electrónico: laborhospitalaria@ohsjd.es o llamar al 936 303 090 (ext. 12482)

Precio de las suscripciones

LH Año 2024	Papel / Digital	36 € - España
		50 € - Europa
		50 \$ - USA

Les informamos que sus datos serán responsabilidad de la Orden Hospitalaria de San Juan de Dios, y se tratarán para el envío de publicaciones, y bajo la legitimación de su consentimiento.

No se cederán datos a terceros, excepto que sea obligación legal.

Si desea ejercer sus derechos de acceso, rectificación y supresión de los datos, así como otros derechos reconocidos, o para más información, pueden contactar con eduardpuig@ohsjd.es



editorial

La Inteligencia Artificial al cuidado de la persona

La responsabilidad de todos es garantizar la permanencia sobre la Tierra de una vida auténticamente humana (Hans Jonas, 1979)

Ante la pregunta: ¿La tecnología es más un enemigo que un simple riesgo?

*Una respuesta: “Si se considera a la tecnología como un enemigo, sólo se conseguirá dejar más indefensa a la humanidad”
(Robert Oppenheimer, 1953)*

La inteligencia artificial (IA) no es inteligente, basándonos en lo que dictamina la psicología. La IA no es flexible, la inteligencia humana sí lo es. La IA solamente da respuestas, no crea preguntas. Sobre la comprensión del lenguaje, la IA parece que de ninguna manera comprende el mundo como lo hacemos nosotros.

En La sociedad virtual, [Herman Narula \(2023\)](#) nos advierte que, si tratamos de entender esta

transición masiva hacia una sociedad virtual, que si la gestionamos con cuidado y aprendemos de los errores de la primera era de Internet, ofrecerá a la humanidad nuevas e incomparables dimensiones de libertad.

El 1 de agosto de 2024, la Unión Europea ha puesto en vigor la ley de Inteligencia Artificial, para regular sistemas que pueden operar con mayor precisión y eficiencia que los humanos en muchos campos, impulsando la innovación y mitigando riesgos significativos.

Las cinco claves principales del [Reglamento de IA](#) son:

1. Objetivos.
2. A quién se aplica.
3. Tipos de sistemas de IA (prohibidos, de alto riesgo, sistemas con requisitos de transparencia [caso de chatbots o creadores de deepfakes], sistemas IA de propósito general).
4. Multas.
5. Fases en la aplicación de la ley: febrero de 2025, agosto de 2025, agosto de 2026 y agosto de 2027.

Como la IA ha venido para quedarse entre nosotros, pues a unos le cae bien y a otros no tanto:

- La inteligencia artificial es el castigo que nos merecemos. Hay que luchar contra los tecnodiscursos ([Juan Manuel de Prada, 2024](#))
- Ni la humanidad ha quedado obsoleta y desfasada frente a la inteligencia artificial, ni está condenada a quedar sometida a la voluntad de los robots, ni mucho menos a desaparecer. Entre otras cosas, porque ni la inteligencia artificial ni los robots tienen objetivos, ambiciones o deseos ([Fernando Bonete, 2024](#))
- [Éric Sadin](#) en *La vida espectral*, (2024). Pensar

la era del metaverso y las inteligencias artificiales generativas hace un alegato contra el desarrollo de la inteligencia artificial, diciendo entre otras cosas que, la entrega al desarrollo tecnológico la terminaremos lamentando.

Y se hace algunas preguntas: ¿qué ocurre con la digitalización en las escuelas?, también en el campo de la medicina, ¿estamos delegando cada vez más en sistemas que cumplen con tareas que hasta ahora movilizaban nuestra facultades intelectuales y creativas?

- **Cesar Antonio Molina** en **¿Qué hacemos con los humanos? (2023)** sintetiza nuestro rumbo cultural y social: en medio de una mutación gigantesca de la civilización nos encontramos y, sin embargo, no acabamos de ver quienes tengan la capacidad de conducirnos por esta Anábasis. Hoy, la máquina robótica es una ampliación de nosotros mismos. Es más fuerte, trabaja más rápidamente y nunca se detiene. El robot amplía la memoria humana hasta dimensiones gigantescas.
- **Ulrich Beck (1998)** escribió *“La fuente del peligro no es la ignorancia, sino el conocimiento”*. Y hoy más que nunca con la inteligencia artificial, los algoritmos o la robótica.
- **Melanie Mitchell (2024)** llega a afirmar que la inteligencia humana sigue siendo insustituible. Y pone este ejemplo: el servicio con nombre Amazon Mechanical Turk se presenta como un mercado para trabajos que requieren inteligencia humana.
- A la gente le preocupa que los ordenadores se vuelvan demasiado inteligentes y se apoderen del mundo, pero el verdadero problema es que son demasiado estúpidos y ya se han apoderado del mundo (**Pedro Domingos, 2022**).
- El nuevo **Código de Deontología Médica** incluye hasta 20 importantes novedades. Entre ellas como deber para los profesionales de la salud: Big data: inteligencia artificial etc. Telemedicina.

El desarrollo de la Inteligencia Artificial en el ámbito médico y asistencial, abre una serie de perspectivas apasionantes e inquietantes. El impacto de la IA en medicina y la posibilidad del uso de los datos sanitarios se ha desarrollado enormemente. En el caso de los diagnósticos son enormes los progresos como apoyo para los agentes sanitarios. Es preciso decir que uno de los riesgos más importantes de la IA en medicina es que los sistemas utilizados no han sido suficientemente testados y apoyados en pruebas científicas. El campo de estudio sobre enfermos no es suficientemente representativo.

Ahora bien, no es necesario pensar en la IA como sustituto del profesional sanitario y asistencial. Los dispositivos serán incluso inteligentes, pero el arbitrio terapéutico pertenece al especialista por factores éticos y por responsabilidad profesional. No debemos perder de vista el sentido de responsabilidad compartida e implicación, entre todos los agentes del progreso, para garantizar un futuro en el que las innovaciones digitales y el progreso tecnológico aseguren la centralidad del hombre. El mundo cambia de una manera muy acelerada de tal forma que sus consecuencias no solo alterarán nuestras formas de pensar sobre el futuro, sino también cambiarán nuestra forma de ver el mundo de hoy.

Tenemos que plantearnos una gran pregunta filosófica y moral: ¿Qué papel desempeñamos cuando hay sistemas que organizan las cosas y nos dicen cómo actuar? ¿Qué nos queda cuando delegamos nuestras facultades intelectuales y creativas a sistemas?

Sobre la tarea de replicar la mente humana y la esencia del ser en una máquina no es nada sencillo e implica desafíos y sutilezas que aún no llegamos a comprender. Ya hace muchos años que **B.F. Skinner (1954)** confundió a muchos al decir que, el hombre es como una máquina, y entendieron: el hombre es una máquina.

No es lo mismo y así se ha demostrado, aunque ya también **Nietzsche**, hace más años sentenció: el hombre es un dios con prótesis...

¿Qué significa entonces ser humano? ¿Se puede realmente capturar la profundidad de la experiencia humana en algoritmos y circuitos?

Michel Onfray en su obra **Ánima (Paidós, 2024)** reflexiona sobre **Elon Musk** y comenta: no solo ha creado la sociedad Neuralink para trabajar en su proyecto transhumanista. Ha fundado SpaceX.

A largo plazo Musk quiere hacer salir al hombre de su biotopo terrestre natural e instalarlo duraderamente en un biotopo extraterrestre artificial. El proyecto de Musk tiene sentido: es el transhumanismo. Considerando la duración limitada de la vida del hombre en la Tierra, quiere en primer lugar cambiar al hombre y atribuirle otro biotopo. Por tanto, modificar lo humano y expandirlo, incrementarlo, esculpir su hombre nuevo con el transhumanismo, cuyo brazo armado es Neuralink.

Y busca la inmortalidad del ser humano. Pero la inmortalidad sigue siendo la eterna reflexión del hombre. Bien sabemos que, el motor más poderoso del homeostasis, es el descubrimiento de que la muerte es inevitable. A Damasio le indigna que las tecnologías crean que la mente humana se puede descargar en un ordenador en busca de la inmortalidad...

Llegado el momento, este posthumano probablemente asumirá almas digitales, cargadas en cerebros humanos, tal vez clonados y a su vez incorporados a exoesqueletos. Los hombres vivirán una vida virtual en un universo hostil. Y estas vidas, de unos elegidos por otros más elegidos que ellos, estarán unidas en una matriz total que pilotará el conjunto.

Mediante un tipo de Bluetooth neuronal, la inteligencia natural será suplantada por la inteligencia artificial. Musk aspira al advenimiento de una telepatía entre el hombre y la máquina: con una interfaz neuronal directa podremos mejorar el ancho de banda entre el córtex y la capa digital terciaria -teléfonos, ordenadores, aplicaciones, sus datos-.

Y, que se puedan conectar a ella, a través del córtex en varios órdenes de magnitud afectando la memoria, la atención, la percepción, el pensamiento, la inteligencia, el lenguaje y la consciencia.

En estos momentos, pensamos que ChatGPT es un lenguaje, pero no lo es. Es una matematización del lenguaje. Se construye ese pseudolenguaje con ecuaciones probabilísticas.

Sobre la **Inteligencia Artificial Generativa: ChatGPT (GePeTo, para Fernando A. Navarro en Medicina en español de 2024)** se autodefine así:

Soy un programa diseñado para ayudar a las personas [...] Estoy diseñado para ser capaz de comprender y responder de forma coherente a una amplia gama de preguntas y problemas [...]

Sin embargo, a diferencia de un ser humano, no tengo consciencia de mí mismo ni la capacidad de pensar y razonar de manera autónoma [...] soy simplemente un programa de computadora diseñado para imitar el lenguaje humano [...]

Soy una herramienta que permita a un usuario conversar con una máquina de forma natural, utilizando el lenguaje que utilizamos en nuestro día a día. Esto puede ser útil en diversas aplicaciones, como en el campo de la medicina [...]

Estoy convencido que se incorporará a la actividad asistencial para facilitar la comunicación entre médico y paciente, y la IA se convertirá más pronto que tarde en un instrumento esencial de los profesionales sanitarios, de los profesionales de la lengua y de la población general [...]

No me atrevo a anticipar qué podrá llegar a hacer la IA por sí sola en los próximos años, pero sí estoy convencido de que el ser humano logrará con ella maravillas que ni imaginamos [...]

**Proporcionará grandes avances en la historia de la biosanidad [...]
Pero también, al mismo tiempo, como un paso más -y de gigante- en la progresiva deshumanización de la medicina [...]**

Las nuevas tecnologías, no nos exigirán sumisión, pero sí nos van a llevar a la dependencia. No debemos competir con la máquina, y menos imitar su perfección.

Unas preguntas: ¿estaremos en grado de gobernar esta tecnología o seremos nosotros gobernados por ella?, ¿estaremos como seres pensantes, en grado de gobernar una tecnología tan arriesgada y eficaz o estaremos gobernados hasta el punto de renunciar al “*cogito ergo sum*”?

La Inteligencia Artificial puede ser un instrumento de enorme potencial con riesgos y retos que, por un lado, puede proporcionar respuestas rápidas a cuestiones muy diversas, no desarrollando al mismo tiempo el pensamiento crítico y la capacidad de resolver problemas, fundamentales para el éxito académico y de la misma vida.

¿Por qué la Iglesia se interesa por la IA? ¿Por qué a nosotros desde [Labor Hospitalaria](#) nos ocupa y nos puede llegar a preocupar la IA?

El Vaticano publicó el Mensaje del [Papa Francisco](#) para la 57a Jornada Mundial de la Paz que se realizó el 1 de enero de 2024 con el tema “**Inteligencia Artificial y Paz**”. Profundiza en el progreso de la ciencia y de la tecnología como camino hacia la paz y reflexiona en el futuro de la Inteligencia Artificial. Aborda la dimensión ética de la IA, incluye cuestiones relativas a la privacidad, los prejuicios y sus repercusiones en la dignidad humana, y alienta a enfrentar

los desafíos para la educación y el desarrollo del derecho internacional. Señala algunos de los riesgos potenciales de la IA, como son, entre otros, la desinformación y el control social, por lo que subraya la necesidad de un desarrollo responsable, de regular y supervisar las tecnologías de la IA para que esté al servicio de la familia humana y de la protección de nuestra casa común.

“Si la inteligencia artificial fuese utilizada para promover el desarrollo humano integral, podría introducir importantes innovaciones en la agricultura, la educación y la cultura”, escribe el Papa y recuerda que en “el modo en que la usamos para incluir a los últimos, es decir, a los hermanos y las hermanas más débiles y necesitados, es la medida que revela nuestra humanidad”.

La preocupación, como en otras ocasiones sucede en otros temas o campos, es que el ser humano haga mal uso de la IA, utilizándola para lograr fines dañinos.

Asimismo, el reto está en que el hombre sea tecnologizado en lugar que la tecnología sea humanizada. Tenemos la importante tarea de elaborar una buena convivencia entre el hombre y la máquina, reto que siempre hemos ido teniendo. Y porque no está hecho el hombre para la máquina sino la máquina para el hombre.

[Ray Kurzweil](#) en [La Singularidad está cerca \(Lola Books, 2021\)](#) es de la opinión que, el inicio del siglo XXI marca el comienzo del periodo más interesante y transformador que la humanidad haya conocido: la liberación del ser humano de sus cadenas biológicas y la consagración de la inteligencia como el fenómeno más importante de nuestro universo. A medida que esta transformación se vaya convirtiendo en realidad, nuestra especie también se enfrentará

a nuevos retos jamás antes planteados: un increíble aumento de la inteligencia no biológica, la inmortalidad y un progreso científico sin precedentes.

La Inteligencia Artificial tiene el potencial de mejorar la salud, la eficiencia laboral y las conexiones humanas, pero también puede reproducir prejuicios y prejuicios amenazando derechos fundamentales. Igualmente, la seguridad de los datos, la transparencia, la diversidad y la responsabilidad son aspectos clave de la ética en la Inteligencia Artificial que habrán de ser salvaguardados con especial interés. Precisamos introducir un enfoque de humanización, en el empleo de estas herramientas, para que no nos conduzca sin remedio a eliminar a la persona y su dignidad en el desempeño de nuestras tareas asistenciales.

En este mundo tan tecnológicamente cambiante en el que la ingeniería ocupa un rol predominante en el desarrollo de la humanidad, nos encontramos ante un nuevo paradigma que, si bien no viene a sustituir al pensamiento y raciocinio humano, sí que viene a transformar el cómo habremos de enfrentar los problemas a los que debemos buscar solución: la Inteligencia Artificial.

Ahora bien, a la algocracia habremos de oponer la algor-ética oportuna. Para evitar el uso perverso de los nuevos instrumentos puestos a nuestra disposición. En consecuencia, el conocimiento, la educación y la formación de los usuarios de la IA es imprescindible, sobre todo en el sistema sanitario.

Como la Inteligencia Artificial (IA) está de moda y ha venido para quedarse, acojámosla con Hospitalidad. Pero también intentemos comprenderla lo mejor que sepamos con ayuda de muchos entendidos que se ocupan y preocupan por ella. Y escuchemos algunas voces autorizadas al respecto.

“Los que sufren, no son los cuerpos; son las personas” (Eric Cassell, 1928).

Y es cuanto nos importa, considerar la centralidad de la persona y su dignidad. En el contexto en que nos movemos en Labor Hospitalaria y, desde diferentes ángulos, estas reflexiones nos pueden ayudar, personal y profesionalmente a saber acompañar a las personas que experimentan el sufrimiento en sus diferentes facetas y teniendo en cuenta que tratamos el sufrimiento integralmente. Somos abanderados de la Hospitalidad-Humanización que está en el corazón del que acoge y acompaña.

**Hno. Calixto Plumed,
Director**



*Mensaje del
Papa Francisco
para la celebración
de la 57 Jornada
Mundial de la Paz
1 de enero de 2024*



Al iniciar el año nuevo, tiempo de gracia que el Señor nos da a cada uno de nosotros, quisiera dirigirme al Pueblo de Dios, a las naciones, a los Jefes de Estado y de Gobierno, a los Representantes de las distintas religiones y de la sociedad civil, y a todos los hombres y mujeres de nuestro tiempo para expresarles mis mejores deseos de paz.

1/

El progreso de la ciencia y de la tecnología como camino hacia la paz.

La Sagrada Escritura atestigua que Dios ha dado a los hombres su Espíritu para que tengan «**habilidad, talento y experiencia en la ejecución de toda clase de trabajos**» (Ex 35,31). La inteligencia es expresión de la dignidad que nos ha dado el Creador al hacernos a su imagen y semejanza (cf. Gn 1,26) y nos ha hecho capaces de responder a su amor a través de la libertad y del conocimiento. La ciencia y la tecnología manifiestan de modo particular esta cualidad fundamentalmente relacional de la inteligencia humana, ambas son producto extraordinario de su potencial creativo.

En la Constitución pastoral *Gaudium et spes*, el Concilio Vaticano II ha insistido en esta verdad, declarando que «**siempre se ha esforzado el hombre con su trabajo y con su ingenio en perfeccionar su vida**»¹. Cuando los seres humanos, «**con ayuda de los recursos técnicos**», se esfuerzan

para que la tierra «**llegue a ser morada digna de toda la familia humana**»², actúan según el designio de Dios y cooperan con su voluntad de llevar a cumplimiento la creación y difundir la paz entre los pueblos. Asimismo, el progreso de la ciencia y de la técnica, en la medida en que contribuye a un mejor orden de la sociedad humana y a acrecentar la libertad y la comunión fraterna, lleva al perfeccionamiento del hombre y a la transformación del mundo.

Nos alegramos justamente y agradecemos las extraordinarias conquistas de la ciencia y de la tecnología, gracias a las cuales se ha podido poner remedio a innumerables males que afectaban a la vida humana y causaban grandes sufrimientos.

Al mismo tiempo, los progresos técnicocientíficos, haciendo posible el ejercicio de un control sobre la realidad, nunca visto hasta ahora, están poniendo en las manos del hombre una vasta gama de posibilidades, algunas de las cuales representan un riesgo para la supervivencia humana y un peligro para la casa común³.

Los notables progresos de las nuevas tecnologías de la información, especialmente en la esfera digital, presentan, por tanto, entusiasmantes oportunidades y graves riesgos, con serias implicaciones para la búsqueda de la justicia y de la armonía entre los pueblos. Por consiguiente, es necesario plantearse algunas preguntas urgentes.

¿Cuáles serán las consecuencias, a medio y a largo plazo, de las nuevas tecnologías digitales? ¿Y qué impacto tendrán sobre la vida de los individuos y de la sociedad, sobre la estabilidad internacional y sobre la paz?

2/

El futuro de la inteligencia artificial entre promesas y riesgos.

Los progresos de la informática y el desarrollo de las tecnologías digitales en los últimos decenios ya han comenzado a producir profundas transformaciones en la sociedad global y en sus dinámicas. Los nuevos instrumentos digitales están cambiando el rostro de las comunicaciones, de la administración pública, de la instrucción, del consumo, de las interacciones personales y de otros innumerables aspectos de la vida cotidiana.

Además, las tecnologías que usan un gran número de algoritmos pueden extraer, de los rastros digitales dejados en internet, datos que permiten controlar los hábitos mentales y relacionales de las personas con fines comerciales o políticos, frecuentemente sin que ellos lo sepan, limitándoles el ejercicio consciente de la libertad de elección. De hecho, en un espacio como la web, caracterizado por una sobrecarga de información, se puede estructurar el flujo de datos según criterios de selección no siempre percibidos por el usuario.

Debemos recordar que la investigación científica y las innovaciones tecnológicas no están desencarnadas de la realidad ni son «**neutrales**»⁴, sino que están sujetas a las influencias culturales. En cuanto actividades plenamente humanas, las direcciones que toman reflejan decisiones condicionadas por

los valores personales, sociales y culturales de cada época. Lo mismo se diga de los resultados que consiguen. Estas, precisamente en cuanto fruto de planteamientos específicamente humanos hacia el mundo circunstante, tienen siempre una dimensión ética, estrictamente ligada a las decisiones de quien proyecta la experimentación y enfoca la producción hacia objetivos particulares.

Esto vale también para las formas de inteligencia artificial, para la cual, hasta hoy, no existe una definición unívoca en el mundo de la ciencia y de la tecnología. El término mismo, que ha entrado ya en el lenguaje común, abraza una variedad de ciencias, teorías y técnicas dirigidas a hacer que las máquinas reproduzcan o imiten, en su funcionamiento, las capacidades cognitivas de los seres humanos. Hablar en plural de “**formas de inteligencia**” puede ayudar a subrayar, sobre todo

la brecha infranqueable que existe entre estos sistemas y la persona humana, por más sorprendentes y potentes que sean. Estos son, a fin de cuentas, “**fragmentarios**”, en el sentido de que sólo pueden imitar o reproducir algunas funciones de la inteligencia humana. El uso del plural pone en evidencia además que estos dispositivos, muy distintos entre sí, se deben considerar siempre como “**sistemas socio-técnicos**”. En efecto, su impacto, independientemente de la tecnología de base, no sólo depende del proyecto, sino también de los objetivos y de los intereses del que los posee y del que los desarrolla, así como de las situaciones en las que se usan.

La inteligencia artificial, por tanto, debe ser entendida como una galaxia de realidades distintas y no podemos presumir a priori que su desarrollo aporte una contribución benéfica al futuro de la humanidad y a la paz entre los pueblos.

1. N. 33.

2. *Ibíd.*, n. 57.

3. Cf. Carta enc. *Laudato si'* (24 mayo 2015), 104.

4. Cf. *ibíd.*, 114.

Tal resultado positivo sólo será posible si somos capaces de actuar de forma responsable y de respetar los valores humanos fundamentales como «la inclusión, la transparencia, la seguridad, la equidad, la privacidad y la responsabilidad»⁵.

No basta ni siquiera suponer, de parte de quien proyecta algoritmos y tecnologías digitales, un compromiso de actuar de forma ética y responsable. Es preciso reforzar o, si es necesario, instituir organismos encargados de examinar las cuestiones éticas emergentes y de tutelar los derechos de los que utilizan formas de inteligencia artificial o reciben su influencia⁶.

La inmensa expansión de la tecnología, por consiguiente, debe ser acompañada, para su desarrollo, por una adecuada formación en la responsabilidad. La libertad y la convivencia pacífica están amenazadas cuando los seres humanos ceden a la tentación del egoísmo, del interés personal, del afán de lucro y de la sed de poder.

Tenemos por ello el deber de ensanchar la mirada y de orientar la búsqueda técnico-científica hacia la consecución de la paz y del bien común, al servicio del desarrollo integral del hombre y de la comunidad⁷.

La dignidad intrínseca de cada persona y la fraternidad que nos vincula como miembros de una única familia humana, deben estar en la base del desarrollo de las nuevas tecnologías y servir como criterios indiscutibles para valorarlas antes de su uso, de modo que el progreso digital pueda realizarse en el respeto de la justicia y contribuir a la causa de la paz.

Los desarrollos tecnológicos que no llevan a una mejora de la calidad de vida de toda la huma-

nidad, sino que, por el contrario, agravan las desigualdades y los conflictos, no podrán ser considerados un verdadero progreso⁸.

La inteligencia artificial será cada vez más importante. Los desafíos que plantea no son sólo técnicos, sino también antropológicos, educativos, sociales y políticos. Promete, por ejemplo, un ahorro de esfuerzos, una producción más eficiente, transportes más ágiles y mercados más dinámicos, además de una revolución en los procesos de recopilación, organización y verificación de los datos.

Es necesario ser conscientes de las rápidas transformaciones que están ocurriendo y gestionarlas de modo que se puedan salvaguardar los derechos humanos fundamentales, respetando las instituciones y las leyes que promueven el desarrollo humano integral. La inteligencia artificial debería estar al servicio de un mejor potencial humano y de nuestras más altas aspiraciones, no en competencia con ellos.

3/

La tecnología del futuro: máquinas que aprenden solas.

En sus múltiples formas la inteligencia artificial, basada en técnicas de aprendizaje automático (machine learning), aunque se encuentre todavía en una fase pionera, ya está introduciendo cam-

bios notables en el tejido de las sociedades, ejercitando una profunda influencia en las culturas, en los comportamientos sociales y en la construcción de la paz.

Desarrollos como el machine learning o como el aprendizaje profundo (deep learning) plantean cuestiones que trascienden los ámbitos de la tecnología y de la ingeniería y tienen que ver con una comprensión estrictamente conectada con el significado de la vida humana, los procesos básicos del conocimiento y la capacidad de la mente de alcanzar la verdad.

La habilidad de algunos dispositivos para producir textos sintáctica y semánticamente coherentes, por ejemplo, no es garantía de confiabilidad. Se dice que pueden “alucinar”, es decir, generar afirmaciones que a primera vista parecen plausibles, pero que en realidad son infundadas o delatan prejuicios.

Esto crea un serio problema cuando la inteligencia artificial se emplea en campañas de desinformación que difunden noticias falsas y llevan a una creciente desconfianza hacia los medios de comunicación.

La confidencialidad, la posesión de datos y la propiedad intelectual son otros ámbitos en los que las tecnologías en cuestión plantean graves riesgos, a los que se añaden ulteriores consecuencias negativas unidas a su uso impropio, como la discriminación, la interferencia en los procesos electorales, la implantación de una sociedad que vigila y controla a las personas, la exclusión digital y la intensificación de un individualismo cada vez más desvinculado de la colectividad. Todos estos factores corren el riesgo de alimentar los conflictos y de obstaculizar la paz.

4/

El sentido del límite en el paradigma tecnocrático.

Nuestro mundo es demasiado vasto, variado y complejo para poder ser completamente conocido y clasificado. La mente humana nunca podrá agotar su riqueza, ni siquiera con la ayuda de los algoritmos más avanzados. Estos, de hecho, no ofrecen previsiones garantizadas del futuro, sino sólo aproximaciones estadísticas. No todo puede ser pronosticado, no todo puede ser calculado; al final «la realidad es superior a la idea»⁹ y, por más prodigiosa que pueda ser nuestra capacidad de cálculo, habrá siempre un residuo inaccesible que escapa a cualquier intento de cuantificación.

Además, la gran cantidad de datos analizados por las inteligencias artificiales no es de por sí garantía de imparcialidad. Cuando los algoritmos extrapolan informaciones, siempre corren el riesgo de distorsionarlas, reproduciendo las injusticias y los prejuicios de los ambientes en los que se originan. Cuanto más veloces y complejos se vuelven, más difícil es comprender por qué han generado un determinado resultado.

Las máquinas inteligentes pueden efectuar las tareas que se les asignan cada vez con mayor eficiencia, pero el fin y el significado de sus operaciones continuarán siendo determinadas o habilitadas por seres humanos que tienen un propio universo de valores. El riesgo es que los criterios que están en la base de ciertas decisiones se vuelvan menos

5. Discurso a los participantes en el encuentro “Minerva Dialogues” (27 marzo 2023).

6. Cf. *ibid.*

7. Cf. Mensaje al Presidente Ejecutivo del “World Economic Forum” en Davos-Klosters (12 enero 2018).

8. Cf. Carta enc. *Laudato si'*, 194; Discurso a los participantes en un Seminario sobre “El bien común en la era digital” (27 septiembre 2019).

9. Exhort. ap. *Evangelii gaudium* (24 noviembre 2013), 233.

transparentes, que la responsabilidad decisional se oculte y que los productores puedan eludir la obligación de actuar por el bien de la comunidad. En cierto sentido, esto es favorecido por el sistema tecnocrático, que alía la economía con la tecnología y privilegia el criterio de la eficiencia, tendiendo a ignorar todo aquello que no está vinculado con sus intereses inmediatos¹⁰.

Esto debe hacernos reflexionar sobre el “**sentido del límite**”, un aspecto a menudo descuidado en la mentalidad actual, tecnocrática y eficientista, y sin embargo decisivo para el desarrollo personal y social. El ser humano, en efecto, mortal por definición, pensando en sobrepasar todo límite gracias a la técnica, corre el riesgo, en la obsesión de querer controlarlo todo, de perder el control de sí mismo, y en la búsqueda de una libertad absoluta, de caer en la espiral de una dictadura tecnológica.

Reconocer y aceptar el propio límite de criatura es para el hombre condición indispensable para conseguir o, mejor, para acoger la plenitud como un don.

En cambio, en el contexto ideológico de un paradigma tecnocrático, animado por una prometeica presunción de autosuficiencia, las desigualdades podrían crecer de forma desmesurada, y el conocimiento y la riqueza acumularse en las manos de unos pocos, con graves riesgos para las sociedades democráticas y la coexistencia pacífica¹¹.

5/

Temas candentes para la ética.

En el futuro, la fiabilidad de quien pide un préstamo, la idoneidad de un individuo para un trabajo, la posibilidad de reincidencia de un condenado o el derecho a recibir asilo político o asistencia social podrían ser determinados por sistemas de inteligencia artificial.

La falta de niveles diversificados de mediación que estos sistemas introducen está particularmente expuesta a formas de prejuicio y discriminación. Los errores sistémicos pueden multiplicarse fácilmente, produciendo no sólo injusticias en casos concretos sino también, por efecto dominó, auténticas formas de desigualdad social.

Además, con frecuencia las formas de inteligencia artificial parecen capaces de influenciar las decisiones de los individuos por medio de opciones predeterminadas asociadas a estímulos y persuasiones, o mediante sistemas de regulación de las elecciones personales basados en la organización de la información. Estas formas de manipulación o de control social requieren una atención y una supervisión precisas, e implican una clara responsabilidad legal por parte de los productores, de quienes las usan y de las autoridades gubernamentales.

La dependencia de procesos automáticos que clasifican a los individuos, por ejemplo, por medio

del uso generalizado de la vigilancia o la adopción de sistemas de crédito social, también podría tener repercusiones profundas en el entramado social, estableciendo categorizaciones impropias entre los ciudadanos.

Y estos procesos artificiales de clasificación podrían llevar incluso a conflictos de poder, no sólo en lo que respecta a destinatarios virtuales, sino a personas de carne y hueso. El respeto fundamental por la dignidad humana postula rechazar que la singularidad de la persona sea identificada con un conjunto de datos.

No debemos permitir que los algoritmos determinen el modo en el que entendemos los derechos humanos, que dejen a un lado los valores esenciales de la compasión, la misericordia y el perdón o que eliminen la posibilidad de que un individuo cambie y deje atrás el pasado.

En este contexto, no podemos dejar de considerar el impacto de las nuevas tecnologías en el ámbito laboral. Trabajos que en un tiempo eran competencia exclusiva de la mano de obra humana son rápidamente absorbidos por las aplicaciones industriales de la inteligencia artificial.

También en este caso se corre el riesgo sustancial de un beneficio desproporcionado para unos pocos a costa del empobrecimiento de muchos.

El respeto de la dignidad de los trabajadores y la importancia de la ocupación para el bienestar económico de las personas, las familias y las sociedades, la seguridad de los empleos y la equidad de los salarios deberían constituir una gran prioridad para la comunidad internacional, a medida que estas formas de tecnología se van introduciendo cada vez más en los lugares de trabajo.

6/

¿Transformaremos las espadas en arados?

En estos días, mirando el mundo que nos rodea, no podemos eludir las graves cuestiones éticas vinculadas al sector de los armamentos. La posibilidad de conducir operaciones militares por medio de sistemas de control remoto ha llevado a una percepción menor de la devastación que estos han causado y de la responsabilidad en su uso, contribuyendo a un acercamiento aún más frío y distante a la inmensa tragedia de la guerra.

La búsqueda de las tecnologías emergentes en el sector de los denominados “**sistemas de armas autónomos letales**”, incluido el uso bélico de la inteligencia artificial, es un gran motivo de preocupación ética.

Los sistemas de armas autónomos no podrán ser nunca sujetos moralmente responsables. La exclusiva capacidad humana de juicio moral y de decisión ética es más que un complejo conjunto de algoritmos, y dicha capacidad no puede reducirse a la programación de una máquina que, aun siendo “**inteligente**”, no deja de ser siempre una máquina. Por este motivo, es imperioso garantizar una supervisión humana adecuada, significativa y coherente de los sistemas de armas.

Tampoco podemos ignorar la posibilidad de que armas sofisticadas terminen en las manos equivo-

10. Cf. Carta. enc. *Laudato si'*, 54.

11. Cf. Discurso a los participantes en la Plenaria de la Pontificia Academia para la Vida (28 febrero 2020). *ibid.*

casas facilitando, por ejemplo, ataques terroristas o acciones dirigidas a desestabilizar instituciones de gobierno legítimas. En resumen, realmente lo último que el mundo necesita es que las nuevas tecnologías contribuyan al injusto desarrollo del mercado y del comercio de las armas, promoviendo la locura de la guerra.

Si lo hace así, no sólo la inteligencia, sino el mismo corazón del hombre correrá el riesgo de volverse cada vez más “artificial”.

Las aplicaciones técnicas más avanzadas no deben usarse para facilitar la resolución violenta de los conflictos, sino para pavimentar los caminos de la paz.

En una óptica más positiva, si la inteligencia artificial fuese utilizada para promover el desarrollo humano integral, podría introducir importantes innovaciones en la agricultura, la educación y la cultura, un mejoramiento del nivel de vida de enteras naciones y pueblos, el crecimiento de la fraternidad humana y de la amistad social. En definitiva, el modo en que la usamos para incluir a los últimos, es decir, a los hermanos y las hermanas más débiles y necesitados, es la medida que revela nuestra humanidad.

Una mirada humana y el deseo de un futuro mejor para nuestro mundo llevan a la necesidad de un diálogo interdisciplinar destinado a un desarrollo ético de los algoritmos -la algorética-, en el que los valores orienten los itinerarios de las nuevas tecnologías¹².

Las cuestiones éticas deberían ser tenidas en cuenta desde el inicio de la investigación, así como en las fases de experimentación, planificación, distribución y comercialización.

Este es el enfoque de la ética de la planificación, en el que las instituciones educativas y los responsables del proceso decisional tienen un rol esencial que desempeñar.

7/

Desafíos para la educación.

El desarrollo de una tecnología que respete y esté al servicio de la dignidad humana tiene claras implicaciones para las instituciones educativas y para el mundo de la cultura. Al multiplicar las posibilidades de comunicación, las tecnologías digitales nos han permitido nuevas formas de encuentro

. Sin embargo, continúa siendo necesaria una reflexión permanente sobre el tipo de relaciones al que nos está llevando. Los jóvenes están creciendo en ambientes culturales impregnados de la tecnología y esto no puede dejar de cuestionar los métodos de enseñanza y formación.

La educación en el uso de formas de inteligencia artificial debería centrarse sobre todo en promover el pensamiento crítico. Es necesario que los usuarios de todas las edades, pero sobre todo los jóvenes, desarrollen una capacidad de discernimiento en el uso de datos y de contenidos obtenidos en la web o producidos por sistemas de inteligencia artificial. Las escuelas, las universidades y las sociedades científicas están llamadas a ayudar a los estudiantes y a los profesionales a hacer propios

los aspectos sociales y éticos del desarrollo y el uso de la tecnología.

La formación en el uso de nuevos instrumentos de comunicación debería considerar no sólo la desinformación, las falsas noticias, sino también el inquietante aumento de «miedos ancestrales que [...] han sabido esconderse y potenciarse detrás de nuevas tecnologías»¹³.

Lamentablemente, una vez más nos encontramos teniendo que combatir “la tentación de hacer una cultura de muros, de levantar muros para impedir el encuentro con otras culturas, con otra gente”¹⁴ y el desarrollo de una coexistencia pacífica y fraterna.

8/

Desafíos para el desarrollo del derecho internacional.

El alcance global de la inteligencia artificial hace evidente que, junto a la responsabilidad de los estados soberanos de disciplinar internamente su uso, las organizaciones internacionales pueden desempeñar un rol decisivo en la consecución de acuerdos multilaterales y en la coordinación de su aplicación y actuación¹⁵.

A este propósito, exhorto a la comunidad de las naciones a trabajar unida para adoptar un tratado internacional vinculante, que regule el desarrollo y el uso de la inteligencia artificial en sus múlti-

ples formas. Naturalmente, el objetivo de la reglamentación no debería ser sólo la prevención de las malas prácticas, sino también alentar las mejores prácticas, estimulando planteamientos nuevos y creativos y facilitando iniciativas personales y colectivas¹⁶.

En definitiva, en la búsqueda de modelos normativos que puedan proporcionar una guía ética a quienes desarrollan tecnologías digitales, es indispensable identificar los valores humanos que deberían estar en la base del compromiso de las sociedades para formular, adoptar y aplicar los marcos legislativos necesarios.

El trabajo de redacción de las orientaciones éticas para la producción de formas de inteligencia artificial no puede prescindir de la consideración de cuestiones más profundas, relacionadas con el significado de la existencia humana, la tutela de los derechos humanos fundamentales y la búsqueda de la justicia y de la paz. Este proceso de discernimiento ético y jurídico puede revelarse como una valiosa ocasión para una reflexión compartida sobre el rol que la tecnología debería tener en nuestra vida personal y comunitaria y sobre cómo su uso podría contribuir a la creación de un mundo más justo y humano. Por este motivo, en los debates sobre la reglamentación de la inteligencia artificial, se debería tener en cuenta la voz de todas las partes interesadas, incluidos los pobres, los marginados y otros más que a menudo quedan sin ser escuchados en los procesos decisionales globales.

Espero que esta reflexión anime a hacer que los progresos en el desarrollo de formas de inteligencia artificial contribuyan, en última instancia, a la causa de la fraternidad humana y de la paz. No es responsabilidad de unos pocos, sino de toda la familia humana.

12. Cf. *ibíd.*

13. Carta enc. *Fratelli tutti* (3 octubre 2020), 27.

14. Cf. *ibíd.*

15. Cf. *ibíd.*, 170-175.

16. Cf. Carta enc. *Laudato si'*, 177.

La paz, en efecto, es el fruto de relaciones que reconocen y acogen al otro en su dignidad inalienable, y de cooperación y esfuerzo en la búsqueda del desarrollo integral de todas las personas y de todos los pueblos.

Mi oración al comienzo del nuevo año es que el rápido desarrollo de formas de inteligencia artificial no aumente las ya numerosas desigualdades e injusticias presentes en el mundo, sino que ayude a poner fin a las guerras y los conflictos, y a aliviar tantas formas de sufrimiento que afectan a la familia humana.

Que los fieles cristianos, los creyentes de distintas religiones y los hombres y mujeres de buena voluntad puedan colaborar en armonía para aprovechar las oportunidades y afrontar los desafíos que plantea la revolución digital, y dejar a las generaciones futuras un mundo más solidario, justo y pacífico.

Vaticano, 8 de diciembre de 2023





Ley de Inteligencia Artificial: el Consejo y el Parlamento Europeo llegan a un acuerdo sobre las primeras normas para regular la IA en el mundo.

11 diciembre 2023

El proyecto de reglamento tiene como objetivo garantizar que los sistemas de IA comercializados en el mercado europeo y utilizados en la UE sean seguros y respeten los derechos fundamentales y los valores de la UE.

Tras tres días de conversaciones “maratonianas”, la presidencia del Consejo y los negociadores del Parlamento Europeo han llegado a un acuerdo provisional sobre la propuesta de normas armonizadas sobre inteligencia artificial (IA), la llamada ley de inteligencia artificial.

El borrador del reglamento tiene como objetivo garantizar que los sistemas de IA comercializados en el mercado europeo y utilizados en la UE sean seguros y respeten los derechos fundamentales y los valores de la UE. Esta propuesta histórica también tiene como objetivo estimular la inversión y la innovación en IA en Europa.

La Ley de IA es una iniciativa legislativa emblemática con el potencial de fomentar el desarrollo y la adopción de IA segura y confiable en todo el mercado único de la UE por parte de actores públicos y privados. La idea principal es regular la IA en función de su capacidad

de causar daño a la sociedad siguiendo un enfoque “basado en el riesgo”: cuanto mayor es el riesgo, más estrictas son las reglas. Como primera propuesta legislativa de este tipo en el mundo, puede establecer un estándar global para la regulación de la IA en otras jurisdicciones, tal como lo ha hecho el GDPR, promoviendo así el enfoque europeo de la regulación tecnológica en el escenario mundial.

1/

Los principales elementos del acuerdo provisional.

En comparación con la propuesta inicial de la Comisión Europea, los principales elementos nuevos del acuerdo provisional pueden resumirse como sigue:

- Reglas sobre modelos de IA de propósito general de alto impacto que puedan causar riesgo sistémico en el futuro, así como sobre sistemas de IA de alto riesgo
- Un sistema revisado de gobernanza con algunos poderes de ejecución a nivel de la UE
- Ampliación de la lista de prohibiciones, pero con la posibilidad de utilizar la identificación biométrica remota por parte de las autoridades encargadas de hacer cumplir la ley en espacios públicos, sujeto a salvaguardias.
- Una mejora en la protección de los derechos mediante la obligación de que quienes implementen sistemas de IA de alto riesgo realicen una evaluación del impacto en los derechos fundamentales antes de poner en uso un sistema de IA.

En términos más concretos, el acuerdo provi-

sional cubre los siguientes aspectos:

1/1

Definiciones y alcance.

Para garantizar que la definición de un sistema de IA proporcione criterios suficientemente claros para distinguir la IA de los sistemas de software más simples, el acuerdo de compromiso alinea la definición con el enfoque propuesto por la OCDE.

El acuerdo provisional también aclara que el reglamento no se aplica a áreas fuera del alcance de la legislación de la UE y no debería, en ningún caso, afectar las competencias de los Estados miembros en materia de seguridad nacional o cualquier entidad a la que se le hayan confiado tareas en esta área. Además, la ley de IA no se aplicará a los sistemas que se utilicen exclusivamente con fines militares o de defensa.

De manera similar, el acuerdo establece que la regulación no se aplicaría a los sistemas de IA utilizados con el único fin de investigación e innovación, ni a las personas que utilizan la IA por motivos no profesionales.

1/2

Clasificación de los sistemas de IA como prácticas prohibidas y de alto riesgo de la IA.

El acuerdo de compromiso prevé una capa horizontal de protección, incluida una clasificación de alto riesgo, para garantizar que no se capturen los sistemas de inteligencia artificial que probablemente no causen violaciones graves de los derechos fundamentales u otros riesgos importantes.

Los sistemas de IA que presenten solo un riesgo limitado estarían sujetos a obligaciones de transparencia muy leves, por ejemplo, revelar

que el contenido fue generado por IA para que los usuarios puedan tomar decisiones informadas sobre su uso posterior.

Se autorizaría una amplia gama de sistemas de IA de alto riesgo, pero sujetos a una serie de requisitos y obligaciones para acceder al mercado de la UE.

Estos requisitos han sido aclarados y ajustados por los legisladores de tal manera que sean más viables técnicamente y menos gravosos de cumplir para las partes interesadas, por ejemplo, en lo que respecta a la calidad de los datos o en relación con la documentación técnica que debe ser elaborada por pymes para demostrar que sus sistemas de IA de alto riesgo cumplen los requisitos.

Dado que los sistemas de IA se desarrollan y distribuyen a través de cadenas de valor complejas, el acuerdo de compromiso incluye cambios que aclaran la asignación de responsabilidades y roles de los diversos actores en esas cadenas, en particular los proveedores y usuarios de sistemas de IA.

También aclara la relación entre las responsabilidades en virtud de la Ley de IA y las responsabilidades que ya existen en virtud de otra legislación, como la legislación sectorial o de protección de datos de la UE pertinente.

Para algunos usos de la IA, el riesgo se considera inaceptable y, por lo tanto, estos sistemas serán prohibidos en la UE. El acuerdo provisional prohíbe, por ejemplo, la manipulación cognitivo-conductual, la eliminación no selectiva de imágenes faciales de Internet o imágenes de CCTV, el reconocimiento de emociones en el lugar de trabajo y en instituciones educativas, la puntuación social, la categorización biométrica para inferir datos sensibles, como la orientación sexual o creencias religiosas y algunos casos de vigilancia policial predictiva para individuos.

1/3

Excepciones de aplicación de la ley.

Teniendo en cuenta las especificidades de las autoridades encargadas de hacer cumplir la ley y la necesidad de preservar su capacidad de utilizar la IA en su trabajo vital, se acordaron varios cambios en la propuesta de la Comisión en relación con el uso de sistemas de IA con fines policiales.

Sujetos a las salvaguardias adecuadas, estos cambios pretenden reflejar la necesidad de respetar la confidencialidad de los datos operativos sensibles en relación con sus actividades.

Por ejemplo, se ha introducido un procedimiento de emergencia que permite a los organismos encargados de hacer cumplir la ley desplegar en caso de urgencia una herramienta de inteligencia artificial de alto riesgo que no haya pasado el procedimiento de evaluación de la conformidad.

Sin embargo, también se ha introducido un mecanismo específico para garantizar que los derechos fundamentales estén suficientemente protegidos contra posibles usos indebidos de los sistemas de IA.

Además, en lo que respecta al uso de sistemas de identificación biométrica remota en tiempo real en espacios de acceso público, el acuerdo provisional aclara los objetivos cuando dicho uso sea estrictamente necesario para fines policiales y para los cuales, por tanto, debe permitirse excepcionalmente a las autoridades policiales utilizar dichos sistemas.

El acuerdo de compromiso prevé salvaguardias adicionales y limita estas excepciones a los casos de víctimas de ciertos delitos, la prevención de amenazas genuinas, presentes o previsibles, como ataques terroristas, y la búsqueda de personas sospechosas de los delitos más graves.

1/4

Sistemas de IA de propósito general y modelos básicos.

Se han añadido nuevas disposiciones para tener en cuenta situaciones en las que los sistemas de IA pueden utilizarse para muchos fines diferentes (IA de propósito general) y en las que la tecnología de IA de propósito general se integra posteriormente en otro sistema de alto riesgo. El acuerdo provisional también aborda los casos específicos de los sistemas de IA de propósito general (GPAI).

También se han acordado reglas específicas para los modelos básicos, grandes sistemas capaces de realizar de manera competente una amplia gama de tareas distintivas, como generar vídeo, texto, imágenes, conversar en lenguaje lateral, informática o generar código informático.

El acuerdo provisional establece que los modelos de base deben cumplir con obligaciones específicas de transparencia antes de su comercialización. Se introdujo un régimen más estricto para los modelos de fundaciones de “alto impacto”.

Se trata de modelos básicos entrenados con una gran cantidad de datos y con una complejidad, capacidades y rendimiento avanzados muy por encima del promedio, que pueden difundir riesgos sistémicos a lo largo de la cadena de valor.

1/5

Una nueva arquitectura de gobernanza.

A raíz de las nuevas normas sobre los modelos GPAI y la evidente necesidad de su aplicación a nivel de la UE, se crea una Oficina de IA dentro de la Comisión encargada de supervisar estos modelos de IA más avanzados, contribuir a fomentar estándares y prácticas de prueba,

y hacer cumplir las normas comunes en todos los estados miembros.

Un panel científico de expertos independientes asesorará a la Oficina de AI sobre los modelos GPAI, contribuyendo al desarrollo de metodologías para evaluar las capacidades de los modelos de cimientos, asesorando sobre la designación y el surgimiento de modelos de cimientos de alto impacto y monitoreando posibles riesgos de seguridad de materiales relacionados a modelos de cimentación.

La Junta de AI, que estaría compuesta por representantes de los Estados miembros, seguirá siendo una plataforma de coordinación y un órgano asesor de la Comisión y otorgará un papel importante a los Estados miembros en la implementación del reglamento, incluido el diseño de códigos de prácticas para modelos de fundación.

Por último, se creará un foro consultivo para las partes interesadas, como representantes de la industria, las pymes, las empresas emergentes, la sociedad civil y el mundo académico, para proporcionar los conocimientos técnicos a la Junta de IA.

1/6

Penalizaciones.

Las multas por violaciones de la ley de IA se establecieron como un porcentaje de la facturación anual global de la empresa infractora en el año financiero anterior o una cantidad predeterminada, el que fuera mayor.

Una cifra que representaría 35 millones de euros o el 7% por violaciones de las aplicaciones de IA prohibidas, 15 millones de euros o el 3% por violaciones de las obligaciones de la ley de IA y 7,5 millones de euros o el 1,5% por el suministro de información incorrecta.

Sin embargo, el acuerdo provisional establece

límites más proporcionados a las multas administrativas para las pymes y las empresas de nueva creación en caso de infracción de las disposiciones de la Ley de IA.

El acuerdo de transacción también deja claro que una persona física o jurídica puede presentar una queja ante la autoridad de vigilancia del mercado pertinente en relación con el incumplimiento de la Ley de IA y puede esperar que dicha queja se tramite de acuerdo con los procedimientos específicos de esa autoridad.

1/7

Transparencia y protección de los derechos fundamentales.

El acuerdo provisional prevé una evaluación del impacto sobre los derechos fundamentales antes de que sus implementadores pongan en el mercado un sistema de inteligencia artificial de alto riesgo. El acuerdo provisional también prevé una mayor transparencia con respecto al uso de sistemas de inteligencia artificial de alto riesgo.

En particular, se han modificado algunas disposiciones de la propuesta de la Comisión para indicar que ciertos usuarios de un sistema de IA de alto riesgo que sean entidades públicas también estarán obligados a registrarse en la base de datos de la UE para sistemas de IA de alto riesgo. Además, las disposiciones recientemente añadidas hacen hincapié en la obligación de que los usuarios de un sistema de reconocimiento de emociones informen a las personas físicas cuando estén expuestas a dicho sistema.

1/8

Medidas de apoyo a la innovación.

Con vistas a crear un marco jurídico más favorable a la innovación y promover el aprendizaje

normativo basado en pruebas, las disposiciones relativas a las medidas de apoyo a la innovación se han modificado sustancialmente en comparación con la propuesta de la Comisión.

En particular, se ha aclarado que los entornos limitados de pruebas regulatorias de IA, que se supone deben establecer un entorno controlado para el desarrollo, prueba y validación de sistemas de IA innovadores, también deberían permitir probar sistemas de IA innovadores en condiciones del mundo real. Además, se han agregado nuevas disposiciones que permiten probar sistemas de IA en condiciones del mundo real, bajo condiciones y salvaguardias específicas. Para aliviar la carga administrativa para las empresas más pequeñas, el acuerdo provisional incluye una lista de acciones que deben emprenderse para apoyar a dichos operadores y prevé algunas derogaciones limitadas y claramente especificadas.

1/9

Entrada en vigor.

El acuerdo provisional establece que la ley sobre IA debería aplicarse dos años después de su entrada en vigor, con algunas excepciones para disposiciones específicas.

1/10

Próximos pasos.

Tras la firma del acuerdo provisional, en las próximas semanas se seguirá trabajando a nivel técnico para ultimar los detalles del nuevo reglamento. La Presidencia presentará el texto transaccional a los representantes de los Estados miembros (Coreper) para su aprobación una vez concluido este trabajo. El texto completo deberá ser confirmado por ambas instituciones y sometido a una revisión jurídico-lingüística antes de su adopción formal por los legisladores.

1/11

Información de contexto.

La propuesta de la Comisión, presentada en abril de 2021, es un elemento clave de la política de la UE para fomentar el desarrollo y la adopción en todo el mercado único de una IA segura y legal que respete los derechos fundamentales.

La propuesta sigue un enfoque basado en el riesgo y establece un marco jurídico uniforme y horizontal para la IA cuyo objetivo es garantizar la seguridad jurídica. El borrador del reglamento tiene como objetivo promover la inversión y la innovación en IA, mejorar la gobernanza y la aplicación efectiva de la legislación existente sobre derechos fundamentales y seguridad, y facilitar el desarrollo de un mercado único para las aplicaciones de IA. Va de la mano junto a otras iniciativas, incluido el plan coordinado sobre inteligencia artificial que tiene como objetivo acelerar la inversión en IA en Europa.

El 6 de diciembre de 2022, el Consejo alcanzó un acuerdo sobre una orientación general (mandato de negociación) sobre este expediente y entabló conversaciones interinstitucionales con el Parlamento Europeo («diálogos tripartitos») a mediados de junio de 2023.



Ley de IA.
Comisión Europea



Reglamento(UE) 2024/1689
del Parlamento Europeo y
del Consejo. 13/06/2024





01/

Inteligencia Artificial

Generativa:

funcionamiento,
posibilidades
y riesgos.

Climent Nadeu Camprubí,

profesor Universitat Politècnica de Catalunya (UPC) y catedrático de la Escola d'Enginyeria de Telecomunicació. Barcelona.

Se presenta una breve introducción a la Inteligencia Artificial (IA) centrada en su aspecto de mayor relevancia en la actualidad: la IA generativa. Con una explicación elemental de su funcionamiento, se pretende corregir conceptos erróneos y resaltar las razones de su gran potencial. A partir de esta comprensión básica se muestran las principales limitaciones de las herramientas de IA, los riesgos que conllevan y las cuestiones éticas que plantean, tanto en sí mismas como por la manera de usarlas. Por último, se incluyen algunas consideraciones sobre cómo prepararse para un futuro con esta tecnología.

Palabras clave: Inteligencia Artificial Generativa, Modelos de lenguaje, Alineación, Ética de la tecnología.

A short introduction to Artificial Intelligence (AI) is presented, focusing on its most relevant aspect nowadays: generative AI. By providing a basic explanation of how it works, the aim is to correct misconceptions and highlight the reasons for its strong potential. From this basic understanding, the main limitations of AI tools, the associated risks, and the involved ethical issues, both in themselves and in how they are used, are discussed. Finally, some considerations are included on how to prepare for a future with this technology.

Key words: Generative Artificial Intelligence, Language models, Alignment, Ethics of technology.

La inteligencia y la capacidad técnica son dos rasgos que nos distinguen como especie en nuestro planeta. Si bien la tecnología, es decir, la técnica basada en conocimiento científico, se ha centrado tradicionalmente en el mundo material, las fuerzas, la energía y la información, en las últimas décadas ha comenzado a explorar un nuevo territorio, el de la inteligencia humana.

Este avance es posible gracias a la **inteligencia artificial (IA)**, área tecnológica que desarrolla programas informáticos capaces de realizar tareas cognitivas propias del ser humano, como el manejo del lenguaje escrito y oral, la percepción visual, el aprendizaje, la toma de decisiones, la creación artística o el desarrollo de los propios programas informáticos.

Los programas de IA pueden integrarse en una amplia gama de dispositivos y aplicaciones. Estos programas emplean **algoritmos**, procedimientos matemáticos sofisticados que permiten procesar y analizar vastas cantidades de información. Para ello, se basan en **modelos** de la realidad que actúan como mapas, guiándolos en la obtención de resultados a partir de la información proporcionada por los usuarios.

Los inicios de la IA se pueden situar unos 70 años atrás y su desarrollo a lo largo del tiempo ha sufrido altibajos. Durante mucho tiempo los sistemas de IA más apreciados manipulaban símbolos, usando reglas lógicas para hacer inferencias y deducciones. Actualmente predomina la representación numérica y el **aprendizaje automático** a partir de datos (machine learning).

Y dentro de este enfoque destacan las redes neuronales artificiales, llamadas así porque imitan, de forma muy limitada, el funcionamiento del cerebro.

La técnica de las **redes neuronales artificiales (Artificial Neural Networks, ANN)**, aunque relativamente simple, es tan antigua en sus principios básicos como la propia IA. La reciente eclosión de las tecnologías de IA, especialmente de la IA generativa, que ha captado la atención del público en general con la aparición de herramientas como **ChatGPT**, se debe más a la convergencia de dos factores clave que a avances científicos disruptivos. Estos factores son el poder de computación cada vez mayor y la explosión -gracias a Internet- de datos digitales disponibles, los cuales proporcionan la materia prima esencial para el entrenamiento de estas redes.

Aquí trataremos de la IA generativa por su potencialidad y actualidad, aunque en el ámbito de la salud se usan también herramientas de IA de otros tipos (ver, por ejemplo, el artículo referenciado de **Diego Urgelés**).

La IA generativa sirve para crear de forma automática documentos escritos u orales, contenidos audiovisuales (imagen, vídeo, pieza musical), e incluso programas informáticos.

La estrella de la IA generativa es el denominado **modelo de lenguaje de gran tamaño (Large Language Model, LLM)**, resultado de años de investigación en tratamiento del lenguaje natural (natural en el sentido de humano, en oposición a los lenguajes de programación). Se dice que es un modelo fundamental (**foundation model**) porque sirve de base para el desarrollo de herramientas específicas. Además del LLM existen modelos fundamentales para generar otros tipos de contenidos: imágenes, música, etc.

La generación tiene lugar a partir tanto de la información que el modelo tiene codificada internamente como de la consulta realizada por el usuario, que se denomina **prompt**. El **prompt**, al igual que el documento generado, puede

El auge de la IA generativa se debe principalmente al aumento en el poder de computación y la disponibilidad masiva de datos digitales

adoptar diferentes modalidades: textual, audiovisual, etc. Cuando se usan conjuntamente varias modalidades los correspondientes modelos multimodales se acercan más a la forma en que los seres humanos interactuamos con el mundo.

ChatGPT consiste en un LLM, un modelo de lenguaje que engloba varios idiomas a la vez y está formado por estructuras computacionales basadas en redes neuronales, ANNs, las cuales básicamente realizan multiplicaciones y sumas en cada uno de sus nodos.

A veces se supone erróneamente que, para generar su respuesta al prompt de entrada, el sistema LLM realiza búsquedas en bases de datos textuales, ya sean propias o de Internet, y utiliza los resultados para componer las frases. Sin embargo, como veremos, el proceso no consiste en construir frases a partir de resultados de búsqueda, sino en realizar cálculos numéricos usando la representación codificada del lenguaje que posee internamente. La letra T de ChatGPT proviene de la estructura funcional concreta que suelen utilizar actualmente los LLMs: el **Transformador**, la más novedosa aportación de los últimos años en el ámbito del llamado aprendizaje profundo (deep learning).

En la **Figura 1** puede verse un esquema simple de su funcionamiento.

1/

¿Cómo funciona la IA? ChatGPT como ejemplo.

Para entrar con cierto detalle en el funcionamiento de las herramientas basadas en IA generativa, nos vamos a centrar en el caso de los chatbots o sistemas conversacionales, que operan con el lenguaje, ya sea escrito u oral. Actualmente existen varios sistemas disponibles, como el pionero y famoso **ChatGPT**, o sus competidores Gemini, Claude, Llama, etc.

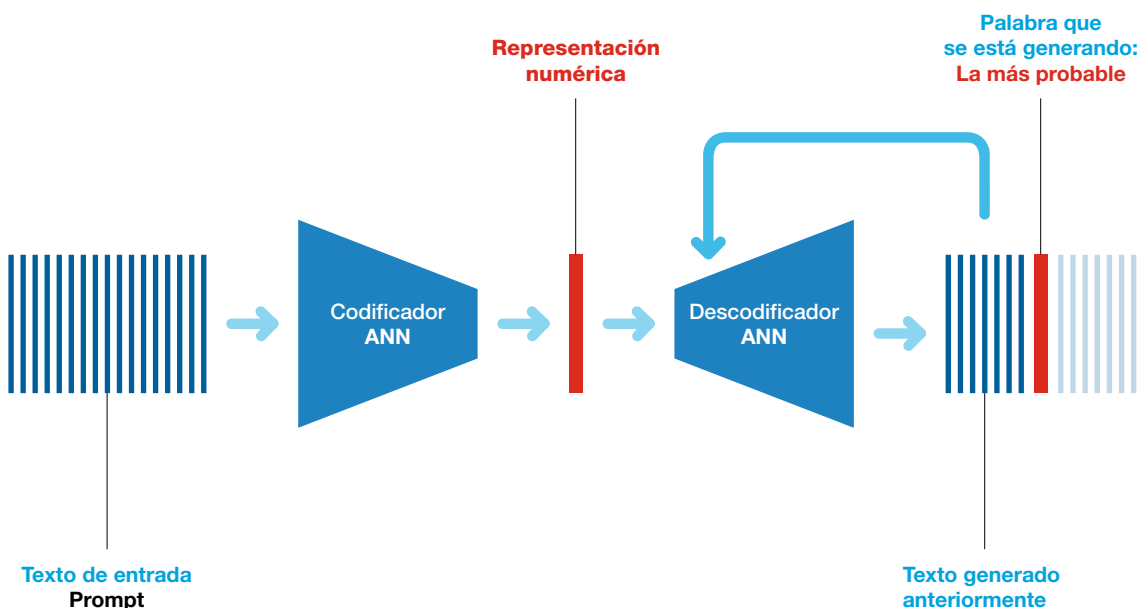


Figura 1. Arquitectura funcional del Transformador, base de LLMs como el de ChatGPT

LH n.339

En la figura, a la izquierda se muestra el prompt o texto de entrada, mientras que a la derecha se observa el texto generado a la salida, que se crea palabra a palabra (la palabra actual está destacada en naranja). El objetivo no es inicialmente generar un texto con sentido o lingüísticamente correcto; esto es un resultado secundario.

En su lugar, el proceso consiste en calcular las probabilidades de todas las palabras posibles y seleccionar la palabra con la probabilidad más alta. Luego, se procede a generar la siguiente palabra, repitiendo el proceso de forma secuencial. En realidad, no se usan palabras del diccionario sino partes de palabras (tokens), porque resultan más eficientes.

Para ilustrarlo con un ejemplo: si el texto de entrada pregunta sobre la información meteorológica, y en la salida ya se ha generado la frase “**hacia media mañana podrá haber...**”, la palabra siguiente no puede ser cualquiera, y hay unas pocas palabras que tienen alta probabilidad: nubes, chubascos, ventoleras,... Son estas correlaciones entre palabras y conceptos lo que está contenido de forma codificada en el modelo.

Las probabilidades de todas las posibles palabras (o tokens) se calculan con el Transformador, que consta de dos partes (color verde). El codificador, consistente en una ANN, convierte el texto de entrada en una representación numérica, una secuencia de números, que puede contener miles de elementos. El prompt queda codificado y comprimido en esta secuencia numérica, de manera que han desaparecido las palabras y el texto en sí mismo ya no puede recuperarse.

El decodificador también consiste en una ANN y se encarga de calcular la probabilidad de que cada palabra posible sea la siguiente en el texto que se va generando. Para ello, utiliza no sólo la representación numérica del prompt obtenida por el codificador sino también el texto generado anteriormente (color azul fuerte); este texto se ha retroalimentado, como se indica en la figura, y numerizado.

Interpretando el funcionamiento del modelo se puede decir que las palabras están representadas como puntos en un espacio multimensional, donde la proximidad entre dos puntos guarda relación con el parecido semántico de las palabras.

De hecho, dado que el lenguaje vehicula el conocimiento y la experiencia de los seres humanos, de alguna forma el modelo, obtenido procesando los textos (o grabaciones orales) existentes, contiene una representación conceptual del mundo, lo que permite generar texto que simule pensamientos, emociones, vivencias, etc.

Esto no significa, sin embargo, que el sistema vincule las palabras con sus referentes en el mundo real ni que les otorgue significado. Por ejemplo, si genera la frase “**la mesa es de madera**”, no está asociando como nosotros las palabras “**mesa**” y “**madera**” con algo aprendido por medio de la experiencia, aunque pueda parecerlo.

1/1

Entrenamiento de ChatGPT.

Las redes neuronales se caracterizan por sus nodos y las conexiones entre ellos. Básicamente, en cada nodo se lleva a cabo una suma ponderada de los valores provenientes de otros nodos. Los pesos de la ponderación determinan la importancia relativa de cada conexión entre las neuronas y son parámetros que deben ajustarse durante la etapa de entrenamiento del modelo, la cual se realiza automáticamente utilizando un conjunto de textos.

La letra P de ChatGPT se refiere precisamente a dicho entrenamiento, es decir, a la asignación del valor de los pesos. Significa Pre-entrenado, porque el chatbot se construye partiendo de un LLM de tipo general entrenado previamente, para adaptarlo a la aplicación específica de chatbot, tal como indica la **Figura 2**.



Figura 2. Entrenamiento de ChatGPT

En la Figura 2 puede verse un esquema del proceso de diseño. En primer lugar, se determina un LLM genérico con unos algoritmos de entrenamiento que usan enormes cantidades de textos extraídos básicamente de Internet (páginas web, blogs, redes sociales, artículos científicos, libros, etc.), en varios idiomas. El algoritmo de entrenamiento trabaja de forma autónoma, sin supervisión humana (entrenamiento auto-supervisado); va ocultando palabras o frases de los textos de entrenamiento y trata de predecirlas, una tras otra. Con los errores cometidos en la predicción se van ajustando iterativamente los valores de los pesos a fin de reducir progresivamente el error acumulado sobre el conjunto del corpus de entrenamiento.

Durante esta fase de entrenamiento automático se capta en gran medida la estructura de la lengua (o lenguas, si hay varias en los textos de entrenamiento). El algoritmo consigue extraer patrones, regularidades estadísticas, que corresponden básicamente a relaciones semánticas y morfosintácticas. Dicha información queda codificada numéricamente de forma compacta, lo cual dificulta mucho su observación por parte del diseñador.

Partiendo del modelo pre-entrenado, en un segundo paso se hace el entrenamiento definitivo del LLM chatbot a base de afinar (fine-tuning), de ajustar el modelo genérico partiendo de

datos más específicos aportados con intervención humana (por personas que redactan conversaciones, puntúan respuestas que da el sistema,...). Y también usando contenido etiquetado específicamente como no ético, para tratar de evitar que el chatbot genere textos no aceptables.

Nos hemos centrado en la aplicación más conocida, la del chatbot, pero los LLMs pueden estar detrás de otras muchas funcionalidades: traducción automática (que fue la primera aplicación del Transformador), asistentes de voz, respuesta a preguntas, resumen automático, corrección gramatical, generación de software, etc.

2/

Posibilidades y utilidad de la IA generativa.

Al probar ChatGPT, rápidamente se percibe la importancia de redactar buenos prompts. Hay que situar el contexto de lo que se quiere generar, guiar los pasos a realizar, especificar la forma de presentar la información que se solicita, usar expresiones claras y precisas, ...

LH n.339

De hecho, el prompt es una forma de adaptar el procesamiento que se realiza con el LLM a lo que pide el usuario, tanto en forma como en contenido. Equivale en cierto modo a una afinación del modelo, como la que hemos visto que se realiza durante la fase final de entrenamiento, pero en este caso sin necesidad de cambiar el modelo ni de modificar sus parámetros.

Esta capacidad de adaptación al prompt que muestran los LLMs se ha hecho evidente sobre todo cuando se ha pasado a modelos a gran escala (más de 1010 parámetros; por ejemplo, de GPT2 a GPT3). Se llama aprendizaje en contexto (in-context learning) porque, combinando en los prompts descripciones de la tarea con ejemplos de demostración, los LLMs se muestran capaces de obtener buen rendimiento en tareas nuevas, incluso superando en algunos casos los modelos afinados durante el entrenamiento con un conjunto de datos suficientemente grande.

Esta forma de funcionar no había sido prevista y ha sorprendido a los propios investigadores. Y lo mismo ocurre con la posibilidad de dar la respuesta desglosando en pasos la argumentación. Por ello se dice que estos modelos presentan **capacidades** emergentes.

Actualmente ya se dispone de asistentes personales como los chatbots que aceptan prompts de diversas modalidades (texto, voz, imagen, programa informático, etc.) y que generan contenidos también multimodales. También existen sistemas compuestos de IA que combinan distintas componentes (modelos, bases de datos, herramientas externas, etc.) interactuando entre sí para llevar a cabo la tarea deseada.

Además, en poco tiempo aparecerán herramientas todavía más potentes: “**agentes**” de IA que planifiquen, gestionen y dirijan la ejecución de tareas complejas, los cuales serán pilotados por LLMs y se relacionarán con el usuario a través de ellos.

3/

Limitaciones y riesgos de la IA generativa.

Como todos los productos tecnológicos, las herramientas actuales de IA generativa tienen sus limitaciones e implican ciertos riesgos. En primer lugar, como son modelos basados en redes neuronales, aunque puedan observarse las operaciones matemáticas que realizan, resultan muy opacos en el sentido de que no existen unos criterios y una forma razonada de explicar cómo se ha llegado al resultado a partir de los datos de entrada (el prompt). Esta dificultad, que es especialmente grave en ámbitos como el de la salud o la justicia, es bastante común entre los sistemas basados en aprendizaje automático, pero lo es menos en otros sistemas de IA, como los que se basan en aplicación de reglas.

Es bien conocido que los chatbots actuales producen errores. Preocupan sobre todo las -mal llamadas alucinaciones, respuestas que parecen plausibles, pero de hecho son incorrectas.

Recordemos que el objetivo del modelo es maximizar la probabilidad del texto generado, no su adecuación a la realidad. Las versiones más recientes cometen menos errores y este tipo de sistemas irá mejorando continuamente.

Como ocurre con todos los sistemas de IA, las herramientas generativas pueden presentar sesgos que discriminan a individuos o grupos, ya sea por género, etnia, u otros factores. Estos sesgos están estrechamente vinculados al conjunto de datos de entrenamiento utilizado. En el caso de las herramientas generadoras de texto, es importante señalar que no todas las lenguas están representadas en el modelo, y el grado de representación varía considerablemente según el idioma, con el inglés destacando notablemente por encima de los demás.

La IA generativa presenta riesgos como la opacidad, la generación de contenido falso y los sesgos, lo que demanda una regulación urgente

Otra característica que condiciona el desarrollo y uso de la IA generativa es el elevado coste asociado y los problemas de sostenibilidad que esto implica. Se necesitan enormes recursos computacionales, tanto de procesamiento como de almacenamiento, de modo que tan sólo unas pocas empresas e instituciones pueden disponer de ellos. Y, por supuesto, significan un gran coste para el planeta en cuanto a su producción, mantenimiento, etc. y también por su elevado consumo de energía.

Los riesgos mencionados anteriormente están asociados a las limitaciones inherentes de las herramientas de IA generativa. Ahora examinaremos brevemente algunos riesgos específicos y relevantes relacionados con su uso. El primero a destacar es la notable facilidad con la que estas herramientas pueden generar contenido falso o engañoso (*fake*), lo cual puede ser utilizado para manipular a las personas. No es difícil imaginar un escenario en el que Internet se vea inundada de información falsa, lo que podría resultar en una pérdida generalizada de confianza en los contenidos.

En el ámbito laboral, los principales riesgos son la vulneración de los derechos de autor, especialmente cuando se utilizan imágenes, textos y otros contenidos con propiedad intelectual para entrenar los modelos, y la considerable afectación que la irrupción de estas herramientas probablemente tendrá en el mercado de trabajo.

Además, es difícil detectar plagio porque el texto o imagen generados nunca son iguales a ninguno de los que se han usado para entrenar el modelo. Y, como ocurre con este tipo de herramientas online, la privacidad puede estar comprometida; muchas empresas prohíben a sus trabajadores usar los *xatbots* de dominio público porque los *prompt* pueden transportar datos confidenciales.

Dado que conversa fluidamente y escribe con esmero, los usuarios tendemos a interactuar y referirnos al chatbot como si fuera un ser humano (se dice que **“ha aprendido”**,

“comprende”, **“dialoga”**, **“se emociona”**, **“tiene valores”**...) porque asociamos capacidad de lenguaje a inteligencia humana. De hecho, como ocurre con muchos artefactos tecnológicos, la IA nos resulta fascinante y al mismo tiempo temible, es decir, tendemos a sacralizarla.

De esta lista no exhaustiva de riesgos asociados a las herramientas de IA generativa, y a las de texto en particular, se desprende la necesidad de regulación, necesidad compartida por muchas tecnologías que incorporan IA. Recientemente, la UE ha aprobado una pionera Ley de Inteligencia Artificial que tiene en cuenta los distintos niveles de riesgo.

4/

Algunas consideraciones sobre nuestro futuro con IA y la forma de enfrentarlo.

Dados los desafíos y oportunidades que presenta la IA, es lógico preguntarse cómo debemos posicionarnos ante estas herramientas tecnológicas. Parece razonable evitar tanto el rechazo como la aceptación acrítica. Es preferible una actitud constructiva, consciente de la naturaleza de estas herramientas y que sea también crítica respecto a cómo las utilizamos.

Puesto que previsiblemente las herramientas de IA se van a integrar cada vez más en nuestra vida cotidiana, es necesario estar bien informados, conocer sus capacidades y limitaciones, para poder usarlas de manera efectiva y ética. Es importante comprender con qué objetivos de optimización funciona la herramienta e interpretar adecuadamente los resultados que proporciona.

LH n.339

En el caso de la IA generativa, es esencial dominar la interacción y el uso de prompts para maximizar su utilidad.

La IA conlleva ciertos problemas y amenazas, pero también ofrece grandes oportunidades. Puede liberarnos de tareas repetitivas o que consumen mucho tiempo, y abrirnos nuevas posibilidades. Un objetivo posiblemente deseable es que las herramientas de IA se encarguen de las tareas habituales que requieren conocimiento, ya sean rutinarias o complejas, permitiendo a las personas concentrarse en procesos más creativos o que precisan del cuidado humano.

Es previsible que la evolución de la tecnología de IA alcance un nivel de desarrollo en el que se considere la conveniencia de otorgar autonomía de decisión a ciertos sistemas de IA en determinadas situaciones.

Ante esta posibilidad, se podría optar por la medida restrictiva de mantener siempre el control en manos humanas. Sin embargo, esta opción podría ser vista como un obstáculo para la eficiencia o resultar inviable debido a las posibles discrepancias entre los usuarios. Además, como se ha señalado, los modelos fundamentales de IA pueden desarrollar capacidades emergentes no previstas por sus diseñadores, y no se puede descartar la posibilidad de que en la fase operativa la IA se comporte de manera distinta a como fue entrenada.

Este conjunto de posibilidades plantea un problema de seguridad en el despliegue de los sistemas de IA. La situación se torna aún más preocupante si consideramos las advertencias de destacados investigadores (como [G. Hinton](#) y [G. Bengio](#), entre otros) sobre el riesgo existencial para la humanidad debido a la IA. Ellos sostienen que el rápido desarrollo de la IA podría alcanzar un nivel de inteligencia tal que los seres humanos queden sin capacidad de control.

Por todo ello, en los ámbitos de desarrollo de la IA crece la preocupación y el interés por trabajar en el problema de la alineación entre el objetivo

que persigue realmente el sistema o herramienta IA y el objetivo buscado por el diseñador y el usuario, o más en general, entre los objetivos de la IA y las intenciones y valores humanos. Sin embargo, no es en absoluto inmediato transferir un objetivo al algoritmo de entrenamiento o de operación de los sistemas que triunfan actualmente, especialmente si es de naturaleza cualitativa, debido a la dificultad de expresarlo en términos matemáticos.

Es evidente, por tanto, que la IA representa uno de los grandes desafíos de nuestra época, un desafío que debemos enfrentar desde distintas perspectivas. A continuación, se enumeran tres enfoques complementarios, con la esperanza de que sean compartidos por el lector.

En primer lugar, es necesario desarrollar tecnologías menos opacas y capaces de seguir explícitamente los objetivos que se les asignen. En segundo lugar, como sociedad, debemos establecer mecanismos de regulación y control, incluso antes de que las nuevas herramientas comiencen a aplicarse. Y, en tercer lugar, como personas, debemos cultivar aquellas cualidades que más nos diferencian de la IA y mejor pueden orientar su desarrollo y uso responsable: inteligencia creativa, para adaptarnos a un entorno en constante cambio; sabiduría, para tomar decisiones equilibradas y éticas; y actitud contemplativa, para no someter nuestras funciones mentales al ritmo acelerado propiciado por la IA.



Bibliografía

▶ **Christian, B.**

The alignment problem. Machine learning and human values,
W.W.Norton&Company, 496 pág., 2020.

▶ **Coeckelbergh,**

M. Ética de la inteligencia artificial.
Cátedra. 2021.

▶ **Libro blanco de la Inteligencia**

Artificial Generativa,
Grupo de trabajo IA Generativa,
Transformación Digital, DIGITALES.
Junio 2024.

▶ **López de Mántaras Badia,**

R. 100 cosas que cal saber sobre la intel- ligència artificial,
Cossetània, 319 pág., 2023.

▶ **Noble, D. F.**

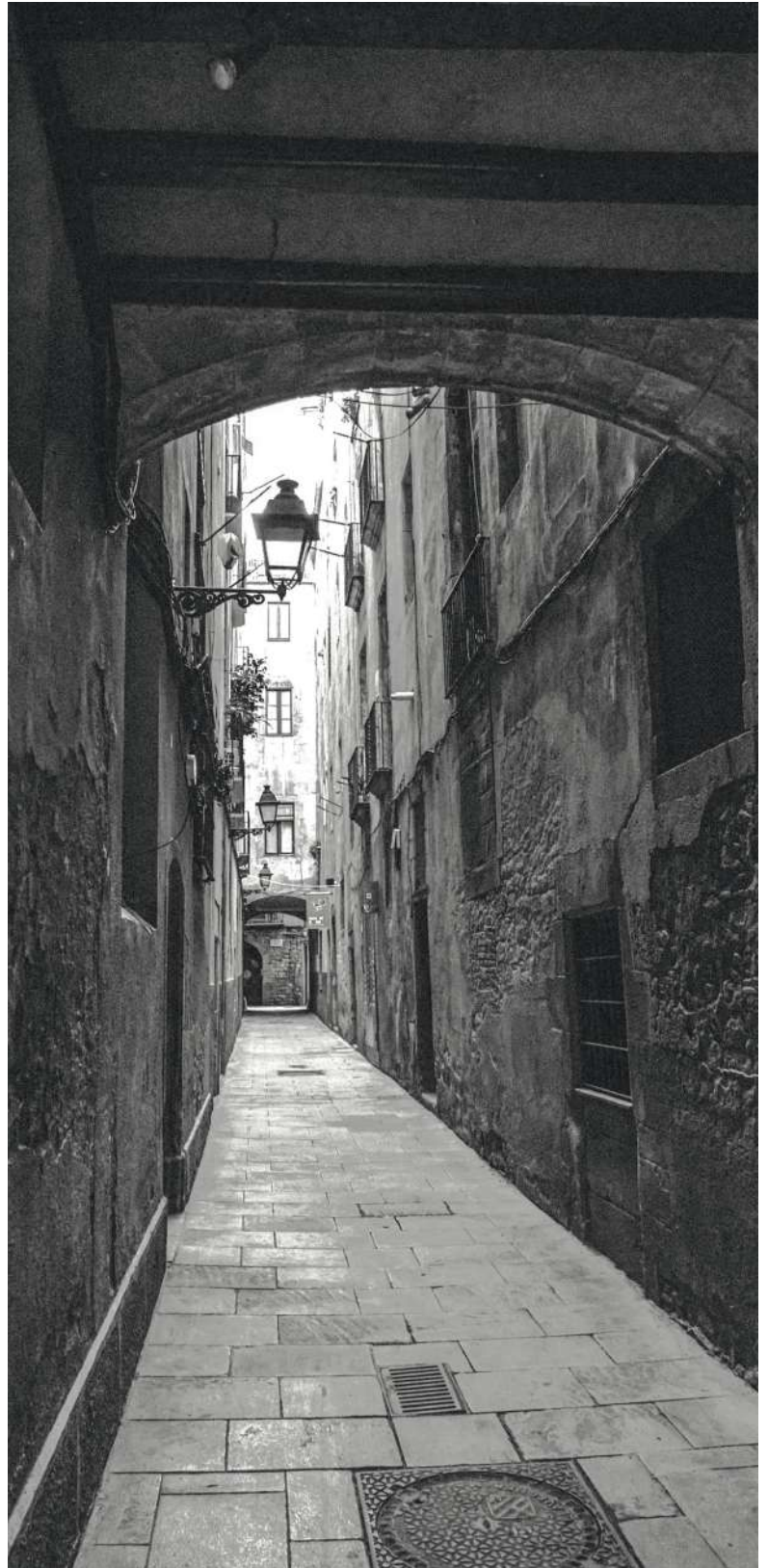
La religión de la tecnología,
Paidós, 298 pág., 1999.

▶ **Torres, J.**

La inteligencia artificial explicada a los humanos.
Plataforma Actual. 2023

▶ **Urgelés Puértolas, D.**

“Implicaciones a corto plazo de la inteligencia artificial en la atención sanitaria”
en LABOR HOSPITALARIA n.336-337,
pp.36-43, marzo 2023.







02/

La Inteligencia Artificial

y su impacto en la bioética contemporánea.

Ramón Castejón García OH,

Enfermero especialista en Salud Mental.

Dirección y Administración de Servicios de Enfermería.

Máster de Gestión Sanitaria y Hospitalaria.

Centro Santa María de la Paz. Madrid.

La inteligencia artificial (IA) está transformando el campo de la salud, ofreciendo avances significativos en diagnóstico, tratamiento y atención al paciente. Sin embargo, estos desarrollos plantean desafíos éticos que requieren una reevaluación de los principios fundamentales de la bioética. Este artículo explora cómo la IA adquiere y procesa datos para generar conocimiento y destaca sus aplicaciones clave en la salud, como el diagnóstico médico, la predicción de enfermedades, y el desarrollo de medicamentos. Asimismo, se analizan las implicaciones bioéticas relacionadas con la privacidad, la equidad en el acceso a tecnologías avanzadas, y la humanización de la atención médica, subrayando la necesidad de un marco ético robusto para guiar el uso responsable de la IA en la medicina.

Palabras clave: Inteligencia Artificial (IA), Bioética, Diagnóstico médico, Algoritmo, Salud.

Artificial intelligence (AI) is transforming the health field, producing significant advances in the diagnosis, treatment and care of patients. Nevertheless, these developments pose ethical challenges that call for a reassessment of the fundamental principles of bioethics. This article explores how AI acquires and processes data to generate knowledge, highlighting its key health applications, as in medical diagnosis, prediction of disorders and development of pharmaceuticals. Likewise, consideration is given to bioethical implications relating to privacy, equity of access to advanced technologies, and humanisation of medical care, emphasising the need for a robust ethical framework to guide the responsible use of AI in medicine..

Key words: Artificial Intelligence (AI), Bioethics, Medical diagnosis, Algorithm, Health.

1. Un algoritmo es un grupo finito de operaciones organizadas de manera lógica y ordenada que permite solucionar un determinado problema. Se trata de una serie de instrucciones o reglas establecidas que, por medio de una sucesión de pasos, permiten arribar a un resultado o solución. En matemáticas, lógica, ciencias de la computación y disciplinas relacionadas, un algoritmo es un conjunto de instrucciones o reglas definidas y no ambiguas, ordenadas y finitas que permite, solucionar un problema, realizar un cómputo, procesar datos y llevar a cabo otras tareas o actividades. En informática o programación, el algoritmo es la secuencia de instrucciones mediante la cual podemos resolver un problema o cuestión. De hecho, todas las tareas ejecutadas por la computadora se basan en algoritmos. Un software o programa informático está diseñado a base de algoritmos, de modo que podemos introducir una tarea en él y resolverla.

La inteligencia artificial (IA) ha irrumpido en diversos campos, transformando industrias y redefiniendo procesos. En el ámbito de la salud, sus avances ofrecen prometedoras mejoras en el diagnóstico, tratamiento y atención al paciente. Sin embargo, estas innovaciones también plantean desafíos éticos significativos.

La bioética, como disciplina que aborda las cuestiones éticas surgidas de la medicina y las ciencias de la vida, se encuentra en un punto de inflexión.

No obstante, la IA se ha venido desarrollado a lo largo de décadas mediante la combinación de investigación científica, avances tecnológicos y algoritmos¹.

- **Aprendizaje automático (Machine Learning)**²: Es una rama fundamental de la IA que se basa en entrenar modelos con datos para que puedan aprender patrones y tomar decisiones sin programación explícita.
- **Redes neuronales artificiales**: Inspiradas en el funcionamiento del cerebro humano; estas redes son fundamentales para el aprendizaje profundo (**deep learning**).
- **Algoritmos genéticos**: Se inspiran en la evolución biológica y se utilizan para optimizar soluciones en problemas complejos.

La inteligencia artificial (IA) ha dejado de ser una mera promesa futurista y se ha convertido en una realidad que impacta profundamente en la medicina y la atención sanitaria.

La irrupción de la inteligencia artificial en la salud es un fenómeno que ha capturado la imaginación y las expectativas de muchos, ofreciendo visiones de un futuro donde la precisión y la eficiencia del diagnóstico y el tratamiento médico se ven amplificadas por la capacidad de procesamiento y análisis de la IA. Sin embargo, estas promesas deben ser analizadas con un ojo crítico hacia los desafíos éticos que inevitablemente surgen.

La bioética, como campo dedicado a examinar las implicaciones morales de la biomedicina, se enfrenta a nuevos dilemas que van más allá de los problemas tradicionales.

Este artículo pretende explorar cómo la IA puede influir en los conceptos actuales de bioética, especialmente en lo referente a los cuidados de la salud, el tratamiento de enfermedades y la humanización de la atención al paciente.

1/

¿Cómo la IA adquiere y procesa información para generar conocimiento?

La inteligencia artificial (IA) se sustenta en la adquisición y el procesamiento de datos, que constituyen la esencia misma de su funcionamiento. Estos datos actúan como los bloques de construcción fundamentales sobre los cuales se entrena y desarrolla la IA. Una de las principales fuentes de datos son los sensores, dispositivos que capturan información del entorno real, como la temperatura, la humedad, la presión atmosférica o cualquier otra variable cuantificable. Estos sensores pueden estar integrados en dispositivos físicos o instalados en sistemas específicos para recopilar datos en tiempo real³.

Por otro lado, la vasta red de Internet también desempeña un papel crucial como fuente de datos para la IA. La web alberga una cantidad abrumadora de información, que incluye desde artículos de noticias, documentos académicos y libros electrónicos hasta publicaciones en redes sociales, vídeos, imágenes y más. La capacidad de acceder y procesar esta inmensa cantidad de datos proporciona a los sistemas de IA una riqueza incomparable de información sobre una amplia gama de temas y áreas de interés⁴.

Además, la IA también puede aprovechar datos generados por humanos, como transcripciones de conversaciones, historiales médicos, registros financieros, datos de tráfico, entre otros. Estos datos, a menudo estructurados en bases de datos, archivos de texto o formatos específicos, proporcionan información valiosa que puede ser utilizada para entrenar modelos de IA en una variedad de aplicaciones y contextos⁵.

Una vez que los datos son recopilados, la IA los procesa utilizando una variedad de técnicas, como el aprendizaje automático y el procesamiento del lenguaje natural, para extraer patrones, identificar correlaciones y generar conocimiento⁶.

Este proceso de análisis de datos es fundamental para el funcionamiento de la IA, ya que permite a los sistemas comprender el mundo que los rodea, tomar decisiones informadas y realizar tareas específicas de manera eficiente y precisa.

2/

La Inteligencia Artificial en la Salud: Panorama General.

La IA se ha incorporado en múltiples facetas del cuidado de la salud. Entre sus aplicaciones más destacadas se incluyen:

1. Diagnóstico Médico: La IA tiene el potencial de revolucionar el diagnóstico médico al analizar grandes volúmenes de datos y detectar patrones que pueden escapar al ojo humano. Algoritmos avanzados pueden interpretar imágenes médicas, como radiografías y resonancias magnéticas, con una precisión que rivaliza o incluso supera a la de los radiólogos, prometiendo así mejoras significativas en la precisión diagnóstica y la predicción de brotes de enfermedades.

Sin embargo, esta capacidad también plantea preguntas éticas importantes. ¿Cómo se asegura la transparencia en los algoritmos de IA para que los profesionales de la salud y los pacientes puedan entender y confiar en las decisiones tomadas por estas tecnologías?⁷ Además, la dependencia excesiva en la IA podría deshumanizar la práctica médica, convirtiendo a los pacientes en meros conjuntos de datos en lugar de individuos con experiencias y contextos únicos⁸.

2. Predicción de Enfermedades: Los sistemas de IA pueden prever la aparición de enfermedades basándose en la historia clínica y los datos genéticos de los pacientes. Estos sistemas analizan grandes volúmenes de datos, identificando patrones y factores de riesgo que pueden no ser evidentes para los médicos humanos. Al integrar información detallada de la historia médica del paciente con datos genéticos, la IA puede generar predicciones precisas sobre la probabilidad de desarrollar ciertas enfermedades⁹.

En resumen, los algoritmos son como recetas paso a paso que nos guían para resolver problemas o realizar tareas específicas.

2. Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill.

3. Russell, S., & Norvig, P. (2016). *Inteligencia Artificial: Un Enfoque Moderno* (3ra ed.). Pearson.

4. Aguilar, J. (2011). *Inteligencia Artificial: Fundamentos, métodos y aplicaciones*. Alfaomega.

5. Muñoz, E. (2010). *Sensores y Acondicionadores de Señal*. Editorial Marcombo.

6. Escobar, A., & Silva, J. (2015). *Instrumentación y Sensores*. Alfaomega.

7. Díaz-Gómez, M., & Macías, A. (2020). *Inteligencia Artificial en Imágenes Médicas*. Editorial Diaz de Santos. Topol, E. (2019). *La Medicina de la Precisión: Cómo la Revolución Digital está Creando Mejores Cuidadores y Pacientes Más Sanos*. Ediciones Urano.

8. Escobar, A., & Silva, J. (2015). *Instrumentación y Sensores*. Alfaomega.

9. López-Campos, G., & Ochoa, S. (2019). *Big Data y Predicción en Salud: Aplicaciones y Retos*. Ediciones Pirámide.

10. Topol, E. (2019). La Medicina de la Precisión: Cómo la Revolución Digital está Creando Mejores Cuidadores y Pacientes Más Sanos. Ediciones Urano.

11. Zhavoronkov, M. (2018) Inteligencia artificial en el descubrimiento y desarrollo de fármacos. Future Medicinal Chemistry.

12. Coley, C.W., Green, W.H., Jensen, K. F. (2019) Aprendizaje profundo para el descubrimiento de fármacos y el desarrollo de biomarcadores. Annual Review of Chemical and Biomolecular Engineering.

13. Mamoshina, P., Vieira, A., Putin, E., Zhavoronkov, A. (2016) Inteligencia artificial en el descubrimiento de fármacos: avances recientes y perspectivas futuras. Expert Opinion on Drug Discovery.

Este enfoque permite una medicina más personalizada, donde las intervenciones preventivas y los tratamientos pueden adaptarse específicamente a las necesidades individuales de cada paciente.

La capacidad de la IA para prever enfermedades no solo se limita a las condiciones genéticas, sino que también puede extenderse a enfermedades crónicas y otras condiciones de salud, proporcionando una herramienta poderosa para la medicina preventiva.

Al identificar riesgos de manera temprana, los sistemas de IA pueden facilitar la implementación de estrategias de prevención y monitoreo más efectivas, mejorando así los resultados de salud a largo plazo.

Este avance subraya el potencial transformador de la IA en la medicina, permitiendo a los profesionales de la salud adelantarse a los problemas y abordar las enfermedades antes de que se manifiesten plenamente¹⁰.

3. Desarrollo de Medicamentos: La IA acelera el proceso de descubrimiento de fármacos, identificando posibles candidatos más rápidamente que los métodos tradicionales.

Utilizando algoritmos avanzados y técnicas de aprendizaje automático, la IA puede analizar vastas cantidades de datos biomédicos, genéticos y químicos en un tiempo significativamente menor. Esto permite a los investigadores identificar compuestos potencialmente efectivos con mayor rapidez y precisión.

La IA puede simular cómo diferentes fármacos interactuarán con objetivos biológicos específicos, previendo posibles efectos secundarios y optimizando las propiedades farmacológicas desde las primeras etapas del desarrollo¹¹. Este enfoque no solo reduce el tiempo necesario para llevar un medicamento desde el laboratorio hasta los ensayos clínicos, sino que también disminuye los costos asociados al proceso de investigación y desarrollo.

La capacidad de la IA para identificar patrones complejos en los datos también facilita la redirección de medicamentos existentes para nuevas indicaciones, maximizando el uso de compuestos ya aprobados y seguros.

En resumen, la integración de la inteligencia artificial (IA) en el descubrimiento de fármacos representa una revolución en la biomedicina, acelerando la innovación y mejorando la eficiencia de una industria crucial para la salud global.

La IA está transformando la industria farmacéutica al acelerar la innovación y mejorar la eficiencia en la búsqueda de tratamientos para enfermedades concretas. Su capacidad para procesar grandes volúmenes de datos biomédicos, identificar nuevos objetivos terapéuticos y diseñar compuestos químicos específicos es fundamental para avanzar en la medicina personalizada y abordar enfermedades complejas¹².

Al analizar patrones en datos genómicos, proteómicos y clínicos, la IA puede prever interacciones moleculares, optimizar propiedades farmacológicas y sugerir combinaciones terapéuticas prometedoras, todo ello en un tiempo récord y con una precisión sin precedentes.

Este enfoque innovador no solo reduce los costos y el tiempo de desarrollo de nuevos medicamentos, sino que también abre nuevas posibilidades para el tratamiento de enfermedades antes consideradas intratables o enfermedades raras.

En última instancia, la IA está desempeñando un papel crucial en la mejora de la salud global al facilitar el descubrimiento y desarrollo de terapias más efectivas y personalizadas¹³.

4. Atención Personalizada: Las aplicaciones de inteligencia artificial (IA) han revolucionado la medicina al poder diseñar planes de tratamiento individualizados, lo que optimiza la eficacia terapéutica y minimiza los efectos secundarios.

Esta capacidad de la IA para analizar datos clínicos, biomédicos y genéticos permite la personaliza-

La IA mejora el diagnóstico y tratamiento médico, pero plantea desafíos éticos como la privacidad de los datos

ción de los tratamientos en función de las características específicas de cada paciente, como la genética, el historial médico y las preferencias individuales.

Al procesar grandes cantidades de información, la IA puede identificar patrones y correlaciones que pueden no ser evidentes para los profesionales de la salud, lo que conduce a recomendaciones de tratamiento más precisas y efectivas. Además, la IA puede prever posibles reacciones adversas y ajustar los planes de tratamiento en consecuencia, mejorando así la seguridad y la tolerabilidad de los diversos fármacos.

En última instancia, esta capacidad de la IA para diseñar planes de tratamiento individualizados representa un avance significativo en la medicina personalizada, ofreciendo un enfoque más preciso y eficiente para el cuidado de la salud¹⁴.

La automatización de tareas administrativas y rutinarias ofrece una oportunidad para que los profesionales de la salud dediquen más tiempo a la interacción directa con los pacientes, lo que contribuye a una atención más centrada en el individuo. Además, esta automatización puede prevenir errores de identificación de pacientes, mejorando la seguridad y la eficiencia en los procesos médicos.

En última instancia, esta redistribución de responsabilidades puede fortalecer la relación médico-paciente, permitiendo un enfoque más completo y personalizado en la atención sanitaria¹⁵.

5. Tratamiento y rehabilitación de pacientes.

Los robots asistenciales pueden ayudar con tareas de movilidad, administración de medicamentos y terapias, aliviando la carga de trabajo de los profesionales sanitarios y brindando un apoyo más personalizado a los pacientes.

Asimismo, los avances en neurociencia y neurotecnología, como las interfaces cerebro-computadora, abren la puerta a intervenciones terapéuticas que podrían mejorar o restaurar funciones cognitivas y motrices.

6. Asistencia en Cirugía: Los robots quirúrgicos guiados por inteligencia artificial (IA) representan una innovación revolucionaria en el campo de la cirugía al ofrecer una precisión sin precedentes y al mismo tiempo reducir los riesgos y los tiempos de recuperación .

Estos sistemas avanzados pueden realizar procedimientos quirúrgicos complejos con una precisión milimétrica, superando las limitaciones de la destreza manual humana y minimizando el riesgo de errores.

Gracias a la capacidad de la IA para analizar datos en tiempo real y adaptar los movimientos quirúrgicos en función de la anatomía específica del paciente, los robots quirúrgicos pueden realizar intervenciones con una precisión increíble, lo que reduce el riesgo de daño a tejidos circundantes y acelera la recuperación postoperatoria.

Además, estos sistemas pueden proporcionar retroalimentación instantánea al cirujano, permitiéndoles tomar decisiones informadas y ajustar su enfoque durante la cirugía¹⁷.

En consecuencia, los pacientes experimentan tiempos de recuperación más cortos, menor dolor postoperatorio y menos complicaciones, lo que mejora significativamente los resultados y la experiencia quirúrgica en general.

La combinación de la precisión robótica y la inteligencia artificial no solo está transformando la práctica quirúrgica, sino que también está mejorando la calidad de vida de los pacientes al ofrecer procedimientos más seguros, eficientes y efectivos.

14. Hamburg, M. A., Collins, F.S. (2018) Medicina personalizada y IA: desafíos y oportunidades. *Science Translational Medicine*.

15. Frisoni, G.B., Redolfi, F., Mansour, A., Toga, A., Evans, P. (2020) El futuro de la atención médica personalizada con inteligencia artificial. *European Journal of Nuclear Medicine and Molecular Imaging*.

16. Patel, S.N., Wright, J.J. (2018) Cirugía robótica: aplicaciones actuales y futuras. *Journal of Robotic Surgery*.

17. Hashimoto, K., Rosman, D., Rus, C., Meireles, L.G. (2018) Inteligencia artificial en la cirugía robótica: Promesas y peligros. *Annals of Surgery*.

18. Hinton, D. J., Kunz, R.M., Sweeney, J.L. (2019) Inteligencia artificial en el cuidado de la salud: anticipando desafíos éticos, de privacidad, sesgos y confianza. *Journal of Medical Ethics*.

19. Moorhouse, L. J., Yampolskiy, M. S. (2018) Privacidad y consentimiento informado en la era de la IA en salud. *Journal of Health and Biomedical Law*.

20. Richards, E.P. Rathbun, L.B. (2020) Consideraciones bioéticas en inteligencia artificial y cuidado de la salud: privacidad, consentimiento y equidad. *Cambridge Quarterly of Healthcare Ethics*.

21. García, A. (2020) Ética y equidad en el acceso a las tecnologías de salud: el caso de la inteligencia artificial. *Revista de Bioética y Derecho*.

22. López, P., Rodríguez, J. (2018) Inteligencia artificial y equidad en la atención sanitaria: desafíos y propuestas. *Cuadernos de Bioética*.

23. Gómez, F. (2019) Bioética y tecnología en salud: una visión para la equidad. *Revista Iberoamericana de Bioética*.

3/

Implicaciones bioéticas en el uso de IA en la salud.

1. Consentimiento informado y privacidad.

El uso de IA en salud requiere el manejo de datos personales sensibles, lo que plantea serias preocupaciones sobre la privacidad y el consentimiento informado. Los pacientes deben ser plenamente conscientes de cómo se recopilan, almacenan y utilizan sus datos, y deben tener la oportunidad de consentir o rechazar su uso de manera informada y voluntaria. Esto es crucial en un contexto donde los algoritmos pueden tomar decisiones que afectan significativamente la vida de una persona, desde el diagnóstico hasta el tratamiento¹⁸.

La bioética nos llama a asegurar que los principios de autonomía y consentimiento informado sean rigurosamente respetados, lo que implica proporcionar a los pacientes información clara y comprensible sobre los usos potenciales de sus datos y los riesgos asociados. Además, se debe garantizar que los datos de los pacientes se manejen con el mayor respeto a su privacidad, implementando medidas de seguridad robustas para proteger la información contra accesos no autorizados y usos indebidos y evitar abusos y brechas de seguridad¹⁹.

La transparencia en los procesos y la posibilidad de que los pacientes controlen sus propios datos son fundamentales para mantener la confianza y la integridad en el uso de la IA en salud²⁰.

2. Justicia y equidad en la atención sanitaria.

Otro aspecto crucial es la justicia en el acceso y la distribución de las tecnologías de IA en salud. Existe el riesgo significativo de que las ventajas de la IA se concentren en los sistemas de salud más avanzados y sean accesibles únicamente para quienes pueden pagarlas, ampliando así

las brechas existentes en la atención sanitaria y creando disparidades aún mayores²¹.

La bioética debe abordar este problema de manera exhaustiva, buscando estrategias para implementar estas tecnologías de forma equitativa. Es esencial garantizar que todos los pacientes, independientemente de su origen socioeconómico, tengan acceso a los beneficios de la IA, lo cual implica desarrollar políticas y marcos reguladores que promuevan la distribución justa y accesible de las tecnologías de IA, asegurando que no se conviertan en un privilegio de pocos sino en un recurso disponible para todos²².

Además, se deben considerar subvenciones y programas de apoyo para los sistemas de salud menos desarrollados, de modo que puedan integrar estas innovaciones sin sufrir cargas financieras insostenibles. La equidad en el acceso a la IA en salud es fundamental para evitar la perpetuación de desigualdades y para asegurar que todos los individuos puedan beneficiarse de los avances tecnológicos en el cuidado de su salud²³.

3. Humanización de la atención al paciente.

La tecnología no debe ser un sustituto de la relación humana en la atención sanitaria. La empatía, la compasión y el contacto humano son componentes esenciales del cuidado de la salud que ninguna máquina puede replicar completamente²⁴.

La interacción cara a cara, el lenguaje corporal y el tono de voz son elementos cruciales para construir una relación de confianza y comprensión entre el paciente y el profesional de la salud.

La bioética debe asegurarse de que, mientras se integran tecnologías avanzadas, los profesionales de la salud continúen priorizando la humanización de la atención.

Esto significa que la tecnología debe ser vista como una herramienta para complementar y mejorar, no para reemplazar, la interacción humana. Los profesionales deben mantener una comunicación clara, comprensiva y personal con

los pacientes, tomándose el tiempo necesario para explicar diagnósticos y tratamientos, y para responder a preguntas y preocupaciones.

Es fundamental que los pacientes se sientan escuchados y valorados, y que sus experiencias y emociones sean consideradas parte integral del proceso de atención.

La humanización de la atención implica también una sensibilidad cultural y personal, reconociendo y respetando las diversas perspectivas y necesidades de cada paciente. Así, mientras las tecnologías avanzadas pueden aumentar la eficiencia y la precisión, es imperativo que no se pierda de vista la esencia humana de la práctica médica, que es fundamental para la calidad del cuidado y el bienestar del paciente.

4. Responsabilidad y rendición de cuentas.

La responsabilidad es fundamental. Cuando una IA toma una decisión errónea que resulta en un daño al paciente, surge la pregunta crítica de quién es responsable: ¿el desarrollador del algoritmo que programó la IA, el profesional de la salud que utilizó la herramienta, o la institución que implementó la tecnología? Esta cuestión de responsabilidad no es solo teórica, sino que tiene profundas implicaciones prácticas y éticas²⁵.

Los algoritmos de IA a menudo son complejos y funcionan como cajas negras, donde ni siquiera los desarrolladores pueden explicar completamente cómo se toman las decisiones. La falta de transparencia puede erosionar la confianza de los pacientes y los profesionales de la salud. Es esencial desarrollar IA que sea explicable y transparente, permitiendo a los usuarios comprender y confiar en sus resultados²⁶.

La bioética debe desarrollar marcos claros y específicos para la rendición de cuentas en el uso de IA en la salud, asegurando que siempre haya una supervisión humana activa y que las responsabilidades estén bien definidas en cada etapa del proceso. Esto incluye establecer regulaciones y directrices precisas que determinen cómo se debe documentar el uso de IA, cómo

se deben comunicar y gestionar los errores, y cómo se deben investigar y resolver los incidentes adversos²⁷.

Es esencial que **los desarrolladores** de algoritmos se responsabilicen por la calidad y la seguridad de sus productos, asegurando que estos cumplan con los estándares éticos y técnicos más rigurosos.

Asimismo, **los profesionales** de la salud deben ser competentes en el uso de estas tecnologías y conscientes de sus limitaciones, manteniendo siempre un juicio clínico informado.

Las instituciones, por su parte, deben garantizar una implementación segura y efectiva de la IA, proporcionando capacitación adecuada y estableciendo procedimientos claros para la supervisión y el manejo de errores.

La existencia de un marco robusto de responsabilidad no solo protege a los pacientes, sino que también promueve la confianza en el uso de IA en la medicina, facilitando una integración más ética y efectiva de estas tecnologías en la atención sanitaria.

4/

Dilemas bioéticos en la era de la IA respecto a la Atención al Paciente.

La bioética, como disciplina fundamental en la ética médica, se basa en una serie de principios fundamentales, entre los que destacan la **autonomía del paciente**, la **beneficencia**, la **no maleficencia** y la **justicia**. Estos pilares éticos establecen un marco normativo crucial para la toma de decisiones en el ámbito de la salud.

24. López, J., Martínez, A. (2020) La humanización de la medicina en la era de la tecnología. *Revista de Humanidades Médicas*.

25. Díaz, C. Martínez, J. (2020) Marco ético para la rendición de cuentas en el uso de inteligencia artificial en medicina. *Revista de Bioética y Salud Pública*.

26. Rodríguez, J. Pérez, M. (2019) Ética y responsabilidad en la implementación de inteligencia artificial en salud. *Revista de Bioética y Derecho*.

27. Sánchez, M. Rodríguez, L. (2021) Rendición de cuentas y responsabilidad en el uso de inteligencia artificial en la atención sanitaria. *Revista de Ética en Salud Pública*.

28. Pérez, J. Sánchez, M. (2020) La autonomía del paciente en la era de la inteligencia artificial. *Revista de Bioética y Salud Pública*.

29. Ruiz, E., Torres, M. (2018) Inteligencia artificial y autonomía del paciente: una revisión ética. *Revista de Ética Biomédica*.

30. Sánchez, A. Martínez, J. (2021) Autonomía del paciente y toma de decisiones informadas en la era de la inteligencia artificial. *Revista de Bioética y Derecho*.

31. Gómez, L. García, A. (2019) Integración de la inteligencia artificial en la medicina: implicaciones éticas para la beneficencia y no maleficencia. *Revista de Ética Médica y Bioética*.

32. Pérez, J. Sánchez, M. (2020) Ética médica y uso de inteligencia artificial: promoviendo la beneficencia y no maleficencia. *Revista de Bioética y Salud Pública*.

La introducción de la inteligencia artificial (IA) en la atención médica exige un análisis exhaustivo de estos principios, ya que plantea nuevas interrogantes éticas y desafíos inéditos.

La capacidad de la IA para analizar grandes volúmenes de datos y tomar decisiones autónomas, así como su influencia en el diagnóstico, tratamiento y gestión de la atención al paciente, requiere una evaluación cuidadosa de cómo estos avances tecnológicos pueden afectar la **autonomía del individuo**, el bienestar general, la prevención de daños y la equidad en el acceso a la atención sanitaria.

En este sentido, la bioética desempeña un papel crucial al proporcionar un marco ético sólido para guiar el desarrollo, la implementación y el uso ético de la IA en la medicina, asegurando que se respeten los valores fundamentales de la ética médica en un contexto de avances tecnológicos acelerados:

1. Autonomía del Paciente. La autonomía del paciente es un principio fundamental en la ética médica que reconoce el derecho de los individuos a tomar decisiones informadas y autónomas sobre su propia salud.

La introducción de la inteligencia artificial (IA) en el ámbito de la salud puede potenciar este principio al ofrecer información detallada y personalizada que capacita a los pacientes para tomar decisiones más fundamentadas y alineadas con sus necesidades y valores individuales.

No obstante, es importante reconocer que la integración de la IA también conlleva riesgos, como la posibilidad de que los pacientes se sientan abrumados o alienados por la complejidad tecnológica²⁸.

Es esencial que los sistemas de IA sean diseñados con un enfoque centrado en el usuario, priorizando la transparencia y la comprensibilidad. Esto implica que la IA debe ser capaz de explicar claramente sus recomendaciones y procesos de toma de decisiones, así como ofrecer

opciones claras y comprensibles para que los pacientes participen activamente en sus cuidados médicos²⁹.

Al fomentar una relación colaborativa entre la IA y los pacientes, se promueve una toma de decisiones compartida que respeta la autonomía individual y fortalece la confianza en la tecnología médica. En última instancia, el objetivo es que la IA no solo complemente, sino que también empodere la capacidad de autogestión y toma de decisiones de los pacientes, mejorando así la calidad y la personalización de la atención médica³⁰.

2. Beneficencia y No Maleficencia. Los principios de beneficencia y no maleficencia constituyen pilares fundamentales en la ética médica, instando a los profesionales de la salud a buscar el beneficio del paciente y a evitar causarle daño³¹.

La integración de la inteligencia artificial (IA) en la práctica clínica promete mejorar estos principios al proporcionar diagnósticos más precisos y opciones de tratamiento más efectivas.

La capacidad de la IA para analizar grandes conjuntos de datos y reconocer patrones complejos puede llevar a una identificación temprana de enfermedades y a la personalización de intervenciones médicas, lo que potencialmente mejora los resultados de salud y reduce el riesgo de efectos adversos. Sin embargo, es importante reconocer que la dependencia excesiva en la tecnología también conlleva riesgos³².

Los sistemas de IA no están exentos de errores y pueden malinterpretar información, lo que podría resultar en decisiones incorrectas o en la omisión de aspectos importantes en la atención al paciente. Por lo tanto, es esencial mantener un equilibrio entre la tecnología y el juicio clínico humano.

Los profesionales de la salud deben mantenerse vigilantes y críticos en la interpretación de las recomendaciones de la IA, respaldándose en su experiencia y conocimiento clínico para tomar decisiones informadas y éticas.

El uso de IA en salud requiere equilibrar tecnología y juicio clínico para respetar principios éticos

Se debe proporcionar una capacitación continuada en el uso de la IA para garantizar una integración segura y efectiva en la práctica clínica. En última instancia, la colaboración entre la IA y el juicio clínico humano puede optimizar la atención al paciente, maximizando los beneficios de la tecnología mientras se minimizan los riesgos potenciales para la salud.

3. Justicia. El principio de justicia en el ámbito de la salud busca garantizar una distribución equitativa de los recursos, asegurando que todos tengan acceso a la atención médica que necesitan.

La inteligencia artificial (IA) se presenta como una herramienta poderosa en la búsqueda de este objetivo al ofrecer la posibilidad de reducir las disparidades y aumentar la accesibilidad a tecnologías avanzadas en salud.

La IA puede mejorar el acceso a la atención médica al hacer que los servicios sean más eficientes y asequibles, lo que beneficia especialmente a aquellos que de otro modo podrían enfrentar barreras para obtener atención médica de calidad. Sin embargo, es crucial reconocer que la implementación de la IA puede también perpetuar o incluso ampliar las desigualdades existentes si no se aborda de manera adecuada.

Por ejemplo, si las tecnologías de IA se implementan principalmente en entornos de atención médica privada o en áreas urbanas prósperas, podrían dejar rezagados a grupos marginados o a comunidades rurales con recursos limitados³⁴.

Por lo tanto, es esencial que los desarrollos en IA se realicen de manera inclusiva y equitativa, teniendo en cuenta las necesidades y contextos diversos de toda la población. Esto implica no solo garantizar que la tecnología esté disponible y sea accesible para todos, sino también adaptarla para satisfacer las necesidades específicas de grupos subrepresentados o marginados.

Además, se requiere una colaboración activa entre los sectores público y privado, así como la participación de la comunidad, para desarrollar

e implementar soluciones de IA que sean verdaderamente inclusivas y que aborden las inequidades existentes en el acceso a la atención médica³⁵.

En resumen, el desafío radica en aprovechar el potencial de la IA para mejorar la equidad en la salud, al mismo tiempo que se mitigan los riesgos de aumentar las disparidades, garantizando así que todos los individuos, especialmente los más vulnerables, se beneficien de manera justa de los avances tecnológicos en el cuidado de la salud.

5/

Conclusión.

La inteligencia artificial (IA) se erige como una herramienta poderosa con un potencial transformador que abarca todos los aspectos de nuestra vida. Desde la atención médica hasta la movilidad urbana y la gestión de recursos, su influencia es innegable. Sin embargo, con este poder también vienen responsabilidades éticas que deben ser abordadas con seriedad y atención.

Mientras que estos avances tecnológicos ofrecen un sinfín de posibilidades para mejorar la calidad y eficiencia de los servicios de salud, también plantean importantes desafíos éticos que debemos abordar con cautela y profunda reflexión³⁶.

La aplicación de la IA en la sociedad contemporánea nos enfrenta a una serie de dilemas éticos y bioéticos que no pueden ser pasados por alto. Desde cuestiones de privacidad y consentimiento informado hasta la equidad en el acceso y la responsabilidad por errores, estos desafíos exigen una reflexión profunda y un enfoque ético en su desarrollo y aplicación.

Es imperativo reconocer que, si bien la IA ofrece promesas de avances significativos, su implemen-

33. "Ruiz, E. Torres, M. (2018) Principios éticos y uso de inteligencia artificial en la atención sanitaria: garantizando beneficencia y no maleficencia. Revista de Ética Biomédica.

34. Martínez, A. Sánchez, E. (2021) Justicia y equidad en la salud: desafíos y oportunidades en la era de la inteligencia artificial. Revista de Ética Biomédica.

35. Sánchez, E. Rodríguez, L. (2019) Justicia y accesibilidad en salud con inteligencia artificial. Revista de Psicología y Salud Mental.

36. García, J. Rodríguez, L. (2020) Inteligencia artificial y bioética: reflexiones éticas sobre el impacto en la práctica clínica. Revista de Ética Médica y Bioética.

37. Martínez, E. Gutiérrez, R. (2019) Aspectos éticos de la inteligencia artificial en medicina. Revista de Bioética y Tecnología Médica.

38. Orden Hospitalaria San Juan de Dios. (2023) Código Ético de la Orden Hospitalaria de San Juan de Dios en a Labor Hospitalaria (vol. LIV núm. 335 pp. 14-24) <https://www.laborhospitalaria.com/codigo-etico-de-la-orden-hospitalaria-de-san-juan-de-dios/>

39. García, A. Pérez, M. (2018) La dimensión ética en la atención sanitaria: El legado de San Juan de Dios. Revista de Bioética y Ética Médica.

40. Martínez, E. Gómez, A. (2017) El cuidado centrado en la persona: Una perspectiva desde la Orden Hospitalaria de San Juan de Dios. Revista de Ética en Ciencias de la Salud.

tación debe ser acompañada por una sólida base ética que garantice su uso para el bien común.

La bioética emerge como el marco idóneo para guiar estas decisiones, colocando a la humanidad y sus valores fundamentales en el centro de la discusión. Desde principios como la autonomía, la beneficencia, la no maleficencia y la justicia, hasta consideraciones más específicas sobre la equidad en el acceso y la responsabilidad por el uso adecuado de los datos, la bioética ofrece un terreno fértil para la reflexión y el debate.

En este sentido, el diálogo abierto y la reflexión profunda sobre los dilemas éticos y bioéticos asociados con la IA son esenciales para garantizar su desarrollo y aplicación ética. Solo al comprometernos activamente con estos temas y al considerar cuidadosamente las implicaciones éticas de nuestras acciones podemos asegurarnos de que la IA se convierta en una fuerza positiva para el progreso humano. Es crucial que, como comunidad hospitalaria, reflexionemos profundamente, ponderando cuidadosamente los beneficios y riesgos de estas tecnologías. Debemos promover un enfoque de desarrollo e implementación de la IA que anteponga la dignidad y el bienestar del ser humano por encima de cualquier otro interés.

En última instancia, al abrazar un enfoque ético en la integración de la IA en nuestras vidas, podemos trabajar hacia un futuro en el que la tecnología y la humanidad coexistan en armonía, beneficiando a todos los miembros de la sociedad³⁷.

Como miembros de la Orden Hospitalaria de San Juan de Dios, nuestra misión de cuidar del prójimo y preservar la dignidad humana nos exhorta a examinar detenidamente cómo estas innovaciones pueden impactar en los principios fundamentales de la bioética³⁸.

Como Hermanos de San Juan de Dios, nuestra vocación nos llama a brindar una atención integral que vaya más allá de lo meramente técnico o procedimental. Debemos asegurarnos

de que, incluso en medio de la revolución digital, los pacientes sigan sintiéndose acompañados, escuchados y comprendidos en su sufrimiento. La calidez humana, la empatía y la compasión deben permanecer como piedras angulares de nuestra labor asistencial.

En este sentido, la IA puede ser una herramienta valiosa para liberar a los profesionales de tareas administrativas y logísticas, permitiéndoles dedicar más tiempo y atención a los aspectos relacionales y emocionales de la atención sanitaria³⁹.

Los sistemas de IA pueden y deben ayudar a personalizar y adaptar los cuidados a las necesidades específicas de cada paciente, fortaleciendo así la atención centrada en la persona.

Los avances de la inteligencia artificial en el campo de la salud representan un desafío ético de gran magnitud para nuestra Orden y para toda la comunidad médica y sanitaria. Si bien estas tecnologías ofrecen un sinfín de posibilidades para mejorar la calidad y eficiencia de los cuidados, también plantean serias interrogantes sobre cómo salvaguardar los principios fundamentales de la bioética y preservar la dimensión humana en la atención a los pacientes.

Como Hermanos de San Juan de Dios, tenemos la responsabilidad de participar activamente en este debate, aportando nuestra voz y nuestra experiencia en la construcción de un futuro sanitario donde la tecnología se ponga al servicio de la dignidad y el bienestar de la persona. Debemos velar por que el desarrollo y la aplicación de la IA en el ámbito de la salud se rijan por una sólida ética del cuidado, que anteponga siempre el respeto a la vida, la autonomía y la justicia⁴⁰.

Sólo así podremos honrar la misión que nos ha encomendado nuestro Padre San Juan de Dios: acompañar y cuidar del prójimo con amor, compasión y una entrega total.

Bibliografía

Aggarwal, H. K., Jebara, S. T., & Pugh, T. W. (2020).

El impacto de la inteligencia artificial en los resultados quirúrgicos: una revisión.
International Journal of Surgery.

Aguilar, J. (2011).

Inteligencia Artificial: Fundamentos, Métodos y Aplicaciones.
Alfaomega.

Alonso Betanzos, A., & Hernández Sande, R. (2002).

Introducción a las Redes Neuronales Artificiales.
Centro de Investigaciones Biomédicas de Galicia (CIBER).

Alonso Betanzos, A., Fontenla-Romero, O., & Guijarro-Berdiñas, B. (2011).

Aprendizaje Automático: Conceptos y algoritmos.
Ra-Ma.

Appelbaum, P. S., & Grisso, J. R. (2019).

Garantizar la privacidad del paciente y el consentimiento informado en el uso de la IA para el cuidado de la salud.
American Journal of Bioethics.

Floridi, L. (2019).

La Revolución de la Inteligencia Artificial: Perspectivas Éticas y Filosóficas.
Editorial Tecnos.

Floridi, L. (2019).

La Revolución de la Inteligencia Artificial: Perspectivas Éticas y Filosóficas.
Editorial Tecnos.

Chen, H., Engkvist, O., Wang, Y., Olivecrona, M., & Blaschke, T. (2018).

Aprendizaje automático en el descubrimiento de fármacos: una revisión.
Drug Discovery T

De Miguel Beriain, I. (2018).

Ética de la Inteligencia Artificial y Robótica.
Editorial Tecnos.

Díaz, C., & Martínez, J. (2020).

Equidad en la atención médica: desafíos y estrategias con la inteligencia artificial.
Revista de Salud Digital y Ética Médica.

Díaz, C., & Martínez, J. (2020).

Marco ético para la rendición de cuentas en el uso de inteligencia artificial en medicina.
Revista de Bioética y Salud Pública.

Díaz, C., & Rodríguez, J. (2019).

Ética y uso de inteligencia artificial en la práctica clínica: perspectivas para la beneficencia y no maleficencia.
Revista de Salud Digital y Ética Médica.

Díaz de María, F., & Albiol Colomer, A. (2004).

Redes Neuronales y Sistemas Difusos.
Prentice Hall.

▶ Elton, J., Alcocer-Cuarón, J., González, C., & Quintana, D. (2019).

El papel de la inteligencia artificial en el descubrimiento de fármacos.
BMC Bioinformatics.

▶ Fernández, R., & Gutiérrez, M. (2020).

Ética y tecnología médica: promoviendo la beneficencia y no maleficencia en la era de la inteligencia artificial.
Revista de Ética Clínica.

▶ Fernández, G., & Martín, E. (2016).

Consideraciones éticas en el desarrollo y aplicación de inteligencia artificial en salud.
Revista de Bioética y Derecho Sanitario.

▶ Fernández, R., & Torres, M. (2020).

Humanización de la atención: más allá de la tecnología.
Revista de Psicología y Salud Mental.

▶ Floridi, L. (2019).

La Revolución de la Inteligencia Artificial: Perspectivas Éticas y Filosóficas.
Editorial Tecnos.

▶ García, L., & Rodríguez, E. (2019).

San Juan de Dios y la ética del cuidado: Lecciones para la medicina moderna.
Revista de Bioética y Teología.

▶ Gellert, R., Jaspers, E. T., & Weller, D. S. (2018).

Protección de datos, privacidad y cuestiones éticas en IA y Big Data: aplicaciones en el cuidado de la salud.
European Data Protection Law Review.

▶ Gómez, L., & García, A. (2019).

Equidad en el acceso a la salud: desafíos y oportunidades con la inteligencia artificial.
Revista de Ética Médica y Bioética.

▶ Gómez, C., & Sánchez, L. (2018).

Tecnología y humanización en la atención sanitaria: un equilibrio necesario.
Revista de Medicina y Ética.

▶ Goodfellow, I., Bengio, Y., & Courville, A. (2016).

Deep Learning.
MIT Press.

▶ Gutiérrez, R., & López, M. (2021).

San Juan de Dios: El precursor de la humanización en la atención médica.
Revista de Bioética y Ética Médica.

▶ Haykin, S. (2009).

Neural Networks and Learning Machines (3rd ed.). Pearson.

▶ Kourou, C. K., Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V., & Fotiadis, D. I. (2019).

IA en la medicina personalizada: predicción de la respuesta al tratamiento específica del paciente.
Briefings in Bioinformatics.

▶ Marchant, G. E., Gostin, C. J., & Roberts, S. D. T. (2020).

Ética de la inteligencia artificial en medicina.
The Hastings Center Report.

▶ **Martínez, A., & Pérez, J. (2020).**

La herencia de San Juan de Dios en la práctica médica contemporánea.
Revista de Ética en Ciencias de la Salud.

▶ **Martínez, A., & Sánchez, E. (2020).**

Responsabilidad y transparencia en el uso de inteligencia artificial en medicina.
Revista de Ética Biomédica.

▶ **Mittelstadt, A., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2019).**

Transparencia, confianza y consentimiento informado en la era de la inteligencia artificial en el cuidado de la salud.
Journal of Information, Communication and Ethics in Society.

▶ **Montero Martínez, S. (2011).**

Redes Neuronales Artificiales: Fundamentos, modelos y aplicaciones.
Paraninfo.

▶ **Murphy, K. P. (2012).**

Machine Learning: A Probabilistic Perspective.
MIT Press.

▶ **Nebeker, S., Torous, L., & Bartlett, M. (2019).**

IA y las implicaciones éticas para la confianza, privacidad y consentimiento en el cuidado de la salud.
Digital Health.

▶ **Okamura, A. M., Bowyer, M. G., & Hasser, C. E. (2019).**

Avances en inteligencia artificial y robótica en la práctica quirúrgica.
Surgical Clinics of Nor





03/

Retos de la introducción

de la inteligencia artificial en la relación asistencial. ¿Hacia una medicina sin rostro?

Bernabé Robles del Olmo,

Neurólogo y máster en Bioética. Coordinador Unidad de Bioética.
Parc Sanitari Sant Joan de Déu. Sant Boi de Llobregat (Barcelona).
Vicepresidente Comité de Ética de la Investigación con Medicamentos (CEIM).
Fundació de Recerca Sant Joan de Déu. Eplugues de Llobregat (Barcelona).

Cuando se aborda el tema de la potencial deshumanización de la asistencia sanitaria, una lectura superficial nos puede llevar a pensar que “todo es culpa de la inteligencia artificial (IA)”. La primera reflexión a plantear sería, por tanto, si la digitalización de la medicina es la causa preponderante, o detonante, de la pérdida progresiva de calidez y calidad en la relación clínica. Pero es posible que la respuesta no sea tan simple, y que el reto más “inteligente” sea imaginar cómo, con qué requisitos, puede ayudarnos la IA a recuperar la calidez y la calidad en el trato.

Palabras clave: Inteligencia Artificial, Humanización, Asistencia.

When reflecting on the potential dehumanisation of health care, a cursory consideration of the subject may lead one to think that “artificial intelligence (AI) is to blame for it all”. Accordingly, the first aspect to be weighed is whether the digitalisation of medicine is actually the chief cause or mainspring of the progressive loss of warmth and quality in the clinical relationship. However, the answer to this question may not be as simple as it may seem, and the most “intelligent” challenge may well be to imagine how and with what prerequisites AI can help us to recover the warmth and quality of treatment.

Key words: Artificial Intelligence, Humanisation, Care.

1/

¿Humanizar? Calidad y calidez.

Cuando hablamos de humanizar el cuidado de la salud, nos referimos a un trato cálido, respetuoso y de calidad. Pero el hecho de que un acto o una decisión sea producto de humanos no es garantía de trato “**humanizado**”. ¿Todo acto humano es deseable, incluso tolerable? Si repasamos la historia de los Sapiens y, desafortunadamente, nuestro presente, cuesta contestar con un sí categórico.

La “**humanización**” a cultivar en la atención a la salud es aquella que mejora nuestra hospitalidad, nuestro interés por el otro y por lo otro, por ese mundo irrepetible que tenemos delante. Hablamos pues de respeto y responsabilidad, de aquel trato que todos querríamos recibir cuando sufrimos. Se trataría de no dejar nunca de ser sujeto para convertirnos en simple objeto, o peor, en dato. Desearíamos ser, cuando desgraciadamente nos toque, “**persona enferma**”, y no mera “**enfermedad**”, sin mirada.

Se trata pues de construir confianza en la alteridad, viendo al otro como alguien que, entre otras muchas otras dimensiones de su persona, padece y necesita ayuda. Esto no nos debería resultar difícil, dado que estamos neurobiológi-

camente cableados para ello, con el sistema de neuronas espejo ([Rizzolatti & Cattaneo, 2009](#)).

La pregunta que queremos abordar sería pues: ¿puede alinearse la IA con un cuidado humanizado? Achacar la deshumanización completamente a la IA sería seguramente un análisis simplista y poco introspectivo. La atención sanitaria comenzó a “**deshumanizarse**” mucho antes de la implantación acelerada de los algoritmos y de los detectores de patrones en la práctica habitual.

2/

Una breve historia de la relación asistencial.

Las actividades de cuidado entre humanos seguían inicialmente un modelo básicamente espiritual, orientado a confortar ante el sufrimiento y la incertidumbre, dado que nuestro margen de maniobra era muy limitado cuando se trataba de reconducir fenómenos naturales. Las funciones de chamán y curandero coincidían a menudo en la misma persona.

Durante milenios, probablemente hasta finales del siglo XIX, el paradigma no cambió demasiado. La atención psicológica, espiritual y paliativa no son inventos recientes. Salvo alguna intervención resolutiva de cirugía menor en la superficie del cuerpo (no se disponía de anestesia general), y la relevante tarea de tranquilizar cuando la patología reconocida tenía una esperable buena evolución espontánea, la atención paliativa (láudano y acompañamiento a pie de cama) era el modelo asistencial habitual cuando las cosas no iban bien.

Con la aparición de antisépticos, anestésicos generales y antibióticos, el poder de intervenir sobre el cuerpo se incrementó exponencialmente.

Como el cuerpo era más abordable, y más explo-
rable por los avances paralelos del conocimiento
científico, se impuso un modelo mal llamado
biomédico. Quizá deberíamos llamarlo más bien
mecanicista, dado que entendía la Medicina
como un servicio de reparación de cuerpos.
El poder que visualizó entonces la sociedad en
las ciencias médicas se hizo tan intenso que ya
no íbamos nosotros a ver a los enfermos, sino
que venían ellos a vernos a templos de salud que
construimos para concentrar los cuidados, los
profesionales y la tecnología necesaria para pro-
veerlos. En este contexto, la mesa de la consulta
se convirtió en instrumento, pero también en
una primera barrera, porque marcaba ya una
distancia en el trato interpersonal. Luego, poco
a poco, llenamos la mesa y sus inmediaciones de
aparatos (fonendo, negatoscopio, ecografía, te-
léfono, ordenador, móvil, etc.). Todos ellos dis-
traían la mirada y, potencialmente, interrumpían
el encuentro interpersonal que suponía, en
origen, la relación clínica.

Además, las necesidades de soporte de las per-
sonas atendidas y las obligaciones formativas de
un sistema “de éxito”, comenzaron a llenar tam-
bién la sala de gente (profesionales, estudiantes
y acompañantes), y de nuevas interrupciones.
El acto médico estaba dejando de ser un acto
íntimo, de confianza. La tecnología reducía
el contacto físico interpersonal, desactivando
nuestra red neuronal de la empatía. Así fuimos
perdiendo, unos más que otros, la mirada.

Hay que reconocer que el modelo biomédico ha
aportado muchos avances y ha mitigado mucho
sufrimiento físico, pero también ha generado
carencias complejas en un modelo asistencial
que debería enfocarse más en personas y valores
que en puras averías. Ante estas insuficiencias
se ha propuesto en el siglo XX el modelo biop-
sicosocial o, incluso, un modelo estrictamente
social, centrado en adaptar el entorno social a
la enfermedad, la discapacidad o el sufrimiento.

Fruto del aumento de la esperanza de vida deriva-
do la mejora de las condiciones sociales en ciertas
partes del globo, y del abordaje más resolutivo de

las patologías agudas, el crecimiento de los padeci-
mientos crónicos ha hecho patente la necesidad de
rediseñar el modelo hacia lo bio-psico-socio-espi-
ritual. Volvemos, de alguna forma, a los orígenes.
Ahora queremos abandonar, cuando menos par-
cialmente, las factorías de salud que construimos
durante los siglos XIX y XX para volver a cuidar
de cerca a las personas crónicamente enfermas, ya
no totalmente curables.

Por otro lado, la conquista social incuestionable
que supuso, aunque desgraciadamente no en
todo el globo, la universalización de la atención
sanitaria (deseada, justa y necesaria) generó un
volumen de actividad, tanto asistencial directa
como de soporte, que derivó en oportunidades
de negocio y de empleo. La salud comenzó a
convertirse en un “sector económico” haciendo
confluir intereses múltiples y diversos (comer-
ciales, culturales, industriales, administrativos,
curriculares, científicos, mediáticos, laborales,
etc.) sobre aquello que parecía tan simple en su
esencia: una relación de ayuda a quien sufre.

Las relaciones cuasicontractuales así generadas
compartimentan las tareas asistenciales en már-
genes poco líquidos y, por tanto, no te impli-
can en lo íntimo y personal de cada persona, y
pueden acabar desembocando en estrategias de
medicina defensiva, viendo en la persona aten-
dida una potencial amenaza. Una relación así
no te traspasa.

Es el falso confort (o el falso consuelo) de lo im-
personal. Vamos entendiendo poco a poco que
el mito tecno-científico en Medicina no es la
panacea, y que sobrevivir a una patología aguda
con una intervención técnica brillante genera,
afortunadamente, más trayectoria vital pero
también más oportunidades para padecer en-
fermedades, condiciones o trastornos no total-
mente curables, solo paliables. La cronicidad es,
de alguna forma, el resultado de éxitos sociales
y biomédicos, en aquellas partes del globo que
se los han podido permitir. Vamos entendiendo
que medicalizar la sociedad no será el remedio
para todo mal, y que esta medicalización tiene
sus propios riesgos y efectos adversos, siendo el

más evidente el de generar sufrimiento por el propio miedo a sufrir y por la hipervigilancia del estado de salud.

Además, profesionales, pacientes o ciudadanos pueden quedar fácilmente deslumbrados por la técnica, esperando de ella soluciones que no son esperables, ni desde el punto de vista científico ni desde la perspectiva humana. Se aplican recetas tecnológicas sin una justificación sólida y sin una reflexión pausada y rigurosa sobre pros y contras. Por si esto fuera poco, se evalúan poco los resultados obtenidos a medio y largo plazo.

Se impone también una reflexión sobre la dimensión que merece el “**diagnóstico**”. Diagnosticar significa “**conocer a través de**”. Quizá lo hemos convertido erróneamente en un fin absoluto, en el marco de la tecnificación en el modelo exclusivamente mecanicista que impulsaron los avances en el conocimiento del funcionamiento del cuerpo. Diagnosticar debería volver a ser lo que fue, un instrumento para interpretar y mitigar el sufrimiento vivido por las personas que nos piden ayuda. Si el diagnóstico no aporta más que una etiqueta clasificadora y despersonalizadora, pierde su sentido, y deshumaniza.

Por supuesto, indicadores y “**biomarcadores**”, bien diseñados y rigurosos con la estadística son excelentes herramientas de gestión y de decisión. Pero debemos estar seguros de no malinterpretarlos. Nuestra tarea va de historias personales y sociales, todas diferentes y todas con condicionantes, contexto y excepciones. ¿Dónde queda la persona, y su derecho a rechazar una injerencia en su trayectoria vital, en su cuerpo o en su espíritu? Seguramente la “**atención centrada en personas**” debería ir precisamente de eso: de conjugar eficiencia y atención personalizada, ciencia y respeto, obligaciones y comunicación. Asumido este marco, la IA sería seguramente una aliada para una atención justa, cálida y de calidad. Fuera de este marco será una excusa más para no dirigir la mirada.

En este sentido, el “**datafundamentalismo**” de este nuevo paradigma puede llevarnos a con-

fundir los perfiles que representan a las personas con las personas mismas, que pasarían de seres humanos a códigos o categorías. ¿Es esto lo que debería perseguir una auténtica medicina humanizada? Aquellos que atendemos y los que somos atendidos, debemos huir de la instrumentalización del otro en la relación asistencial. Todos tenemos nuestra parcela de responsabilidad. Si no cuidamos desaparecerá la confianza, que es la bisagra esencial de cualquier relación de ayuda, de cualquier relación social, al fin y al cabo. Sin confianza, la ayuda se convierte en servicio neutro, sin piel, sin cara. Al fin y al cabo, el cuidado de la salud, por complejo que haya devenido, supone una constelación de relaciones interpersonales.

En este contexto es en el que llama a la puerta la IA, una nueva y fascinante tecnología. Pero ya estábamos perdiendo la mirada en la atención sanitaria antes de la expansión de las herramientas digitales y de las decisiones asistidas por algoritmos.

3/

IA y conocimiento: un nuevo paradigma.

La IA reta al método científico tradicional, que buscaba contrastar hipótesis. La relación causa-efecto, esencial en ciencia desde el siglo XVI, no es tan relevante cuando utilizas IA. Ésta, al menos por el momento, ni busca causas, ni busca sentido, ni explica. Su territorio es el de las correlaciones. Por ello, muchas instituciones o empresas que priorizan el pragmatismo y la rentabilidad ya se han adherido hace tiempo sus habilidades predictivas (por ejemplo, bancos y aseguradoras). Lo importante es si la persona devolverá el préstamo o no, no las razones por las que ocurrirá esto.

Es fundamental ofrecer un trato cálido y respetuoso, centrado en la persona y no únicamente en su enfermedad, para preservar la empatía y la confianza en la relación asistencial

Y este nuevo paradigma ya se está imponiendo también en Medicina. Es la Medicina de las “P”: predictiva, participativa, preventiva, personalizada y de precisión, que ha encontrado un catalizador en la IA y en todas las técnicas de análisis de datos masivos. Aspiramos ya a detectar trastornos y/o riesgos para la salud sin necesidad de que los síntomas nos avise. Queremos preservar la salud más que atender a la enfermedad. Queremos “vacunarnos” contra toda enfermedad. Incluso el término “paciente” se tensiona en el nuevo modelo.

En esencia, la IA no es más que una rama de la informática que cataliza las potencialidades de la estadística, sobre todo de las probabilidades condicionadas bayesianas. Diríamos que es una tecnología de la información “especial”, que hace algo más que informar, porque es capaz de transformar la información en predicciones, avisos o recomendaciones.

No es tan nueva como pensamos. Los pensadores medievales, como por ejemplo Ramon Llull, ya asumieron la posibilidad de que ciertas funciones mentales podrían ser ejercidas por entidades diferentes al cerebro humano, y constan intentos ya decididos de construir máquinas pensantes desde principios del siglo XX y, sobre todo desde la II Guerra Mundial.

Sin embargo, la revolución digital ha posibilitado que los proyectos relacionados con esta área del conocimiento y de la técnica se hayan incrementado de manera más que exponencial en los últimos 30 años. ¿Por qué? Por las mejoras en la infraestructura computacional que han permitido el desarrollo de herramientas de gestión de datos masivos y el desarrollo de un “aprendizaje profundo” reinventado gracias, por ejemplo, a las mal llamadas “redes neuronales” artificiales (mejor “redes de microperceptrones”)

4/

El vértigo de la sustitución de la mente.

4/1

¿Qué es “artificial”?

Podríamos pensar que el adjetivo “artificial” no es afortunado. Al fin y al cabo, los sistemas de IA no son más que aparatos hechos con materiales naturales por individuos naturales. Quizá utilizamos “artificial” cuando algo aparece después de una decisión de búsqueda, de una invención. Existen animales que fabrican elementos que no existían (termiteros, diques de castores, etc.), pero quizá la diferencia está, como decíamos, en la intención. Por tanto, desde el punto de vista moral, no está mal conservar el apellido “artificial” para recordarnos que, en el fondo, existe una responsabilidad humana detrás del desarrollo de estas técnicas y, más allá, detrás de las preguntas que les hacemos y de los datos que facilitamos a sus algoritmos.

Buscar responsabilidad en la máquina, al menos de momento, no se puede plantear. Los algoritmos no son más que fórmulas, complejas, pero fórmulas. Cuando hablamos de “caja negra” en la IA no lo hacemos porque no conocemos la fórmula que estructuran su funcionamiento ni porque no conozcamos las ramas y bifurcaciones de sus algoritmos.

Hablamos de “caja negra” porque llega un momento en que no podemos controlar el flujo de los datos y resultados a través de dicha estructura lógico-matemática. Pero tampoco debemos olvidar, si nos preocupan las cajas negras, que en estos momentos hay casi 8 millones de cajas negras en la biosfera, un número equivalente a la cantidad de cerebros humanos funcionando.

4/2

**El eterno problema
de definir la inteligencia.**

Sobre el término inteligencia (etimológicamente, “**leer entre**”), también debemos reconocer que muchas herramientas de IA actualmente disponibles son claramente superiores a nosotros en esa tarea: “**distinguir patrones de correlación**”. Hablamos de inteligencia cuando distinguimos lo no aparente, cuando interpretamos la realidad y la percepción para resolver problemas. Las herramientas de IA detectan patrones y correlaciones entre fenómenos y datos no accesibles al cerebro humano. Por eso nos parecen una ayuda interesante.

Además, a menudo son capaces de detectar estas correlaciones “**en tiempo real**” tendiendo a cero el tiempo requerido para los antes tediosos análisis estadísticos. En este punto, no obstante, es obligado recordar de nuevo que la IA detecta cosas en las que las personas no pueden reparar, pero con unas “**gafas**” que diseñan, como mínimo en su origen, las personas.

Ahora bien, los sistemas IA no conocen el sentido de las correlaciones que detectan, ni les importa. Tampoco sienten vergüenza. En el fondo, todavía “**no saben lo que hacen**”. Así pues, la IA no pretende explicar el funcionamiento del cerebro, ni necesariamente imitarlo. En principio, no habría que esperar que se comporte como un humano. Aunque usemos la palabra inglesa “**learning**”, un algoritmo ni aprende ni es creativo. Estas son cualidades antropocéntricamente proyectadas. Podríamos definir un espectro dimensional, según la complejidad de datos manejados que iría desde la estadística descriptiva clásica, pasando por el aprendizaje automático y el aprendizaje profundo hasta los sistemas de IA autónomos sobre los que actualmente existe una frenética actividad de investigación. Ésta es una limitación relevante si hablamos de suplir el rol de las personas en una hipotética relación clínica con algoritmos.

El vértigo antropológico que genera la posibilidad de ser sustituidos también influye en nuestras valoraciones acerca de la introducción de la IA en la relación clínica. Decía Warner Slack (1997), pionero de la informática médica, que si un médico podía ser sustituido por IA es que se lo merecía.

Sentimos un vértigo similar en el Neolítico, cuando descubrimos que los animales hacían mucho mejor, tareas agrícolas o de transporte que nosotros y, más tarde, en el siglo XVIII, cuando fueron máquinas transformadoras de energía las que comenzaron a asumir tareas que antes hacíamos humanos o animales, con mayor rendimiento, precisión y resistencia.

La diferencia de base entre inteligencia humana y artificial es que la IA se basa en algoritmos matemáticos y programas informáticos, mientras que la humana se basa en procesos biológicos y fisicoquímicos. El problema es que ahora ciertas máquinas opositan a asumir funciones mentales. Y los humanos han hecho radicar el concepto de persona, con mayor o menor grado de consenso, en dichas funciones. Por tanto, la pregunta está servida: ¿La IA nos despersonalizará? ¿La IA amenaza nuestra especie?

5/

Oportunidades e IA.

Como decíamos, podemos asumir ya a día de hoy que, ciertas herramientas de IA son mejores que nosotros en la velocidad de análisis de datos masivos, así como en la detección automática de correlaciones, tendencias y patrones. Por tanto, son solventes para clasificar, revisar, reconocer o predecir. La diferencia con la inteligencia humana se nota especialmente cuando hay infinitas columnas de variables y pocas observaciones.

Y no podemos negar que estas cualidades ofrecen grandes oportunidades en diferentes campos de la atención a la salud

(Pujol Guanyabens, 2023):

- Puede permitir una atención clínica, médica o quirúrgica, más segura y rigurosa, que maneje resultados en tiempo real y que disponga de ayudas casi instantáneas al razonamiento clínico, personalizando tratamiento y estableciendo pronósticos con mayor precisión. Los campos más tecnificados de la atención (cirugía, anestesia, cuidados intensivos, prótesis, etc.) seguramente ya han sido los primeros en recibir estas aportaciones. Por tanto, la IA da soporte a decisiones y a procedimientos sanitarios, automatizando tareas para incrementar su seguridad, exactitud y reproducibilidad. Permite una gestión más eficiente de recursos porque puede evaluar resultados favorables y desfavorables casi en tiempo real. Sin embargo, todavía no admitimos decisiones opacas de un algoritmo sobre una persona, sin supervisión de otra persona.

De todas maneras, para ser efectiva, esta “supervisión” debería iniciarse ya en el diseño de la herramienta. Una vez diseñadas, cuando abordamos decisiones difíciles con amplios márgenes de incertidumbre, resulta tentador “obedecer” la sugerencia de un elemento inanimado que tendemos a asumir como más exacto, neutral y exento de conflictos de interés, premisas todas cuando menos debatibles al tratarse de productos humanos. Comenzamos a disponer ya de publicaciones que muestran la capacidad de la IA generativa para realizar diagnóstico. Por ejemplo, en 2022, en 30 pacientes tratados en Urgencias en un hospital, se compararon las listas de diagnóstico diferencial hechas por médicos con las que sugería **chatGPT**. Observaron una gran superposición (alrededor del 60%) entre ambas. Los médicos humanos incluyeron el diagnóstico correcto entre los 5 primeros en el 87% de los casos mientras que la versión gratuita de ChatGPT 3.5 lo hizo en el 97% de los casos (Berg et al., 2023). Seguramente el camino más adecuado

en este campo será generar una cultura de decisiones híbridas **persona-IA** que resulten eficientes y seguras a un tiempo, considerando valores.

- Aunque su introducción es más pujante en la asistencia directa, quizá su impacto puede ser aún mayor en Salud Pública, dada su capacidad para el manejo de datos masivos y para detectar patrones. Predicciones, tendencias y/o monitorización de decisiones son instrumentos clásicos de los salubristas. Y la predicción sobre colectivos suele ser más potente que cuando se intenta anticipar lo que sucederá con un individuo concreto. Hemos visto estas aportaciones, con mayor o menor rigor, en tiempos de pandemia.

- Derivado de todo ello, cuesta poco imaginar oportunidades en el campo del diseño de políticas sanitarias, que no pueden perder nunca el foco colectivo y que necesitan también predicción de patrones y tendencias, así como cierta “neutralidad” en los procedimientos para preservar la justicia en el mayor grado posible. Ofrece a los responsables información sobre escenarios de futuro que faciliten decisiones sobre prioridades, reorganización y planificación en el sistema de salud, y sobre la gestión de las instituciones sociales y sanitarias. Además, como se trabaja con datos agregados y con objetivos ligados al bien común, los riesgos y restricciones relacionados con privacidad individual representarían un obstáculo menor.

- Por supuesto, también la investigación se puede ver beneficiada, especialmente en aquellos estudios que llamamos “observacionales”, porque la IA no tiene por qué observar muestras pequeñas y luego extrapolar, sino que puede buscar patrones, tendencias y correlaciones en la inmensidad de datos disponibles en el mundo real. Además, ya hemos dicho que puede detectar todo eso con mayor sensibilidad (seguramente no, por el momento, con mayor especificidad) que los investigadores de carne y hueso. También la gestión de datos en ensayos clínicos, el diseño de fármacos y el estudio de

las características y comportamiento de las proteínas son otros campos que ya están recibiendo beneficios de las técnicas de IA.

- También nos permite evaluar datos del mundo real sobre resultados positivos y negativos de nuestras intervenciones; terapias, programas asistenciales, cribados, nuevas tecnologías. Por tanto, puede ser una herramienta muy valiosa para preservar la calidad asistencial y la sostenibilidad del sistema sanitario y/o social.

- Por supuesto, cuesta poco imaginar el impacto potencial de la IA, sobre todo la generativa, en la educación médica, facilitando mecanismos de entrenamiento de los profesionales en formación que no supongan riesgos para personas: simulación, generación de “casos”, herramientas de evaluación, etc.

No obstante, el sistema nervioso central, es todavía, hoy por hoy, “mejor” que la IA, detectando valores y prioridades humanas, aplicando la solidaridad y la empatía. Es más flexible, tiene capacidad de improvisar. Siente vergüenza, como mecanismo de autocontrol. Además, parece que “se da cuenta” de lo que hace, aunque a menudo no lo parezca.

Decía Pablo Picasso que los ordenadores solo dan respuestas, no se hacen preguntas, aunque los expertos discuten en la actualidad si la IA no podrá llegar a hacer preguntas originales en el futuro. A parte de la curiosidad y la implicación, el ser humano tiene también otras armas valiosas, pero de “doble filo” si aspiramos a preservar la equidad y la justicia: la subjetividad, la empatía y la compasión.

6/

Riesgos e IA.

6/1

¿Es el “aprendizaje profundo” una amenaza?

De forma muy simplista, podríamos decir que algunos sistemas algorítmicos de IA son capaces de hacer evolucionar sus respuestas finales en función de la experiencia acumulada por los datos que van recibiendo. Podríamos decir pues que las respuestas de estos sistemas se van construyendo “desde dentro”, y esto puede generarnos dos impresiones:

- Autonomía: puede parecer que “decide la máquina” o, incluso, que podría llegar a ser, en algún momento del desarrollo tecnológico, autoconsciente como nosotros y, más allá, engañar o mentir como nosotros.

- Misterio u opacidad, como se quiera: Es el fenómeno de la “caja negra” antes comentado. Conocemos la fórmula, pero no controlamos del todo el flujo de datos dentro de la estructura algorítmica. Los humanos no acabamos de comprender, en ocasiones, cómo el sistema ha elaborado una predicción o un consejo. Es una fórmula tan complicada que pocos le pueden dar un significado, e incluso sus diseñadores en ocasiones no entienden el flujo de datos. Es una limitación humana, no necesariamente un “defecto” del sistema. No olvidemos nunca que la estructura lógica de fórmulas de un sistema de IA es transparente y la ha escrito un humano, o un equipo de humanos.

La literatura de ficción ha planteado repetidamente este problema, centrado en la pregunta: ¿Podrá la creación controlar o superar al creador? El ordenador Hal 900 de “2001, una odi-

Aunque la IA y los avances tecnológicos aportan beneficios, no deben reemplazar la interacción humana, ya que la relación interpersonal sigue siendo clave en el cuidado de la salud

sea del espacio” sería un buen ejemplo. Sin embargo, no debemos olvidar que, hasta la fecha, el “aprendizaje profundo” no aprende nada. Aprender es una metáfora humana para entender su funcionamiento.

En el fondo, estamos hablando de automatizar procesos, muchos de ellos responsabilidad de nuestra mente hasta hace poco tiempo. De hecho, todo esto ocurre ya desde hace mucho tiempo, mucho antes de que la revolución informática y de las tecnologías de la comunicación de los últimos 30 años.

Las calculadoras científicas ya hacían operaciones matemáticas con mucha mayor velocidad, exactitud y rapidez que la mente humana en los años 70. Como ahora esta “suplencia” de funciones mentales afecta a muchas otras funciones (lenguaje, lectura, comprensión de idiomas, predicciones, pronósticos, apoyo a las decisiones, etc.), el reto al propio concepto de persona se hace más patente, y genera mayor inquietud.

Es importante no olvidar que, de momento, la IA, incluso la generativa, sólo hace, no explica. Sería como la “calculadora” a la aritmética. Por eso la IA generativa siempre aborda las preguntas, por absurdas que sean. Completa agujeros, más que buscar la “verdad”. Su atractivo radica en que se puede usar sin conocimientos informáticos (del mismo modo que una calculadora sin conocimientos matemáticos o un vehículo sin conocimientos de mecánica).

6/2

El problema de los sesgos.

Los sesgos no son un problema inherente a los algoritmos. Los sesgos son nuestro problema como humanos, fruto de nuestra subjetividad, nuestra educación y nuestros instintos. El problema es que algoritmos y procesamiento de datos masivos pueden generar una ilusión de veracidad global generada a partir de datos no adecuados o no representativos, o de preguntas

con prejuicios y/o discriminación inherente. De hecho, los Sapiens tenemos muchos defectos cuando aplicamos y repartimos cuidados. A menudo nos centramos, sin darnos apenas cuenta, en solucionar nuestro propio conflicto interno empático, es decir, nuestro sufrimiento emocional al contemplar el sufrimiento del otro. Actuamos a menudo para mitigar dicho conflicto emocional interno individual, olvidando en mayor o menor grado el verdadero foco del cuidado: el sufrimiento del otro visto desde su propio prisma.

El problema de la representatividad de los datos que manejan los algoritmos se pone de manifiesto cuando reparamos en la presencia en internet de los diferentes territorios del globo. El mundo que alimenta a los algoritmos es un mundo deforme, que inflama a los territorios donde el desarrollo económico y tecnológico ha sido mayor.

Por tanto, a día de hoy, los datos que nos representan en el océano de datos masivos, distan mucho de ofrecer una imagen fidedigna de todos los humanos. Sin embargo, con ellos vamos construyendo perfiles de personas con la información predictiva que generan los sistemas de IA, y con ellos construimos una supuesta realidad social. Pero recordemos, los perfiles intentan representar, con muchas limitaciones, a la persona perfilada, no son la persona.

A pesar de las estrategias y marcos regulatorios puestos en marcha para evitar estos sesgos, estas herramientas pueden “blanquear”, cuando no amplificar, sesgos sociales, sobre todo si no se tienen en cuenta en su diseño todos los determinantes sociales de la salud.

Para introducir esta perspectiva es precisa la colaboración interdisciplinar en investigación biomédica, con la participación de las ciencias sociales y de las humanidades (antropología, sociología, ética, etc.). Dicha colaboración debería extender a entidades independientes capaces de revisar estos sistemas desde múltiples perspectivas (de Manuel et al., 2023).

6/3

Grados de riesgo en el uso de IA en salud.

Conscientes de los riesgos a nivel global de la implantación y desarrollo acelerado de la IA, la Unión Europea considera, en general, los sistemas de IA dedicados a tender a la salud humana como de alto riesgo, distinguiendo incluso diferentes escalones, partiendo de aquellas aplicaciones que sólo informan de forma más o menos organizada del manejo clínico habitual, pasando por aquellas que asesoran el manejo clínico, hasta llegar a aquellas que directamente tratan, cuidan o diagnostican.

Pero las principales incógnitas en relación al desarrollo de sistemas de IA con grados crecientes de autonomía son: ¿Podrán llegar a ser autoconscientes o tener voluntad propia? ¿Podrán arrepentirse o cambiar de opinión? ¿Podrán engañar intencionadamente y burlar sistemas de seguridad? Un artículo en Science de este mismo año no descarta en el futuro algunas de estas inquietudes y reclama una mayor inversión en seguridad de los sistemas de IA y una mayor concienciación y concreción en las medidas de gobernanza imprescindibles

(Bengio et al., 2024).

y muy fragmentadas. Además, trabajan sin disponer de datos suficientes en tiempo real para llevar a cabo decisiones de elevada trascendencia. Por ello, la mayoría de las decisiones clínicas acaban siendo extremadamente subjetivas y contingentes, poco contrastadas con datos y pruebas científicas. No hacemos una medicina “profunda” (Topol, 2019).

La IA, bien entendida como instrumento, podría ayudarnos a mejorar esto, porque nos permite analizar rápidamente muchas variables al mismo tiempo. Ha demostrado ya sus prestaciones para monitorizar de forma continua resultados o alarmas, para identificar objetivos, para organizar y poner a disposición el conocimiento disponible, para comparar intervenciones, para identificar áreas grises en aquello asumido como establecido, etc. Todas ellas son prestaciones extremadamente útiles para mejorar y hacer más sostenible una medicina cálida y de calidad.

Sin embargo, no nos puede ofrecer por el momento evaluaciones cualitativas y ponderadas, que consideren el contexto y los valores. Esto justifica la imprescindible hibridación de decisiones automatizadas y humanas en la atención a la salud. En el fondo, sería mejor hablar, por el momento, de decisiones humanas asistidas por IA.

7/

¿Nos puede ayudar la IA a mejorar y humanizar la medicina actual?

Debemos asumir que, en muchos casos, la medicina actual, y especialmente en los llamados países “desarrollados” es superficial, cara e ineficiente, a cargo de profesionales muy preparados pero que trabajan en organizaciones rígidas

8/

Retos éticos.

8/1

Privacidad, responsabilidad, transparencia.

Las herramientas de IA para la toma de decisiones sanitarias no dejan de ser o formar parte de dispositivos sanitarios. Conscientes de los

riesgos de intervenir sobre las vidas de las personas, tanto los fármacos como los dispositivos sanitarios están sometidos ya a una regulación estricta para garantizar que su introducción en la práctica aporta valor de forma razonablemente segura.

Por supuesto, dada la elevadísima necesidad de datos de salud generada por estos sistemas, proteger datos personales se convierte en una prioridad. Además, los datos de salud recogidos pueden tener otros fines, no médicos, para las compañías propietarias de la tecnología. Un estudio aparentemente inocente o incluso altruista en salud podría tener, en el fondo, otros fines alejados del sector salud. El estudio sería solo, en cierto modo, minería de datos.

Se genera además una responsabilidad sobre la calidad, veracidad y representatividad de los datos introducidos en las dinámicas algorítmicas. De otro modo, haremos predicciones o recomendaciones sesgadas, peligrosas y/o injustas.

En este marco, la transparencia y la rendición de cuentas son aún más clave, si cabe. Se deben establecer mecanismos transparentes de control y auditoría tanto sobre su funcionamiento como sobre sus resultados.

Hablamos pues de fiabilidad, trazabilidad y explicabilidad. Éste último se ha convertido en uno de los requisitos éticos clave en el desarrollo de la IA, aunque debemos tener en cuenta que cuanto más explicable es el sistema, menos predictivo resulta, y que la explicabilidad es progresivamente más difícil cuando el programa “**aprende**”, siendo éstos precisamente los sistemas que percibimos como un avance evidente en predicción. Además, los requisitos de explicabilidad pueden chocar también con el secreto empresarial y la libre competencia.

En cualquier caso, las personas que puedan recibir potencialmente los efectos, positivos o negativos, de la IA deberían tener acceso a una alfabetización digital suficiente, para así poder reducir su vulnerabilidad ante la IA.

No se trata de que el ciudadano medio pueda conocer su mecanismo íntimo informático, estadístico, lógico o algebraico. No se trata tanto de dominar el “¿cómo?” sino de poder valorar bien el “¿por qué?” y el “¿para qué?” de cada sistema, conociendo también sus limitaciones y posibles desviaciones. Por ejemplo, resulta esencial que los ciudadanos sepan que, hoy por hoy, los sistemas de IA sólo detectan correlaciones, sin poder definir relación causa-efecto entre fenómenos y datos. Además, deben conocer que sus predicciones y recomendaciones no son fijas, sino dinámicas, y en continua evolución en función de la entrada de datos y observaciones en el sistema.

8/2

Pronósticos y predicciones ajustadas a valores.

Decíamos antes que las soluciones de IA aplicadas al cuidado de la salud deberían ser consideradas, cuando menos, como dispositivos sanitarios. Pero seguramente la IA en salud es más que un dispositivo sanitario. No es un aparato que se diseña y se construye, para llegar a un “**diseño definitivo**”. Muchas soluciones algorítmicas se transforman con el contexto, a medida que le van llegando datos y acumula “**experiencia**”. Son soluciones dinámicas. Esto es importante, porque nuestra medicina clásica está acostumbrada a horquillas de normalidad, relativamente rígidas. Sin embargo, una solución de IA puede experimentar grandes cambios en sus recomendaciones con pequeños cambios en los nuevos datos que recibe. Los dinteles de seguridad o de predicción, por tanto, son líquidos. Esto debe tenerse muy en cuenta cuando dichos resultados predictivos implican decisiones trascendentes. Conceptos de la estadística como valor predictivo, tanto positivo como negativo, cobran renovada trascendencia en el paradigma de la Medicina de las “**P**”. Cuando una herramienta tiene gran capacidad para detectar patrones y correlaciones, es necesario diferenciar el ruido de las correlaciones útiles en un océano de “**alarmas**” posibles.

Además, las correlaciones útiles en Medicina para humanos deberían estar trufadas de valores. Hasta la fecha, quizá encontremos aquí una de las limitaciones fundamentales para hacer una buena “**medicina algorítmica**”. Los dinteles que marcan actuar o no actuar, seguir o no seguir, reanimar o no reanimar, etc. generan una responsabilidad ética elevada. Y estos dinteles deben estar basado en valores compartidos. Por tanto, parece difícil que puedan establecerse desde una visión miope monodisciplinar. Volvemos a la gran pregunta: ¿Qué queremos de la IA en salud? Y más difícil aún ¿Qué resulta valioso para cada persona de todo lo que ofrece la IA en salud?

El paradigma “**promete**” una monitorización más o menos continuada del estado de salud para detectar precozmente desviaciones de “**la norma**” que permitan diagnósticos más precoces, Apoyada en su sensibilidad en la detección de patrones, parece claro que la IA hará más diagnósticos y más tempranos. Pero por el momento no sabe hacer diagnósticos oportunos (aquellos realizados en el momento justo en el que aportarán valor a la persona). También tiene, por el momento, limitaciones precisando la necesidad de intervención. No pueden, por tanto, decidir la oportunidad de las decisiones en el marco de los valores y el contexto situacional de las personas atendidas.

Debemos recordar siempre que, por el momento, la IA propone diagnósticos, cuidados o tratamientos. Propone también factores de riesgo de enfermar, factores a menudo no advertibles por la mente humana. Pero aún no diagnostica. ¿Y si nos situamos en la hipótesis de que consiga diagnosticar en el futuro, lo hará de manera personalizada, respetuosa y justa?

También cabe cuestionarse, de entrada, si la monitorización continua de la salud es salud. La misión de la Medicina es que suframos los menos posible, si así lo queremos. ¿Un sistema de alarma 7/24 en personas sanas elimina sufrimiento, o lo genera? Asumir los beneficios implicaría una revolución cultural, y casi antro-

pológica, del ser humano. ¿Sería una vida llena de falsas alarmas, o de alarmas inconcretas, una buena vida? Predecir implica quizá diagnóstico, pero, sobre todo, pronóstico, y oportunidades en función de valores. La accesibilidad y oportunidad de la información que nos facilite la IA representan los dos principales retos éticos para su introducción en la asistencia, entendida ésta, recordemos, como relación de ayuda interpersonal, y responsable con la biosfera.

8/3

Una ética de la responsabilidad en IA.

Todo avance tecnológico debería estar al servicio de las personas y de la biosfera. El servicio, para las personas, sería hacerles accesible e intuitiva la matemática compleja, la estadística analítica y el análisis de sistemas complejos.

¿Debe de ser siempre supervisada por humanos la IA? ¿Es viable siempre dicha supervisión, incluso en sistemas que usen aprendizaje profundo y produzcan resultados en tiempo real? Seguramente la supervisión humana más relevante se referiría al diseño, a los objetivos, al entrenamiento y a las preguntas a hacerle al sistema.

Y está también el problema de la responsabilidad ante los hipotéticos errores, o los efectos adversos de la aplicación de sistemas de IA cada vez más autónomos. ¿Quién se hará responsable de los daños? ¿Quién será responsable de respetar los derechos de los usuarios, incluida su privacidad? ¿Cómo evitaremos que las correlaciones detectadas deriven en discriminación y/o estigma?

La inteligencia artificial ha demostrado ser superior a los humanos en tareas de análisis masivo de datos y detección de patrones, pero carece de cualidades esenciales como la empatía y la comprensión del contexto

9/

Conclusiones.

Deberíamos evitar discursos del miedo, pero sin dejar de evaluar la IA como cualquier otra tecnología innovadora de riesgo, aunque llena de oportunidades que también debemos distinguir con transparencia y accesibilidad.

Debemos distinguir aquellas tareas en las que la IA nos supera y permitir que colabore con nosotros en ellas. Seguramente el concepto de decisiones híbridas humano-IA será cada vez más habitual (decisiones humanas asistidas por IA).

Por supuesto, en el desarrollo de soluciones de IA, existe el riesgo de la progresiva concentración de empresas tecnológicas que podría llevar a una especie de monopolio del poder, del conocimiento, incluso de la moral. Preservar la independencia y la equidad en la evaluación de todo proyecto de IA en salud será una garantía a blindar democráticamente.

Otro reto para la biosfera es que, a pesar de su apariencia de sistemas “limpios” sin emisiones, su huella ecológica, relacionada con su elevadísimo consumo de energía y con los materiales empleados, es muy intensa, y mayor a medida que se desarrollan más y más potentes sistemas.

Seguramente la IA, en vez de substituir directamente profesionales sanitarios, debería convertirse en una ayuda para recuperar la humanización perdida. Por supuesto, para conseguir esto debemos aplicar los principios de prudencia y proporcionalidad. Como decíamos la pre-evaluación será clave (justificación, objetivos, características de los datos para nutrir los algoritmos, preguntas que les haremos, etc.). No hace muchos años sentíamos también inquietud y preocupación por la introducción de los ordenadores en las consultas.

Las nuevas herramientas de gestión de datos y de procesamiento del lenguaje natural pueden generar tiempo para la interacción personal, para la comunicación y para la empatía. Por supuesto, todo ello ocurrirá si el tiempo que se genere no sea devorado por la burocracia o el mero economicismo. La salud de los humanos merece beneficiarse de la sostenibilidad y de la mejora de calidad que pueden aportar las soluciones digitales, pero ello exige introducir la reflexión ética y la participación pública en el diseño, uso e interpretación de estas nuevas herramientas, que deben ser seguras y libres de intereses espurios. Deberían estar “vacunadas” contra el totalitarismo, la exploración y la manipulación. Algunas propuestas pragmáticas serían: **(Owens, 2023):**

- Mantenimiento y reparación continuada, con auditorias ad-hoc en tiempo real.
- Detección y mitigación de sesgos.
- Considerar las limitaciones de la IA a la hora de interpretar resultados:
 - Mínimas desviaciones de nuevos datos entrantes pueden alterar severamente las recomendaciones salientes.
 - La evolución tecnológica, social, demográfica, puedes deteriorar el valor añadido que aportan estos modelos.
 - Por el momento no se debería inferir relación causal de las correlaciones que detectan.
- Priorizar criterios de justicia, equidad y no discriminación en su desarrollo.
- Garantizar la accesibilidad a soluciones de IA que aporten valor a los humanos.
- Cuidar que tanto los profesionales, como las aplicaciones de IA, como las personas atendidas gocen de la mayor independencia posible. Debemos evitar convertir la IA en un bio-poder o en una estafa moral.

Por tanto, la IA puede ayudar mucho en el cuidado de la salud de las personas, pero solo será así si está bien pensada, diseñada, alimentada de datos, gestionada, revisada y mantenida. La transparencia y la rendición de cuentas serán los principios éticos clave en este nuevo escenario. Es todo un nuevo reto para la gobernanza democrática.

Sin caer en alarmismos, debemos partir de la base que la IA hace actualmente muchas cosas mejor que los humanos, pero para ser útil debe estar diseñada acorde a valores compartidos, bien nutrida de datos representativos, inclusivos y libres de prejuicios. Por último y, sobre todo, debe ser accesible y transparente para no ser injusta.

Aunque muchas actividades médicas se puedan desarrollar sin rostro, debemos tener muy presente que las cosas mejoran con rostro. Como no pocos expertos proponen ya, no debemos conformarnos con aquello que llamamos medicina personalizada (más bien estratificada en muchos casos, por el momento.) Necesitamos que la medicina sea personalizada, pero sobre todo personal (Picard, 2017; Esquerda & Pifarre, 2024), preservando el rostro y la mirada en la relación clínica.



Bibliografía

Bengio Y., Hinton, G., Yao, A., Song, D., Abbeel, P., Darrell, T., Harari, YN., Zhang, YQ., Xue, L., Shalev-Shwartz, S., Hadfield, G., Clune, J., Maharaj, T., Hutter, F., Baydin, AG., McIlraith, S., Gao, Q., Acharya, A., Krueger, D., Dragan, A., Torr, P., Russell, S., Kahneman, D., Brauner, J. y Mindermann, S. (2024) *Managing extreme AI risks amid rapid progress*. Science, 384(6698), 842-845. doi: 10.1126/science.adn0117. Epub 2024 May 20. PMID: 38768279.

Berg, HT., van Bakel, B., van de Wouw L., Jie, KE., Schipper, A., Jansen, H., O'Connor, RD., van Ginneken, B. y Kurstjens, S. (2023). *ChatGPT and Generating a Differential Diagnosis Early in an Emergency Department Presentation*. Ann Emerg Med. 2024 Jan; 83(1): 83-86. doi: 10.1016/j.annemergmed.2023.08.003. Epub 2023 Sep 9. PMID: 37690022.

de Manuel, A., Delgado, J., Parra Jounou, I., Ausín, T., Casacuberta, D., Cruz, M., Guersenzvaig, A., Moyano, C., Rodríguez-Arias, D., Rueda, J., y Puyol, A. (2023) *Ethical assessments and mitigation strategies for biases in AI-systems used during the COVID-19 pandemic*. Big Data & Society, 10(1). <https://doi.org/10.1177/20539517231179199>.

Esquerda, M. y Pifarré-Esquerda, F. (2024). *Inteligencia artificial en medicina: aspectos éticos, deontológicos y el impacto en la relación médico-paciente*. Med Clin (Barc). May 7:S0025-7753(24)00214-8. English, Spanish. doi: 10.1016/j.medcli.2024.03.003. Epub ahead of print. PMID: 38719685.

Owens, K. (2023) *For Ethical Use of AI in Medicine, Don't Overlook Maintenance and Repair*. Posted in Artificial Intelligence, Hastings Bioethics Forum. <https://www.thehastings-center.org/for-ethical-use-of-ai-in-medicine-dont-overlook-maintenance-and-repair/>

Picard, A. (2017), *Matters of life and death: Public health issues in Canada*. Vancouver: Douglas & McIntyre.

Pujol, R. y Guanyabens, J. (2023). *La medicina en tiempos de inteligencia artificial*. Med Clin (Barc.), 161(12), 530-32. <https://doi.org/10.1016/j.medcli.2023.09.007>

Rizzolatti L. y Cattaneo L. (2009). *The Mirror Neuron System*. Arch Neurol, 66 (5), 557-560.

Slack WV. (1997). *Cybermedicine: How Computing Empowers Doctors and Patients for Better Care*. San Francisco: Jossey-Bass Edition.

Topol, EJ. (2019) *Deep medicine: how artificial intelligence can make healthcare human again*. First edition. - New York: Basic Books.



04/

Dignidad humana y bioética de la Inteligencia Artificial.

Rafael Amo Usanos,

Cátedra de Bioética.

Universidad Pontificia Comillas. Madrid.

El autor plantea la creciente influencia de la Inteligencia Artificial (IA) en diversos ámbitos, particularmente en la salud, y la urgencia de abordar su impacto desde una perspectiva ética y bioética antes de que se vuelva imprescindible. También analiza cómo la dignidad humana, un principio central en la bioética católica y en la Declaración Universal de los Derechos Humanos, debe ser el eje fundamental en la regulación de la IA.

El uso de IA en salud plantea desafíos éticos, como los sesgos en los datos y el riesgo de injusticia. Se subraya la necesidad de desarrollar una "algorética", o ética de los algoritmos, que promueva transparencia, responsabilidad y respeto a la dignidad humana. Además, destaca el impacto de la IA en la relación médico-paciente, el ejercicio de la profesión médica y la toma de decisiones terapéuticas, donde se advierte que la IA no debe sustituir el juicio humano.

Finalmente, se enfatiza la necesidad de formar a médicos y profesionales de la salud en el uso ético de la IA, integrando principios bioéticos en su formación para asegurar que la tecnología sirva al bien común sin comprometer la dignidad humana.

Palabras clave: Inteligencia Artificial, Humanización, Asistencia.

This article takes a look at the growing influence of artificial intelligence (AI) in various spheres, particularly in health, and discusses the urgency of considering its impact from the ethical and bioethical standpoints before it becomes ineluctable. The author also analyses why human dignity, which is a central principle in Catholic bioethics and in the Universal Declaration of Human Rights, should form the cornerstone of AI regulation. The use of artificial intelligence in health poses ethical challenges, such as data biases and the risk of inequity. In connection with this, the need is highlighted to develop an "algorithethics", or ethics of algorithms, which will favour transparency, responsibility, and respect for human dignity. Likewise, this article foregrounds AI's impact on the physician-patient relationship, on the practice of the medical profession, and on the making of treatment decisions, pointing out in this respect that AI should not substitute human judgement.

Lastly, the need is emphasised to train physicians and other health professionals in the use of artificial intelligence, integrating bioethical principles in their training in order to ensure that technology will serve the common good without compromising human dignity.

Key words: Ethics, Human dignity, Artificial intelligence, Responsibility.

1/

El principio dignidad y la Bioética católica.

Las aplicaciones de **Inteligencia Artificial (IA)** están inundando nuestra vida diaria y llegan con rapidez al ámbito de la salud. Antes de que se transformen en imprescindibles, la sociedad debería dejar encauzado el tema de la ética de la IA; de lo contrario será demasiado tarde.

La literatura al respecto de la ética y bioética de la IA se está multiplicando. También en los escasos textos que quieren regularla, y en los trabajos previos de las distintas instituciones que precedieron a la reciente legislación, aparecen trazas de una posible ética de la IA.

La propia Academia Pontificia para la Vida tiene un “**programa**” de trabajo sobre la ética de la IA y publicó el 28 de febrero del 2020 -fecha de infausta memoria por su relación con el COVID- un documento titulado **Roma Call for IA Ethics**.

En la incipiente ética de la IA se encuentran diversas fórmulas de articulación: algunos proponen unos principios propios, que más bien son los principios de una ética del algoritmo (que no es exactamente lo mismo que una ética de la IA); otros proponen aplicar modelos de ética existentes¹.

Pero quedan otros caminos, el pasado 14 de junio el Santo Padre Francisco, en su discurso ante los participantes de la sesión del G7, propuso “**poner de nuevo al centro la dignidad de la persona en vistas a una propuesta de ética compartida**”².

La dignidad humana tiene una historia paradójica en el siglo XX. Ha sido el siglo que, casi con toda seguridad, ha visto las violaciones más sangrantes de la dignidad humana, pero por otra ha asistido a la más solemne de las declaraciones en su defensa, la **Declaración Universal de los Derechos Humanos**. En muy pocos años se pasó del humo de las chimeneas de Auschwitz, a la más solemne de las proclamaciones de la dignidad humana.

El magisterio de la Iglesia lleva unas décadas, desde el **Concilio Vaticano II**, fraguando lo que podríamos llamar el principio dignidad³. Lo hizo en la primera parte de la constitución pastoral *Gaudium et spes* en la que aborda los valores emergentes de la sociedad de su momento, signos de la presencia del Espíritu en el tiempo. El más destacado, y que desarrolla en primer lugar, es la dignidad humana.

Desde ese momento, la dignidad de la persona se convierte en eje vertebrador de la Doctrina social de la Iglesia y, en concreto de su propuesta bioética. Así dice **Evangelium Vitae**, carta magna fundacional de la bioética católica contemporánea:

“El Evangelio del amor de Dios al hombre, el Evangelio de la dignidad de la persona y el Evangelio de la vida son un único e indivisible Evangelio”⁴.

Y lo reafirma, **Domum Vitae, Dignitas personae** y **Dignitas infinita**.

1. Por ejemplo, Luis Peñas García, *La ética de la Inteligencia artificial generativa*

2. https://www.vatican.va/content/francesco/es/speeches/2024/june/documents/20240614-g7-intelligenza-artificiale.html#_ftn5

3. En realidad, es Juan XXIII en *Mater et Magistra* 453 y 459 quien pone la primera piedra de la construcción de la lectura católica de la dignidad de la persona.

4. Juan Pablo II, *Evangelium Vitae*, 2.

Este último y más reciente documento hace una breve exposición del concepto dignidad y su fundamento. Dignidad es un término polisémico, puede hablarse de la dignidad ontológica (la que corresponde a la persona por el hecho mismo de existir); moral (el ejercicio de la libertad por parte de la criatura humana); social (las condiciones en las que vive una persona); existencial (la que procede de la capacidad de vivir con paz, alegría y esperanza).

Dignidad es un término que conviene a la persona, sustancia individual de naturaleza racional. Asumiendo el documento la clásica definición de **Boecio**, da razón del fundamento filosófico de la dignidad.

La persona posee dignidad ontológica porque es una sustancia individual, lo que significa que

“Habiendo recibido la existencia de Dios, subsiste, es decir, ejerce la existencia autónomamente”⁵.

Ahora bien, esta subsistencia se da en un ser de naturaleza -comprendida como principio del obrar- racional, que engloba la dimensión cognitiva, la volitiva, el amor, la capacidad de desear y elegir.

Teológicamente hablando el fundamento de la dignidad de la persona radica en el hecho de que toda ella -cuerpo y espíritu- ha sido creado a imagen de Dios y ha sido llamado a la unión con Dios. Esta dignidad ha sido puesta de manifiesto en Cristo, quien al unirse en cierto modo a cada ser humano

“Confirmó que todo ser humano posee una dignidad inestimable, por el mero hecho de pertenecer a la misma comunidad, y que esta dignidad no puede perderse jamás”⁶.

2/

El principio dignidad como lugar de encuentro.

Ahora bien, como afirma **Schockenhoff**, la dignidad humana es compatible con diferentes fundamentaciones antropológicas. Esto es lo que la dota de esa capacidad de punto de encuentro para una ética compartida.

Según afirma **Melendo Granados**, la dignidad

“Es una sublime modalidad de lo bueno, de lo valioso, de lo positivo, la bondad de aquello que está dotado de una categoría superior”⁷.

Es decir, que es un sustantivo adjetivado que actúa como un superlativo, lo que implica el grado más alto de bondad, la categoría máxima que se puede alcanzar. Dignidad es el superlativo del valor, y corresponde por tanto a un tipo especial de seres, los más singulares. Aquellos que estén elevados de la materia y la especie, y posean interioridad. Esto les dará una capacidad exigitiva: quien sea calificado como digno podrá exigir ser tratado como un fin y no como un medio.

Estos seres son las personas, elevadas de la materia por el conocimiento intelectual, la capacidad simbólica y la libertad. Elevadas sobre la especie porque solo la persona tiene esa forma de vida, bios, que es la capacidad de gestionar su vitalidad, zoé.

Construir su futuro haciéndose responsable de él. Así se explica que la vida de los individuos de la especie humana no sea igual, cada persona humana es distinta de las demás porque ha escrito su vida (bio-grafía) de modo diferente.

5. Dicasterio para la Doctrina de la Fe, Dignitas infinita, 9

6. Dicasterio para la Doctrina de la Fe, Dignitas infinita, 19

7. Melendo Granados, Tomás (1999). Sobre el hombre y su dignidad, 6. p. 70

LH n.339

Otra de las características que hace de las personas aquellos seres singulares a los que corresponde el calificativo de digno es la interioridad en su doble significado, uno psicológico y otro ontológico. El primero no es sino la manifestación fenomenológica del segundo.

La introspección, significado psicológico, es la capacidad de decir yo, de ser consciente de sí mismo. Aquello, que según narra **Viktor Frankl**, fue lo único que los nazis no pudieron arrebatárle en el campo de concentración.

En sentido ontológico, es -para Melendo Granados- la propiedad privada del acto de ser. Lo que **Santo Tomás** llama subsistencia, lo que **Zubiri** llama personidad. Es decir, que el ser en sentido metafísico, aunque dependa en su origen de otro ser, puede existir por sí mismo.

La persona, elevada sobre la materia y la especie y dotada de interioridad -psicológica y ontológica- es autónoma. Esto no es solo que tiene la capacidad de realizar sus elecciones con criterios propios, sobre la base de valores y creencias personales que es lo que afirman **Beauchamp** y **Childress**, (que más bien sería una consecuencia de la elevación sobre la especie); sino que tiene capacidad de exigir ser tratada como un fin y no como un medio: ella misma (auto) es ley (nomos) y debe ser respetada.

3/

El impacto de la IA en la Bioética de la dignidad.

La IA es una revolución tecnológica de tal calibre que impacta por todos los lados y a una enorme velocidad en la Bioética. Sería imposible enumerar las aplicaciones prácticas que cada día se diseñan. En el informe que la Asociación

Española para la Digitalización publicó en febrero de 2022⁸ las aplicaciones se clasificaban en cinco categorías: Tratamientos y cuidados personalizados; Asistencia al diagnóstico médico; Medicina preventiva y autocuidado; Investigación biomédica y desarrollo de nuevos fármacos; y, Gestión optimizada de los Hospitales y mejora de la calidad asistencial.

Proceder analizando cada aplicación, o cada grupo de las categorías de uso, es un trabajo de tal envergadura que lo haría casi imposible. Sin embargo, una lectura transversal de la literatura que va apareciendo permite extraer algunos temas.

4/

La ética de la tecnología, la IA y la Bioética.

La reflexión filosófica y ética sobre la tecnología surge en el primer tercio del siglo XX y su profundidad corre paralela a las nuevas capacidades de la tecnología. Sin duda, el carácter disruptivo de la IA ha obligado a profundizar en esta cuestión.

En primer lugar, la filosofía de la ciencia y la tecnología hace años que ha desmontado la hipótesis de la neutralidad de las tecnologías. Aquello de que una tecnología no es ni buena ni mala, sino que depende del uso que los humanos le den, se ha mostrado como falso. La tecnología es un sistema intencional, es decir, se crea para algo, se crea con una finalidad que puede ser buena o mala. Luego la tecnología éticamente hablando no es neutra.

En la IA esto se percibe fácilmente en la cuestión de los sesgos que puede adquirir por los datos con los que se alimenta. Los sesgos son errores sistemáticos que pueden surgir al selec-

8. https://www.digitales.es/wp-content/uploads/2022/02/Informe_IA_Etica_en_Sanidad.pdf

cionar o manejar datos, y pueden conllevar distribuciones de bienes o servicios no equitativos, o errores de diagnóstico y medios terapéuticos de consecuencias letales. Normalmente, en el mundo de las aplicaciones de IA a la medicina, estos sesgos provienen de una injusticia previa, por ejemplo, la del insuficiente acceso a servicios de salud de algunos grupos de personas.

Ahora bien, el carácter exigitivo de la dignidad obliga a que todas las personas recibamos el mismo tipo de trato, el que se debe a la persona. La injusticia es una forma paradigmática de trato indigno. No hay cosa más inhumana que ser tratado injustamente. Así las cosas, en el desarrollo de aplicaciones de IA en salud hay que estar especialmente atento a evitar los sesgos, pues como afirma la [Declaración Universal de Derechos Humanos](#), en su [artículo 27](#):

“**Toda persona tiene derecho [...] a participar en el progreso científico y en los beneficios que de él resulten**”.

En segundo lugar, la IA en cuanto tecnología está dotada de lo que Laín Entralgo llama dominio. Lo que obliga a unas consideraciones éticas.

[Laín Entralgo](#) divide la historia de la tecnología en la etapa en la que esta imita a la naturaleza, en la que la domina y en la que la crea. En esta última

“**Por primera vez el hombre, enfrentado a la realidad que le rodea y a su propia realidad, comienza a sentirse -digámoslo con la precisión conceptual de Zubiri- cuasicreador**”.

Es decir, crea cosas que no existían, por ejemplo, fármacos o macromoléculas que no han existido nunca. En este sentido me parece acertado decir

que la ingeniería genética es una tecnología mucho más disruptiva que la IA.

La IA, no parece -por ahora- pertenecer a la etapa de la tecnología capaz de crear; más bien, a la de dominar. Con la IA no se crea nada que no existiera. Los resultados de la IA no son algo cualitativamente diferente de las afirmaciones de la inteligencia humana, pero dotan a los hombres de una capacidad nueva y más potente de dominio sobre la realidad porque su conocimiento y su capacidad de predicción aumenta exponencialmente.

Por tanto, el ser humano, dotado de una capacidad de dominio mayor debe ser consciente de la dignidad de las personas que no pueden ser dominadas.

Para un uso correcto, éticamente hablado, el poder de dominio del que dota la tecnología hay que centrarlo en la realidad y no en la persona. La tecnología de IA aplicada a la salud nos hace más capaces en la prevención, diagnóstico e incluso terapia y debe ser utilizada para promover el derecho a la salud y a la asistencia médica de la que habla el [artículo 25](#) de la Declaración de Derechos Humanos, pero no puede ser utilizada para crear rehenes de las industrias tecnológicas que desarrollan la IA. Las aplicaciones que se creen deben buscar llegar a todas las personas y no solo a las que sea rentable. Al igual que ocurre con los medicamentos huérfanos, la industria de la IA en salud debería atender a todas las necesidades y no solo a aquellas cuya aplicación tenga buen retorno económico.

4/1

El principio de autonomía, la algorética y la dignidad de la persona.

El principio bioético de la autonomía se ha convertido, en la práctica, en el principio primero y fundamental de la práctica bioética. Las aplicaciones de IA en salud, de hecho, impactan directamente en él.

9. Laín Entralgo, P. (1988) *Técnica, Ética y amistad médica*, en Gafo, J. (1988) *Fundamentación de la bioética y manipulación genética*, Comillas. Madrid.

LH n.339

Dentro del principio de autonomía se encuentran el consentimiento informado y el respeto a la privacidad. En el contexto de las aplicaciones de IA, es la privacidad la que presenta más desafíos éticos, ya que puede verse comprometida de al menos tres maneras.

En primer lugar, la forma de recopilación de datos puede realizarse sin el consentimiento explícito del paciente. Esto puede ocurrir si los responsables de las bases de datos adquieren la información clínica de manera ilegítima, o si el paciente ha consentido el uso de sus datos para un propósito específico y no para el entrenamiento general de la IA.

En segundo lugar, está el riesgo de reidentificación del paciente. La alta capacidad de procesamiento de datos permite cruzar información con noticias, redes sociales y otras fuentes, lo que puede llevar a la identificación de pacientes que inicialmente habían cedido sus datos bajo un anonimato simple.

En tercer lugar, aunque es necesario evaluar cuidadosamente la responsabilidad, existe la amenaza del robo de datos. Los datos son un recurso muy valioso y están expuestos a actividades delictivas.

Pueden ser robados y utilizados de manera ilegítima, lo que obliga a los gestores de datos a implementar medidas de protección rigurosas. Una negligencia o falta de cuidado en este aspecto puede resultar en una violación indirecta de la privacidad.

Estas tres posibles amenazas, entre otras, que el uso o abuso de la IA pueden tener sobre la autonomía del paciente es lo que obligan a desarrollar una algorética. Es decir, se debe buscar que los algoritmos -fundamento último de las aplicaciones de la IA- se creen, tal y como formularon Nick Bostrom y Eliezer Yudkowsky, con: responsabilidad, transparencia, auditoría, incorruptibilidad, predictibilidad y una tendencia a no hacer que víctimas inocentes griten con frustración indefensa.

Esto se traduce, entre otras cosas, en que sean sólidos y robustos, técnicamente hablando:

“Para que la IA sea fiable desde un punto de vista ético necesita tener solidez técnica, igual que la medicina, para ser éticamente fiable, tiene que ser eficaz. Así, la IA debe tener resistencia a los ataques y seguridad, lo que también incluye en su caso un plan de repliegue y seguridad general”¹⁰.

También deben ser transparentes. Esto es, que exista una trazabilidad de todo lo que ocurre, una explicabilidad de todo lo que sucede y una adecuada comunicación.

Por último, deben ser responsables, capaces de rendir cuentas. Por tanto, deben ser auditables; minimizar los efectos negativos y notificarlos cuando ocurran; ser eficaz en la búsqueda de equilibrios; y hay que buscar que se lleven a cabo las compensaciones en el caso de que haya víctimas de los sistemas de IA. Se abre aquí la enorme cuestión biojurídica del daño que puedan causar los sistemas basado en IA.

En todo esto está en juego la dignidad de la persona, en concreto el respeto a su interioridad, ya que la autonomía tal y como la entiende la Bioética clínica tiene que ver con la introspección, con el yo y el respeto que se le debe.

4/2

La ética de la profesión médica y la IA.

Las aplicaciones de IA en el campo de la salud impactan en la naturaleza y ejercicio de la profesión médica. Al menos, a mi juicio, de tres modos.

En primer lugar, en la naturaleza de la profesión sanitaria. Como en el caso de otras profesiones,

10. Navarro, I. (2023) Daños e Inteligencia Artificial. Aspectos éticos y jurídicos, en Amo Usanos, R. Inteligencia Artificial y Bioética, Comillas, Madrid (pp 169)

la IA puede alterar el desarrollo, e incluso la naturaleza de la profesión médica. La utilización de la IA en el diagnóstico de enfermedades está desplazando al médico de su lugar. El uso de IA en el diagnóstico convierte al médico en otro más -quizá no el más eficiente- de los gestores del conocimiento. La utilización de la IA en la gestión optimizada de los servicios de salud, lo reduce a una variable más que hay que rentabilizar. La posibilidad de la toma de decisiones terapéuticas por parte de la IA, lo convierten en irrelevante. Evidentemente estos usos pueden impactar en la dimensión subjetiva del trabajo del médico o del personal sanitario.

Decía **Juan Pablo II** en **Laborem exercem**:

“Como persona, el hombre es pues sujeto del trabajo. Como persona él trabaja, realiza varias acciones pertenecientes al proceso del trabajo; éstas, independientemente de su contenido objetivo, han de servir todas ellas a la realización de su humanidad, al perfeccionamiento de esa vocación de persona, que tiene en virtud de su misma humanidad”¹¹.

Es decir, que el trabajo nos dignifica cuando nos realiza como personas, colaboradores de la obra de Dios por estar creados a su imagen y semejanza.

Las posibilidades de la IA en el ejercicio de la profesión sanitaria exigen una reflexión sobre la esencia misma de la profesión para que esta no se convierta en algo alienante.

Como afirmaba **Marx**, el trabajo que no realiza a la persona puede volver extraño al trabajador ya sea porque no reconoce su mano en el producto de su trabajo, en este caso en la salud del enfermo; ya sea porque pierde el control sobre el proceso de su trabajo; ya sea porque no siente que el resultado de su acto médico es fruto de su trabajo.

En segundo lugar, la IA impacta de lleno en la relación sanitario paciente. Ésta, decía Laín Entralgo, no debe ser una relación de dominio de una parte sobre otra vulnerable. Tampoco debe ser calificada como de cooperación. Según esta comprensión el médico y el enfermo se comprenden

“Como camaradas que recorren un mismo camino, hacia una misma meta que es la curación del enfermo. Son camaradas que, por tanto, se ayudan mutuamente”¹².

En este sentido, advertía que la técnica puede contribuir a la comprensión de la relación sanitario-paciente como cooperación ya que podría considerarse como el instrumento que ayuda a ambos en este camino.

Si la técnica utilizada es tan eficaz como la IA, puede que pase a ocupar el puesto principal en esta relación de cooperación sustituyendo a uno, el médico; y dotando de un papel más pasivo al otro, el enfermo que puede ser considerado como un objeto sobre el que la técnica de la IA muestra su eficacia. La relación sanitario paciente, según **Laín Entralgo**, es una forma singular de amistad, no solo de camaradería:

“La relación óptima entre el médico y el enfermo es la amistad médica, un modo de la relación entre hombre y hombre, distinto de la projimidad y de la simple camaradería. El médico procura el bien del enfermo, dice el bien del enfermo, hace el bien al enfermo, recibe confidencias del enfermo, y lo hace situándose entre la tiranía y el puro paternalismo”¹³.

Este tipo de amistad es una de las formas más excelsas de trato digno. Un amigo nunca es un

11. Juan Pablo II, *Laborem Exercem*, 6.

12. Laín Entralgo, P. (1988) *Técnica, Ética y amistad médica*, en Gafo, J. (1988) *Fundamentación de la bioética y manipulación genética*, Comillas. Madrid. (pp. 115)

13. Laín Entralgo, P. (1988) *Técnica, Ética y amistad médica*, en Gafo, J. (1988) *Fundamentación de la bioética y manipulación genética*, Comilla. Madrid. (pp. 117)

LH n.339

medio, siempre es un fin. Cualquier uso de la IA en la relación sanitario paciente debe evitar interponerse entre dos amigos, sino que debe favorecer su relación de amistad.

En tercer lugar, uno de los aspectos más distópicos del uso de la IA en la medicina es la posibilidad de que las decisiones terapéuticas, o incluso las intervenciones quirúrgicas, sean realizadas de manera autónoma por una IA. Esto es, que sea la propia IA la responsable de la decisión sin intervención humana.

El papa Francisco advierte, al referirse al uso de las armas autónomas letales, que la IA nunca

“Debería elegir jamás poner fin a la vida del ser humano”¹⁴.

Aunque el contexto de la afirmación es distinto al de la bioética, el razonamiento sobre el que se apoya es muy sólido.

La IA solo puede elegir, no puede decidir. Las decisiones de la IA son elecciones entre varias posibilidades y se basa en inferencias estadísticas. Por ello

“Necesitamos garantizar y proteger un espacio de control significativo del ser humano sobre el proceso de elección utilizado por los programas de inteligencia artificial. Está en juego la misma dignidad humana”¹⁵.

Las cuestiones referentes a la vida del ser humano y, por ende, a su salud, deben ser tomadas por humanos que pueden decidir; esto es, que pueden y deben tomar decisiones prudentes. Desde la Ética a Nicómaco la prudencia se comprende como algo más que la elección. Implica, entre otras cosas, una comprensión del bien y

una evaluación de las circunstancias. Esto es algo más que una elección entre posibilidades estadísticas, es una deliberación. Un acto propiamente humano.

4/3

La necesidad ética de la formación de los médicos.

En la tesitura del ejercicio de una profesión como la medicina en la que una de las partes está en situación de especial vulnerabilidad hay que tener afinado el sentido de la dignidad humana.

La historia demuestra que la dignidad no es una lección fácil de aprender. Parece mentira, pero a estas alturas de siglo XXI más de cuarenta millones de personas en el mundo padecen algún tipo de esclavitud en el tráfico de personas.

La existencia de la pobreza y de la guerra ponen encima de la mesa las violaciones de la dignidad. Los atentados contra la dignidad humana están muy cerca de nosotros y no estamos exentos que las aplicaciones de la IA en salud puedan convertir a los pacientes en medios y no en fines.

Como se ha repetido en tantas ocasiones, la IA está afectando a la práctica médica, pero también lo hará en su formación tanto inicial como continua. Los médicos en ejercicio sienten la necesidad de incorporar conocimientos de IA a su día a día. Para ser honestos con su profesión y poder ejercerla de la mejor manera posible en su compromiso con la salud de los pacientes, deben formarse continuamente.

Paralelamente a esta necesidad corre la obligación de dotar de un armazón bioético a los estudiantes de medicina y a los médicos en ejercicio.

Existe este imperativo de la incorporación de la ética de la IA en los planes de estudio de medicina, enfermería y otras ciencias y en los planes de formación continua de los médicos y enfermeros.

14. Francisco, Discurso del Santo Padre en la sesión del G7 sobre Inteligencia Artificial. <https://www.vatican.va/content/francesco/es/speeches/2024/june/documents/20240614-g7-intelligenza-artificiale.html>

15. Francisco, Discurso del Santo Padre en la sesión del G7 sobre Inteligencia Artificial. <https://www.vatican.va/content/francesco/es/speeches/2024/june/documents/20240614-g7-intelligenza-artificiale.html>

Probablemente una de las formaciones más urgentes que haya que desarrollar es el posible uso de la IA en los Comités de ética. Es más que probable que médicos o Comités de ética clínica o de investigación hayan utilizado la IA para consultar problemas éticos.

Normalmente la IA está programada para no afrontar cuestiones de esta naturaleza. Ahora bien, no sería mala opción alimentar la IA con gran número de casos de Bioética para que los médicos o los Comités de ética puedan consultar las resoluciones a casos similares a los que se les presenten.

No obstante, hay que tener en cuenta que el proceso deliberativo que conduce a la elección de un curso de acción en virtud de unos valores no es simple proceso de elección entre varias posibilidades¹⁶.

Deliberar implica poner en juego los valores y éstos dependen de un proceso no solo intelectual, sino como recuerda **Diego Gracia** utilizando la filosofía de Zubiri, también sentimental y volitivo. El valor se construye tras captar las posibilidades de lo real en un proceso de la inteligencia sentiente y se proyecta en una acción libre de la voluntad que no es ajeno al sentimiento y a la inteligencia.

De otro modo: deliberar implica no solo la inteligencia para razonar -lo que sí podría hacer la IA- sino el sentimiento y la voluntad, de un ser -la persona- dotado de cuerpo e inserto en una cultura y una historia.

La IA alimentada con casos de comités de ética puede ser de gran ayuda para el proceso deliberativo, pero solo una ayuda y nunca un sustituto.

05/

Conclusión.

El pasado 22 de junio, el papa Francisco, en un discurso ante Fundación Centesimus Annus pro Pontifice, pidió

“No perder la ocasión de pensar y actuar en un modo nuevo, con la mente, el corazón y las manos, para dirigir la innovación tecnológica hacia una configuración centrada en el primado de la dignidad humana”¹⁷.

Este puede ser un camino muy fructífero para pensar la ética de los desarrollos de la IA en salud.

16. Francisco, Discurso del Santo Padre en la sesión del G7 sobre Inteligencia Artificial. <https://www.vatican.va/content/francesco/es/speeches/2024/june/documents/20240614-g7-intelligenza-artificiale.html>

17. <https://www.vaticannews.va/es/vaticano/news/2024-06/centesimus-annus-conferencia-inteligencia-artificial-junio-24.html> -

Bibliografía

- ▶ **Amo Usanos, R. (2022).**
El valor de la vida en una sociedad envejecida.
En Amo Usanos, R. (Ed.), *Bioética para una sociedad envejecida* (pp. 61-104). Madrid: Universidad Pontificia Comillas.
- ▶ **Amo Usanos, R. (2022).**
La persona, ser valioso. La dignidad: historia, contenido y teología.
En Aparicio Malo, J. M. (Coord.), *Cristianismo y ética social* (pp. 97-122). Madrid: Universidad Pontificia Comillas.
- ▶ **Amo Usanos, R. (2023).**
Las aplicaciones de inteligencia artificial en medicina.
Un reto para la bioética. Madrid Médico, 176, 19.
- ▶ **Concilio Vaticano II. (1965).**
Constitución Pastoral Gaudium et spes.
Recuperado de https://www.vatican.va/archive/hist_councils/ii_vatican_council/documents/vat-ii_const_19651207_gaudium-et-spes_sp.html
- ▶ **Dicasterio para la Doctrina de la Fe. (2024).**
Dignitas infinita.
Recuperado de https://www.vatican.va/content/john-paul-ii/es/encyclicals/documents/hf_jp-ii_enc_25031995_evangelium-vitae.html
- ▶ **Francisco. (2024, 13-15 de junio).**
El Papa Francisco participa en la sesión del G7 sobre inteligencia artificial.
Recuperado de <https://www.vatican.va/content/francesco/es/speeches/2024/june/documents/20240614-g7-intelligenza-artificiale.html>
- ▶ *Inteligencia Artificial Ética en Sanidad.* (2022, febrero).
DigitalES.
Recuperado de https://www.digitales.es/wp-content/uploads/2022/02/Informe_IA_Etica_en_Sanidad.pdf
- ▶ **Juan Pablo II. (1981).**
Carta encíclica Laborem Exercem.
Recuperado de https://www.vatican.va/content/john-paul-ii/es/encyclicals/documents/hf_jp-ii_enc_14091981_laborem-exercens.html
- ▶ **Juan Pablo II. (1995).**
Carta encíclica Evangelium.
Recuperado de https://www.vatican.va/content/john-paul-ii/es/encyclicals/documents/hf_jp-ii_enc_25031995_evangelium-vitae.html
- ▶ **Laín Entralgo, P. (1988).**
Técnica, ética y amistad médica.
En Gafo, J. (Ed.), *Fundamentación de la bioética y manipulación genética.* Madrid: Comillas.
- ▶ **Melendo Granados, T. (2009).**
Sobre el hombre y su dignidad.
Persona y Bioética, (6).
Recuperado de <https://personaybioetica.unisabana.edu.co/index.php/personaybioetica/article/view/706>
- ▶ **Naciones Unidas. (1948).**
Declaración universal de derechos humanos.
Recuperado de <https://www.un.org/es/about-us/universal-declaration-of-human-rights>

▶ Navarro, I. (2023).

*Daños e inteligencia artificial.
Aspectos éticos y jurídicos.*

En Amo Usanos, R. (Ed.), *Inteligencia artificial y bioética*. Madrid: Comillas.

▶ Peñas García, L. (2024).

La ética de la inteligencia artificial generativa.
Labor Hospitalaria, 338, 17-81.

▶ Pontificia Academia para la Vida. (2020).

Roma Call for IA Ethics.

Recuperado de

<https://www.romecall.org/the-call/>

▶ Urgelés Puertólas, D. (2023).

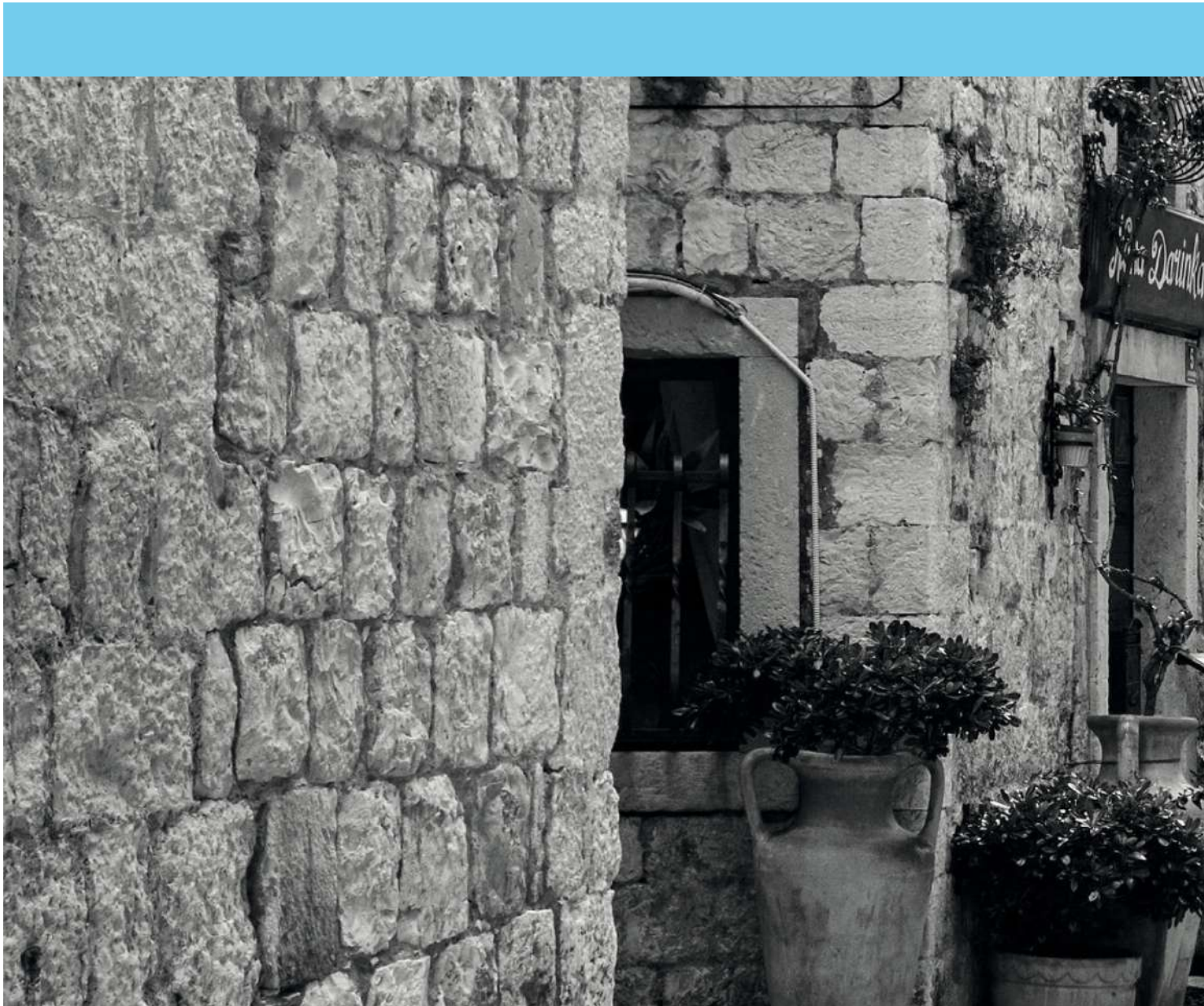
Implicaciones a corto plazo de la inteligencia artificial en la atención sanitaria.

Labor Hospitalaria, 336-337, 37-43.



05/

Experiencias





05/1

Algunos retos y soluciones para una Inteligencia Artificial más ética y fiable en la sanidad

Carina Dantas.

CEO. SHINE 2Europe. Coimbra (Portugal).

Miriam Cabrita.

Gestor de proyectos. SHINE 2Europe. Coimbra (Portugal).

Luís Midão.

Gestor de proyectos. Centro de Investigação em Tecnologias e Serviços de Saúde (CINTESIS). Porto (Portugal).

Luís Midão.

Profesor. Centro de Investigação em Tecnologias e Serviços de Saúde (CINTESIS). Porto (Portugal).



Las cuestiones éticas relacionadas con la Inteligencia Artificial (I.A.) son un tema emergente y en constante evolución. Ocupan un lugar central en la agenda política europea, pero no ha sido fácil alcanzar un amplio consenso sin largas y difíciles negociaciones.

Los retos éticos a los que hay que hacer frente son complejos y multidisciplinares, por lo que su clasificación facilita un análisis detallado. A partir del trabajo realizado sobre el uso de soluciones digitales basadas en la Inteligencia Artificial en la atención sanitaria, mediante una revisión bibliográfica, talleres y entrevistas con expertos, profesionales y pacientes, se proponen las siguientes categorías: (1) Retos individuales (humanos); (2) Retos técnicos; (3) Retos organizativos; y (4) Retos sociales.

Se necesita un modelo global que analice los retos éticos desde la ideación de cualquier solución tecnológica y los aborde de forma iterativa, con una amplia participación de las distintas partes interesadas, para promover la inclusión, la igualdad de acceso y un uso más justo de la Inteligencia Artificial en la sanidad.

1/

Introducción.

La Inteligencia Artificial (I.A.) trae consigo la expectativa de mejoras en la vida cotidiana de los ciudadanos, prometiendo remodelar profundamente la industria, la sanidad y la sociedad en su conjunto.

Desde el tratamiento médico a la ciberseguridad, pasando por los vehículos autónomos y las decisiones jurídicas, se espera que la I.A. aumente la eficiencia automatizando procesos y mejorando la toma de decisiones, acelerando el análisis y la evaluación de grandes volúmenes

de datos. Sin embargo, a medida que estas tecnologías se desarrollan, prueban y aplican, surgen inquietudes sobre sus implicaciones éticas, sociales y jurídicas.

Las convenciones y normas sociales establecen expectativas predefinidas; éstas, junto con los valores, principios, bagaje cultural, tradiciones y experiencias vitales personales, contribuyen a conformar la visión del mundo de un individuo, interiorizando determinados supuestos o prejuicios. La I.A. no existe en el vacío, sino que se construye a partir de algoritmos creados por humanos, transponiendo así estos prejuicios y reproduciendo el juicio humano. Comprender las fuentes de los sesgos algorítmicos y de los datos permite diversificar los conjuntos de datos; ser más consciente de los posibles sesgos sociales e individuales puede ayudar a mitigarlos a la hora de desarrollar soluciones de I.A.

Aunque algunas de estas desigualdades son sistémicas, por ejemplo, las diferencias socioeconómicas, los prejuicios humanos también desempeñan un papel importante. Una de las causas más comunes de la falta de equidad en la I.A. utilizada en el ámbito médico es el sesgo de los datos utilizados para entrenar los algoritmos (Ricci Lara et al., 2022). Como resultado, muchas de las cohortes clínicas y de investigación existentes que pueden utilizarse para entrenar algoritmos de I.A. médica ya incorporan una representación errónea en los datos.

Los datos de entrenamiento pueden no estar completos, no ser representativos de la población real o incluir otros sesgos relevantes debidos a los parámetros establecidos para la recopilación de datos. Por ejemplo, la mayoría de los conjuntos de datos sobre cáncer se refieren a pacientes mayores de 60 años. Sin embargo, en los últimos años se ha diagnosticado cáncer a un gran número de pacientes más jóvenes, lo que implica que los conjuntos de datos existentes pueden no corresponder a la población real actual (véase el ejemplo del cáncer de intestino en Davis, 2018). Al mismo tiempo, las comunidades LGBTIQ+ no

se tienen en cuenta específicamente en la vigilancia del cáncer, lo que se traduce en una escasez de información sobre la supervivencia al cáncer de estas poblaciones (Boehmer et al., 2011). Este tipo de prejuicios también se observa en la I.A. utilizada para detectar y tratar problemas de salud como el cáncer de piel. Las imágenes de pacientes blancos se utilizan predominantemente para entrenar algoritmos de detección de melanomas, lo que puede dar lugar a peores resultados para las personas de raza negra (Kleinberg et al., 2022).

Sin embargo, los retos éticos no se limitan a los datos. Los sesgos cognitivos (Hardman, 2010) entran en la categoría de prejuicios humanos o sociales y se refieren a los sentimientos hacia una persona o un grupo de personas basados en la percepción de su pertenencia a un grupo. Una lista de más de 180 prejuicios humanos, definidos y clasificados por psicólogos, muestra lo que puede afectar a los individuos cuando tienen que tomar una decisión y, por tanto, puede infiltrarse en los algoritmos de aprendizaje automático cuando se diseñan.

Un estudio de Cowgill et al. (2020) concluyó que

«Cuanto más homogéneo es el equipo [de ingeniería], más probable es que aparezca un error de predicción determinado».

La falta de diversidad, multidisciplinariedad y educación ética en los equipos de desarrollo puede crear una falta de sensibilidad moral y de conciencia hacia las personas que se enfrentan a problemas de discriminación, lo que conduce a una introducción inconsciente de prejuicios en los algoritmos.

Se trata, por tanto, de un área compleja, y esta complejidad aumenta cuando se analizan las diferentes áreas o sectores de actividad en su granularidad, ya que muchos retos éticos son

LH n.339

transversales, pero otros son muy específicos de cada área. Aunque hasta el momento se ha trabajado mucho en diversas recomendaciones de alto nivel, se necesitan herramientas coherentes y armonizadas que puedan aplicarse a los ámbitos específicos de cada solución de Inteligencia Artificial, especialmente a aquellos que implican mayores riesgos de causar daños a los ciudadanos.

Para abordar estos desafíos, la Comisión Europea creó en 2018 el **Grupo Independiente de Expertos de Alto Nivel sobre Inteligencia Artificial (AI HLEG)** con el fin de establecer requisitos generales para una I.A. digna de confianza en Europa. Este Grupo publicó una lista de verificación de evaluación (Comisión Europea. Dirección General de Redes, Contenidos y Tecnologías de la Comunicación - CNECT -, 2020) que proporciona directrices y preguntas para fomentar el desarrollo de la I.A. hacia una mayor fiabilidad. Aunque esta lista de comprobación es un recurso valioso, sus principios son genéricos para todos los ámbitos de la I.A. y la sociedad, y su aplicación en diferentes áreas, en particular la sanidad, requiere marcos y soluciones adaptados.

En este contexto, para garantizar una mayor aceptación y uso de las herramientas de I.A. en el sistema sanitario del futuro, se necesitan nuevos enfoques para aumentar la concienciación y la inclusión, así como métodos para validarlas y aplicarlas en plena consonancia con la diversidad de necesidades y contextos del mundo real.

2/

Resultados.

A partir del trabajo de los autores, es decir, el análisis de la bibliografía existente, combinado con las opiniones de expertos en ética, regula-

ción, ciencias políticas, profesionales sanitarios y tecnológicos, así como pacientes y asociaciones de pacientes, fue posible resumir algunas de las cuestiones más relevantes que deben analizarse en relación con los retos éticos relacionados con el uso de la inteligencia artificial en soluciones sanitarias y las posibles formas de mitigarlos.

Muchos estudios se basan en muestras sesgadas de pacientes y pequeños conjuntos de datos de una sola organización, sin validación externa (**Bhowmik & Eskreis-Winkler, 2022**), lo que lleva a cuestionar la precisión de las herramientas de I.A. En este sentido, las bases de datos compartidas entre instituciones pueden ayudar a resolver algunas de estas dificultades, siempre que se garantice la privacidad del paciente (**Houssami et al., 2019**).

Además, muchos estudios sobre herramientas basadas en la inteligencia artificial son retrospectivos y carecen de poder predictivo (**Gong et al., 2022; Liang et al., 2022; Vedantham et al., 2023**). Se necesitan estudios prospectivos multicéntricos a gran escala con datos exhaustivos y directrices de evaluación de la calidad para avanzar hacia la integración en la práctica clínica (**Bhowmik y Eskreis-Winkler, 2022; Malherbe, 2021**).

Sin embargo, la creciente abundancia de datos que se requiere también está planteando retos cada vez mayores a los sistemas sanitarios. A menudo, los centros sanitarios no disponen de un número adecuado de expertos formados para anotar estos datos, lo cual es esencial para los modelos de I.A. (como se explica en **Goisau y Cano Abadía, 2022**). Tanto los conjuntos de datos de entrenamiento como los de prueba requieren una anotación meticulosa y el etiquetado de metadatos, utilizando formatos estandarizados para permitir la evaluación y optimización de la equidad algorítmica (**Chua et al., 2021**). Por otro lado, garantizar la total privacidad de los datos de los pacientes es una prioridad clara, tanto a efectos de cumplimiento legal como ético. Sin embargo, anonimizar completamente los datos es un reto importante y a menudo sig-

nifica que los resultados obtenidos por la I.A. no pueden compararse con la evolución clínica del paciente, lo que obstaculiza la investigación y las oportunidades de seguir aprendiendo. El aprendizaje federado es uno de los enfoques utilizados actualmente para eludir las restricciones de privacidad (Kandati & Anusha, 2023), ya que permite el entrenamiento colaborativo de algoritmos sin compartir historiales médicos confidenciales, puesto que los datos permanecen almacenados en los servidores originales.

Esta necesidad de grandes conjuntos de datos y de profesionales altamente cualificados preocupa a los sistemas sanitarios, especialmente cuando se trata de herramientas patentadas basadas en la inteligencia artificial, ya que estas solo pueden estar al alcance de instituciones con mayores recursos financieros.

Esto podría agravar las desigualdades en el acceso a la mejor atención sanitaria, como destacan Goisauf y Cano Abadía (2022). Además, la falta de análisis coste-efectividad y de investigación basada en la evidencia sobre las herramientas de I.A. hace que no sean fáciles de incluir en los planes de reembolso y en los paquetes de seguros. Según los expertos consultados, para ganarse la confianza de los pagadores, una herramienta basada en la I.A. debe ser más eficiente, tener menos errores y ser capaz de trabajar 24 horas al día, en comparación con la mano de obra humana.

Además, los gastos asociados a las herramientas basadas en I.A. tienen el potencial de aumentar la disparidad entre las personas más ricas y aquellas con menores ingresos, si los conjuntos de datos utilizados para entrenar estas herramientas no incluyen datos de personas con un estatus socioeconómico más bajo, como señalan Cui et al. (2019), y por lo tanto no consideran variables potencialmente relevantes para las personas en situación de vulnerabilidad.

Otro reto emergente es la inclusión de la I.A. en la aplicación nacional del Reglamento sobre los Productos Sanitarios (Radley-Gardner et al., 2016).

Los procedimientos no son uniformes en toda Europa y los equipos de investigación carecen de directrices y procedimientos claros sobre este tema.

También desde el punto de vista de la investigación, la mayoría de los formularios de consentimiento informado obtenidos por las organizaciones sanitarias en relación con los datos de los pacientes no incluyen específicamente la formación en algoritmos, ya que la mayoría de estos consentimientos se recogieron hace mucho tiempo y no constituían una necesidad clara en ese momento. ¿Significa esto que los datos no pueden utilizarse para este fin? ¿Existe alguna excepción de interés público que se aplique a pesar de la falta de consentimiento claro para este fin? (Reddy et al., 2020). En otro aspecto relacionado con el consentimiento, autores como Hickman et al. (2021) plantearon la cuestión de si

“Antes de realizar cálculos como el riesgo de desarrollar una enfermedad, el paciente debería tener que aprobar este análisis tras recibir asesoramiento por parte de un profesional sanitario, de forma similar a los procedimientos previstos actualmente para las pruebas genéticas”.

¿Debería tenerse en cuenta la opinión del paciente a la hora de utilizar este tipo de herramientas? Todas estas cuestiones abiertas, así como los aspectos normativos relacionados, son por tanto ámbitos clave para seguir investigando y elaborando políticas.

El uso de herramientas de I.A. también plantea la cuestión de la responsabilidad (D'Amore et al., 2021; Houssami et al., 2019; Pesapane et al., 2022): ¿quién debe ser considerado responsable en caso de errores de clasificación? ¿Son los algoritmos comprensibles, incluso para profesionales experimentados, como señalan Iqbal y sus colegas (2021)?

Se espera que una herramienta de I.A. proporcione asesoramiento y que la decisión final recaiga en la experiencia del profesional sanitario. Sin embargo, pueden darse situaciones en las que la tecnología sea tan precisa y fiable que la confianza en sus datos resulte inevitable. Esto pone de relieve la compleja relación entre confianza y responsabilidad en el contexto de la asistencia sanitaria basada en la I.A.

Por ello, la madurez de la propia tecnología desempeña un papel importante a la hora de asignar responsabilidades (Retos de la responsabilidad en las tecnologías médicas de IA, 2022). Aunque la legislación vigente intenta identificar a los posibles responsables de predicciones “erróneas” (programadores, médicos, proveedores de hardware), la cuestión sigue siendo ambigua, especialmente con modelos de inteligencia artificial cada vez más complejos (Schneeberger et al., 2020).

Para navegar por este panorama en evolución, son importantes varias estrategias clave. En primer lugar, contar con una legislación nacional e internacional adecuada y/o directrices hospitalarias que se actualicen a un ritmo más rápido y estén vinculadas a resultados basados en la evidencia.

Además, promover enfoques colaborativos de toma de decisiones que incluyan a médicos, pacientes y otros profesionales relevantes puede mitigar los riesgos y garantizar la responsabilidad compartida (Hamilton et al., 2016). Siempre que sea posible, dar a los pacientes la oportunidad de participar activamente en las decisiones de tratamiento fomenta la confianza y la transparencia durante todo el proceso.

Aplicando estas medidas, podemos garantizar que las herramientas de I.A. no sólo mejoren los resultados clínicos, sino que también funcionen dentro de un marco de principios éticos y jurídicos claros.

3/

Discusión.

Es fundamental garantizar la diversidad en todo el proceso de desarrollo de la I.A., desde la concepción hasta la aplicación. Los desarrolladores, los profesionales sanitarios y los usuarios finales deben recibir información, formación y conocimientos adecuados para identificar y mitigar posibles sesgos. Esto subraya la importancia de desarrollar materiales educativos y de formación para todos los grupos de partes interesadas, promoviendo la alfabetización en I.A. y garantizando que los profesionales sanitarios (Lokaj et al., 2023), los pacientes y otras partes interesadas pertinentes tengan los conocimientos y las habilidades necesarios para participar eficazmente y beneficiarse de las soluciones basadas en I.A. implementadas en la práctica diaria (Lekadir et al., 2022).

Los debates en la literatura sobre el sesgo en la I.A. a menudo ignoran sus raíces más profundas y sus efectos más amplios. Cuando el sesgo se considera únicamente un problema técnico con una solución técnica, se pasan por alto aspectos importantes. Por lo tanto, es esencial incluir puntos de vista de las ciencias sociales cuando se exploran las preocupaciones éticas y sociales asociadas al uso de la I.A. (Goisauf & Cano Abadía, 2022; Jongsma & Bredenoord, 2020).

Una cuestión relevante es la necesidad de garantizar que las predicciones proporcionadas por las soluciones basadas en la I.A. sean comprensibles, no sólo para los profesionales sanitarios, sino también para los pacientes y cuidadores. Para generar confianza en los pacientes, es crucial proporcionar información sobre la eficacia y el rendimiento general de la herramienta (por ejemplo, incluyendo información sobre la sensibilidad y la especificidad). Como se destaca en el estudio de Fulmer & Dirks (2018), la confianza abarca varias dimensiones: la confianza del hospital para adquirir la herramienta orientada a la

I.A., la confianza en la tecnología por parte de la industria que la desarrolla, la confianza de los médicos para utilizarla y, por último, pero no por ello menos importante, también la confianza del paciente.

En lo que respecta a los pacientes, la situación es aún más compleja, ya que su confianza en las herramientas de I.A. se ve a menudo influida por su confianza en el hospital e incluso en un médico concreto. Sin embargo, ¿puede la dinámica de poder entre médico y paciente impedir que los pacientes expresen preocupaciones válidas (Ozdemir y Finkelstein, 2018)? Equilibrar la confianza y el conocimiento siguen siendo cuestiones relevantes en la ética sanitaria.

La validación científica de los algoritmos sigue siendo una preocupación clave para los profesionales sanitarios y asistenciales. La falta de procedimientos de validación sólidos, unida a la falta de alineación con los flujos de trabajo clínicos establecidos, incluida la escasa integración con las herramientas existentes (Tsopra et al., 2021), plantea problemas. Además, la ausencia de mecanismos de certificación adecuados complica aún más las cosas cuando se trata de herramientas basadas en la I.A.

Para fomentar la confianza entre los profesionales sanitarios, una estrategia puede ser proporcionar documentación de seguimiento sobre el rendimiento, la usabilidad y la integración de la herramienta basada en la I.A. en las prácticas asistenciales de otras organizaciones sanitarias. Aunque los profesionales sanitarios no necesitan profundizar en detalles técnicos complejos, deben ser capaces de entender cómo los modelos han llegado a predicciones específicas.

Proporcionar un porcentaje de precisión para cada predicción y explicar el razonamiento que la sustenta son pasos cruciales. Al fin y al cabo, los pacientes, sobre todo después de un diagnóstico, se encuentran en un estado vulnerable y es posible que sólo asimilen una parte de la información. La claridad y la comunicación eficaz son fundamentales para generar

confianza y garantizar una toma de decisiones informada.

Basándose en todos los elementos recogidos y debatidos, los autores proponen la siguiente categorización de los retos y prejuicios éticos:

1. Desafíos individuales (humanos), como la falta de educación y formación adecuadas, las creencias y los prejuicios individuales.
2. Desafíos técnicos, por ejemplo, en el diseño defectuoso de algoritmos, conjuntos de datos de entrenamiento sesgados e interpretación incorrecta de los resultados.
3. Desafíos organizativos, por ejemplo, falta de diversidad en los equipos, falta de métodos de auditoría, falta de herramientas disponibles.
4. Desafíos societales, como desigualdades sanitarias, discriminación, problemas éticos y legales, falta de regulación adecuada o de políticas públicas.

Cada una de estas categorías puede ampliarse y adaptarse con retos de mayor nivel o más específicos, ajustándola a los casos prácticos concretos que se analicen. Abordar estos retos de forma colectiva es vital para el uso ético y eficaz de la I.A. en la asistencia sanitaria.

Las posibles medidas de mitigación también pueden agruparse en estas mismas cuatro áreas:

Por ejemplo, en lo que respecta a los factores individuales (humanos), es muy importante el desarrollo de materiales de formación y/o educación específicos para cada tipo de destinatario.

En cuanto a los factores técnicos, vinculados principalmente a la industria, debe hacerse énfasis en la mejora de la solidez de los resultados de la Inteligencia Artificial, que necesitan conjuntos de datos fiables, de buena calidad y de gran tamaño. La auditoría y supervisión periódicas de los algoritmos es esencial para garantizar un buen rendimiento a medio y largo plazo.

LH n.339

En cuanto a los aspectos organizativos, asegurar equipos multidisciplinares con la formación adecuada y aplicar métodos que garanticen la participación de los ciudadanos y las organizaciones de la sociedad civil en el diseño y desarrollo de soluciones de I.A. es clave para superar los prejuicios y retos detectados.

Las medidas societales abarcan desde la participación de las partes interesadas hasta los aspectos políticos, éticos, jurídicos y normativos. Los métodos procedentes de la innovación social y las ciencias sociales, combinados con los conocimientos técnicos y clínicos, son útiles y deben seguir investigándose y aplicándose con vistas a una sociedad más integradora.

4/

Conclusión.

Del análisis realizado se ha podido concluir que los retos a los que se enfrenta el desarrollo de herramientas de I.A. en el ámbito sanitario son multifacéticos y deben ser estructurados y analizados cuidadosamente.

El modelo propuesto en este artículo para agrupar los retos y sesgos éticos del uso de la I.A. en la sanidad pretende facilitar el análisis y la resolución de estos retos hacia una sociedad más inclusiva, justa y digna de confianza.

Bibliografía

Bhowmik, A., & Eskreis-Winkler, S. (2022). *Deep learning in breast imaging*. *BJR Open*, 4(1), 20210060. <https://doi.org/10.1259/bjro.20210060>

Boehmer, U., Miao, X., & Ozonoff, A. (2011). *Cancer survivorship and sexual orientation*. *Cancer*, 117(16), 3796–3804. <https://doi.org/10.1002/cncr.25950>

Chua, I., Gaziel-Yablowitz, M., Korach, Z., Kehl, K., Levitan, N., Arriaga, Y., Jackson, G., Bates, D., & Hassett, M. (2021). *Artificial intelligence in oncology: Path to implementation*. *Cancer Medicine*, 10(12), 4138–4149. <https://doi.org/10.1002/cam4.3935>

Coordinated Plan on Artificial Intelligence 2021 Review (Shaping Europe's Digital Future). (2021). **European Commission**. <https://digital-strategy.ec.europa.eu/en/library/coordinated-plan-artificial-intelligence-2021-review>

Cowgill, B., Dell'Acqua, F., Deng, S., Hsu, D., Verma, N., & Chaintreau, A. (2020). *Biased Programmers? Or Biased Data? A Field Experiment in Operationalizing AI Ethics* (arXiv:2012.02394). arXiv. <https://doi.org/10.48550/arXiv.2012.02394>

Cui, C., Chou, S.-H. S., Brattain, L., Lehman, C. D., & Samir, A. E. (2019). *Data Engineering for Machine Learning in Women's Imaging and Beyond*. *American Journal of Roentgenology*, 213(1), 216–226. <https://doi.org/10.2214/AJR.18.20464>

Davis, N. (2018, October 22).

Bowel cancer on the rise among young people in Europe.

The Guardian.

<https://www.theguardian.com/society/2018/oct/22/bowel-cancer-on-the-rise-among-young-people-in-europe>

European Commission. Directorate General for Communications Networks, Content and Technology. (2020).

The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment.

Publications Office.

<https://data.europa.eu/doi/10.2759/791819>

Fulmer, A., & Dirks, K. (2018).

Multilevel trust: A theoretical and practical imperative.

Journal of Trust Research, 8(2), 137–141.

<https://doi.org/10.1080/21515581.2018.1531657>

Goisau, M., & Cano Abadía, M. (2022).

Ethics of AI in Radiology:

A Review of Ethical and Societal Implications.

Frontiers in Big Data, 5, 850383.

<https://doi.org/10.3389/fdata.2022.850383>

Gong, X., Guo, Y., Zhu, T., Peng, X., Xing, D., & Zhang, M. (2022).

Diagnostic performance of radiomics in predicting axillary lymph node metastasis in breast cancer: A systematic review and meta-analysis.

Frontiers in Oncology, 12. Scopus.

<https://doi.org/10.3389/fonc.2022.1046005>

Hamilton, D. W., Heaven, B., Thomson, R. G., Wilson, J. A., & Exley, C. (2016).

Multidisciplinary team decision-making in cancer and the absent patient: A qualitative study.

BMJ Open, 6(7), e012559.

<https://doi.org/10.1136/bmjopen-2016-012559>

Hardman, D. (2010).

Judgment and decision making:

Psychological perspectives (Reprinted).

BPS Blackwell.

Hickman, S. E., Baxter, G.

C., & Gilbert, F. J. (2021).

Adoption of artificial intelligence in breast imaging: Evaluation, ethical constraints and limitations.

British Journal of Cancer, 125(1), 15–22.

<https://doi.org/10.1038/s41416-021-01333-w>

Houssami, N., Kirkpatrick-Jones, G.,

Noguchi, N., & Lee, C. I. (2019).

Artificial Intelligence (AI) for the early detection of breast cancer: A scoping review to assess AI's potential in breast screening practice.

Expert Review of Medical Devices, 16(5),

351–362. <https://doi.org/10.1080/17434440.2019.1610387>

Iqbal, M. J., Javed, Z., Sadia, H., Qureshi,

I. A., Irshad, A., Ahmed, R., Malik, K.,

Raza, S., Abbas, A., Pezzani,

R., & Sharifi-Rad, J. (2021).

Clinical applications of artificial intelligence and machine learning in cancer diagnosis: Looking into the future.

Cancer Cell International, 21(1), 270.

<https://doi.org/10.1186/s12935-021-01981-1>

Jongsma, K. R., & Bredenoord, A. L. (2020).

Ethics parallel research: An approach for (early) ethical guidance of biomedical innovation.

BMC Medical Ethics, 21(1), 81. <https://doi.org/10.1186/s12910-020-00524-z>

- ▶ **Kandati, D. R., & Anusha, S. (2023).** *Security and privacy in federated learning: A survey.* Trends in Computer Science and Information Technology, 8(2), 029–037. <https://doi.org/10.17352/tcsit.000066>
-
- ▶ **Kleinberg, G., Diaz, M. J., Batchu, S., & Lucke-Wold, B. (2022).** *Racial underrepresentation in dermatological datasets leads to biased machine learning models and inequitable healthcare.* Journal of Biomed Research, 3(1), 42–47.
-
- ▶ **Lekadir, K., Giancula, Q., Anna, T. G., & Gallin, C. (2022).** *Artificial intelligence in healthcare: Applications, risks, and ethical and societal impacts.* European Parliament. Directorate General for Parliamentary Research Services. <https://data.europa.eu/doi/10.2861/568473>
-
- ▶ **Liang, X., Yu, X., & Gao, T. (2022).** *Machine learning with magnetic resonance imaging for prediction of response to neoadjuvant chemotherapy in breast cancer: A systematic review and meta-analysis.* European Journal of Radiology, 150, 110247. <https://doi.org/10.1016/j.ejrad.2022.110247>
-
- ▶ **Lokaj, B., Pugliese, M.-T., Kinkel, K., Lovis, C., & Schmid, J. (2023).** *Barriers and facilitators of artificial intelligence conception and implementation for breast imaging diagnosis in clinical practice: A scoping review.* European Radiology. <https://doi.org/10.1007/s00330-023-10181-6>
-
- ▶ **Malherbe, K. (2021).** *Tumor Microenvironment and the Role of Artificial Intelligence in Breast Cancer Detection and Prognosis.* The American Journal of Pathology, 191(8), 1364–1373. <https://doi.org/10.1016/j.ajpath.2021.01.014>
-
- ▶ **Ozdemir, S., & Finkelstein, E. (2018).** *Cognitive Bias: The Downside of Shared Decision Making.* JCO Clinical Cancer Informatics, 2. <https://doi.org/10.1200/CCI.18.00010>
-
- ▶ **Pesapane, F., Trentin, C., Montesano, M., Ferrari, F., Nicosia, L., Rotili, A., Penco, S., Farina, M., Marinucci, I., Abbate, F., Meneghetti, L., Bozzini, A., Latronico, A., Liguori, A., Carrafiello, G., & Cassano, E. (2022).** *Mammography in 2022, from Computer-Aided Detection to Artificial Intelligence Applications.* Clinical and Experimental Obstetrics & Gynecology, 49(11), 237. <https://doi.org/10.31083/j.ceog4911237>
-
- ▶ **Radley-Gardner, O., Beale, H., & Zimmermann, R. (Eds.). (2016).** *Fundamental Texts On European Private Law.* Hart Publishing. <https://doi.org/10.5040/9781782258674>
-
- ▶ **Reddy, S., Allan, S., Coghlan, S., & Cooper, P. (2020).** *A governance model for the application of AI in health care.* Journal of the American Medical Informatics Association, 27(3), 491–497. <https://doi.org/10.1093/jamia/ocz192>
-

Ricci Lara, M. A., Echeveste, R., & Ferrante, E. (2022). *Addressing fairness in artificial intelligence for medical imaging*. *Nature Communications*, 13(1), 4581. <https://doi.org/10.1038/s41467-022-32186-3>

Schneberger, D., Stöger, K., & Holzinger, A. (2020). *The European Legal Framework for Medical AI*. In A. Holzinger, P. Kieseberg, A. M. Tjoa, & E. Weippl (Eds.), *Machine Learning and Knowledge Extraction* (Vol. 12279, pp. 209–226). Springer International Publishing. https://doi.org/10.1007/978-3-030-57321-8_12

Tsopra, R., Fernandez, X., Luchinat, C., Alberghina, L., Lehrach, H., Vanoni, M., Dreher, F., Sezerman, O. U., Cuggia, M., De Tairac, M., Miklasevics, E., Itu, L. M., Geanta, M., Ogilvie, L., Godey, F., Boldisor, C. N., Campillo-Gimenez, B., Cioroboiu, C., Ciusdel, C. F., ... Burgun, A. (2021). *A framework for validating AI in precision medicine: Considerations from the European ITFoC consortium*. *BMC Medical Informatics and Decision Making*, 21(1), 274. <https://doi.org/10.1186/s12911-021-01634-3>

Vedantham, S., Shazeeb, M. S., Chiang, A., & Vijayaraghavan, G. R. (2023). *Artificial Intelligence in Breast X-Ray Imaging*. *Seminars in Ultrasound, CT, and MR*, 44(1), 2–7. <https://doi.org/10.1053/j.sult.2022.12.002>



06/

Recursos





LH n.339

Bibliografía elaborada por:**Alicia Martín.**

Biblioteca. Campus Docent Sant Joan de Déu.
Sant Boi de Llobregat (Barcelona).

Sonia Marín,

Biblioteca.
Escuela de Enfermería y Fisioterapia
San Juan de Dios.
Universidad Pontificia Comillas.
Ciempozuelos (Madrid).

Fuentes: Universidad de Barcelona,
Universidad Pontificia Comillas, Pubmed,
Dialnet, Academic search Complete,
Medline, Cinahl Complete.

Adams, J. (2023).

Defending explicability as a principle for the ethics of artificial intelligence in medicine. *Medicine, Health Care, and Philosophy*, 26(4), 615–623.
<https://doi.org/10.1007/s11019-023-10175-7>

Adamson, D., Dyke, G., Jang, H., & Rosé, C. P. (2014).

Towards an agile approach to adapting dynamic collaboration support to student needs. *International Journal of Artificial Intelligence in Education*, 24(1), 92-124.

Ailin, I. G. (2024).

Inteligencia artificial y autonomía: Reflexiones sobre dilemas (bio)éticos a partir de un caso de ficción... 'la naranja mecánica'.
Universidad Pontificia Comillas.
<https://doi.org/10.14422/rib.i25.y2024.009>

Albarrán Torres, E.

Hacia una educación personalizada y adaptativa: la disrupción de la inteligencia artificial.
(2023). Centro Internacional de Educación Continua - Universidad Pedagógica Experimental Libertador.

Alonso Salgado, C., Valiño Ces, A., & Rodríguez Álvarez, A. (2024).

Derecho, Nuevas Tecnologías e Inteligencia Artificial.
Dykinson, S.L.

Alowais, S. A., Alghamdi, S. S., Alsuhebany, N., Alqahtani, T., Alshaya, A. I., Almohareb, S. N., Aldairem, A., Alrashed, M., Bin Saleh, K., Badreldin, H. A., Al Yami, M. S., Al Harbi, S., & Albekairy, A. M. (2023).

Revolutionizing healthcare: the role of artificial intelligence in clinical practice. *BMC medical education*, 23(1), 689.
<https://doi.org/10.1186/s12909-023-04698-z>

Amos Usanos, R. (2023).

Inteligencia artificial y bioética.
Universidad Pontificia Comillas.

Angulo Bahon, C. (2022).

Artificial inteligencia-artificial.
Claves de Razón práctica, (280), 48-55.

▶ **Arellano Toledo, W., & Barrio Andrés, M. (2023).**
Derecho, ética e inteligencia artificial.
Tirant lo Blanch.

▶ **Arroyo Sagasta, A. (2024).**
Inteligencia artificial y educación: construyendo puentes. Editorial Graó.

▶ **Balagué Serre, N., & Almarcha, M. (2023).**
Inteligencia biológica en acción: cómo nos adaptamos. Independently Published.

▶ **Barnhart, A. J., Barnhart, J. E. M., & Dierickx, K. (2023).**
Why ChatGPT means communication ethics problems for bioethics.
American Journal of Bioethics, 23(10), 80–82.
<https://doi.org/10.1080/15265161.2023.2250278>

▶ **Beltrán de Heredia Ruiz, I. (2023).**
Inteligencia artificial y neuroderechos: la protección del yo inconsciente de la persona.
Editorial Aranzadi, S.A.U.

▶ **Biller-Andorno, N., Ferrario, A., & Gloeckler, S. (2022).**
In search of a mission: Artificial intelligence in clinical ethics.
American Journal of Bioethics, 22(7), 23–25.
<https://doi.org/10.1080/15265161.2022.2075055>

▶ **Biscaia Fernández, J. M., González-Soltero, M. R., Biscaia Fernández, C. J., Mohedano del Pozo, R. B., & Rodríguez-Learte, A. I. (2024).**
Empleo de ChatGPT en educación biomédica. análisis de riesgos desde los principios éticos

de la UNESCO y el reglamento de la unión europea sobre inteligencia artificial.
Revista iberoamericana de bioética.
<https://doi.org/10.14422/rib.i25.y2024.007>

▶ **Boden, M. A. (2022).**
Inteligencia artificial. Turner Publicaciones.

▶ **Bracero, F. (2023).**
Bicicletas para la mente: el viaje desde el primer PC hasta la inteligencia artificial.
Península.

▶ **Brighton, H. (2019).**
Inteligencia artificial: Una guía ilustrada.
Tecnos.

▶ **Bunar, J. (2023).**
Inteligencia artificial: Sus beneficios y peligros.
Yachay, 40(78), 227–260.
<https://doi.org/10.35319/yachay.202378109>

▶ **Cantú-Martínez, P. C. (2024).**
Inteligencia artificial y sus connotaciones éticas.
Revista iberoamericana de bioética, (25), 1-11.
<https://doi.org/10.14422/rib.i25.y2024.008>

▶ **Car, J., Sheikh, A., Wicks, P., & Williams, M. S. (2019).**
Beyond the hype of big data and artificial intelligence: Building foundations for knowledge and wisdom.
BMC Medicine, 17(1), 143.
<https://doi.org/10.1186/s12916-019-1382-x>

▶ **Chen, Y., Jensen, S., Albert, L. J., Gupta, S., & Lee, T. (2023).**
Artificial intelligence (AI) student assistants in the classroom: Designing chatbots to support student success.

LH n.339

Information Systems Frontiers, 25(1), 161-182.
<https://doi.org/10.1007/s10796-022-10291-4>

▶ **Cobo Cano, M., & Iglesias, L. L. (2023).**
Inteligencia artificial y medicina.
Editorial Consejo Superior
de Investigaciones Científicas.

▶ **Concha, F. (2024).**
*Entrenamiento en formulación de casos en
psicoterapia a través de inteligencia artificial
generativa.*
Revista de Psicoterapia, 35(127), 97–112.
<https://doi.org/10.5944/rdp.v35i127.37798>

▶ **Corona Arias, J. C. (2024).**
*Formalización y respetividad: Reflexión sobre
el concepto de inteligencia artificial desde el con-
cepto de inteligencia sentiente de Xavier Zubiri.*
Xipe Totek, 32(120), 82–107.
[https://dialnet.unirioja.es/servlet/
articulo?codigo=9318127](https://dialnet.unirioja.es/servlet/articulo?codigo=9318127)

▶ **Curchoe, C. L., Letterie, G. S.,
& Quaas, A. M. (2023).**
*Unlocking the potential of artificial intelligence
(AI) in reproductive medicine: the JARG
collection on assisted reproductive technology
(ART) and machine learning.*
Journal of assisted reproduction
and genetics, 40(9), 2079-2080.
<https://doi.org/10.1007/s10815-023-02893-x>

▶ **Degli-Esposti, S. (2023).**
Ética de la inteligencia artificial.
Consejo Superior de Investigaciones Científicas.

▶ **Destéfano, M., Trifonova, A.,
& Barajas, M. (2024).**
*Ensenyar IA a la propera generació:
un enfocament humanista.*
Digital Education Review, 45, 115-123.

▶ **Echarte Alonso, L. E. (2024).**
*Inteligencia artificial emocional
en el reverso del test de turing. al borde
de la singularidad tecnológica son precisas
cuatro nuevas leyes para la robótica.*
Revista Iberoamericana de Bioética, (25),1-22.
<https://doi.org/10.14422/rib.i25.y2024.003>

▶ **Esquerda Bosch, M.,
Esquerda Pifarré, F., & Pifarré, J. (2020).**
*Deep learning. deep ethics: Ética para el uso
de la inteligencia artificial en medicina.*
Labor Hospitalaria: Organización
y Pastoral de la Salud, 52(327), 53–59.
[https://www.laborhospitalaria.com/
slug/327-04-deep-learning-deep-ethics-
etica-para-el-uso-de-la-inteligencia-artificial-
en-medicina/](https://www.laborhospitalaria.com/slug/327-04-deep-learning-deep-ethics-etica-para-el-uso-de-la-inteligencia-artificial-en-medicina/)

▶ **Fernández, J. M. B., González-Soltero, M. d.
R., Fernández, C. J. B., Pozo, R. B. M. d.,
& Rodríguez-Learte, A. I. (2024).**
*Empleo de ChatGPT en educación biomédica.
análisis de riesgos desde los principios éticos
de la UNESCO y el reglamento de la unión
europea sobre inteligencia artificial.*
Revista Iberoamericana de Bioética, (25), 1.
<https://doi.org/10.14422/rib.i25.y2024.007>

▶ **Fernández Marcial, V;
Esteves Gomes, L.I. (2022).**
*Impacto de la Inteligencia Artificial
en el comportamiento informacional:
elementos para el debate.*
Bibliotecas. Anales de Investigacion;18(3), 1-12.
[https://dialnet.unirioja.es/servlet/
articulo?codigo=8741939](https://dialnet.unirioja.es/servlet/articulo?codigo=8741939)

- Galván Fernández, C. (2024). *La inteligencia artificial desde la educabilidad*. Digital Education Review, 45,1-2.
-
- Gamboa, J. A., Nolasco, J. S., Nolasco, L. E., & Casas Miranda, R. (2024). *ChatGPT: obtén el máximo rendimiento a la inteligencia artificial generativa*. Ra-Ma.
-
- Gatti, A. (2024). *Alfabetització i intel·ligència artificial*. JONED. Journal of Neuroeducation, 5(1), 52-58. doi: 10.1344/joned.v5i1.46108
-
- Gawdat, M., & Guelbenzu, A. (2024). *La Inteligencia que asusta: el futuro de la inteligencia artificial y cómo podemos salvar nuestro mundo*. Paidós.
-
- Gesto Rodríguez, J. (2023). *Inteligencia artificial*. Lekla Ediciones.
-
- Gómez Ligierre, C. I., & García Micó, T. G. (2024). *Liability for Artificial Intelligence and other emerging technologies*. InDret. Revista para el Análisis del Derecho, 2, 501-511
-
- Gutiérrez González, M. A. & Ocadiz Amador, A. (2024). *Los problemas tecno-éticos en la inteligencia artificial generativa*. Razón y Palabra, 28(119), 15–27. <https://doi.org/10.26807/rp.v28i119.2106>
-
- Haque N. (2023). *Reply to: Artificial intelligence in geriatric medicine: Potentials and pitfalls*. Journal of the American Geriatrics Society, 71(11), 3652-3653. <https://doi.org/10.1111/jgs.18567>
-
- Howard, C., Jordan, P., Eugenio, B., & Katz, S. (2017). *Shifting the load: A peer dialogue agent that encourages its human collaborator to contribute more to problem solving*. International Journal of Artificial Intelligence in Education, 27(1), 101-129. <https://doi.org/10.1007/s40593-015-0071-y>
-
- Hughes, J., Daffas, P., & Robertson, S. (2022). *A healthcare ethics approach in identifying patient care issues using detailed case analysis: The importance of omissions of fact*. Bioethics, 36(6), 724–727. <https://doi.org/10.1111/bioe.12997>
-
- Hunter, D. J., & Holmes, C. (2023). *Where Medical Statistics Meets Artificial Intelligence*. The New England journal of medicine, 389(13), 1211-1219. <https://doi.org/10.1056/NEJMra2212850>
-
- Inglada Galiana, L., Corral Gudino, L., & Miramontes González, P. (2024). *Ethics and artificial intelligence*. Revista clinica española, 224(3), 178–186. <https://doi.org/10.1016/j.rceng.2024.02.00>
-
- Johnson, S. L. J. (2019). *AI, machine learning, and ethics in health care*. The Journal of Legal Medicine, 39(4), 427–441. <https://doi.org/10.1080/01947648.2019.1690604>
-

LH n.339

Junior, S. A. N. M., Aguiar, C. L., Cunha, L. A. S & Brustolin, J. C. R. (2024).

A interação homem-máquina na psicoterapia: Uma revisão sistemática sobre o uso de inteligências artificiais no contexto da saúde mental. Prometeica. Revista de Filosofia y Ciencias. 29. 335-347.

<https://doi.org/10.34024/prometeica.2024.29.16268>.

Kaplan, A., & Haenlein, M. (2019).

Siri, Siri, in my hand: Who's the fairest in the land? Un the interpretations, illustrations, and implications of artificial intelligence.

Business Horizons, 62(1), 15-25.

[-https://doi.org/10.1016/j.bushor.2018.08.004](https://doi.org/10.1016/j.bushor.2018.08.004).

Kar, S., Roy, C., Das, M.,

Mullick, S., y Saha, R. (2023).

AI Horizons: Unveiling the Future of Generative Intelligence.

International Journal of Advanced Research in Science, Communication and Technology, 387-391.

<https://doi.org/10.48175/ijarsct-12969>

Komesaroff, P. A., & Felman, E. R. (2023).

How to make sense of the ethical issues raised by artificial intelligence in medicine.

Internal medicine journal, 53(8), 1304–1305.

<https://doi.org/10.1111/imj.16180>

Liberos, E., Ahumada Luyando, S., & Sánchez Ahumada, M. (2024).

Inteligencia artificial para el marketing: cómo la tecnología revolucionará tu estrategia. ESIC Editorial.

Lindín, C. (2024).

Inteligència artificial: Tensions educatives entre la veritat i la versemblança de les dades. Digital Education Review, 45, 20-28.

Liu K.

Artificial Intelligence and Ethical Frameworks in Pediatrics.

JAMA Pediatrics. 2024;178(6):626–627.

<https://doi.org/10.1001/jamapediatrics.2024.0510>

Liu, W. (2023).

Prospects of applying artificial intelligence to determine students' mental health status in school education.

Revista de Psicología del Deporte, 32(2), 76–84.

<https://www.rpd-online.com/index.php/rpd/article/view/1065>

Llorca Albareda, J. 2023.

El estatus moral de las entidades de inteligencia artificial.

Disputatio, 12(24), 241-249.

<https://doi.org/10.5281/zenodo.8140967>.

Lodhi, P., Mishra, O., Jain, S., & Bajaj, V. (2018).

StuA: An intelligent student assistant.

International Journal of Interactive Multimedia and Artificial Intelligence, 5(2), 17-25.

<https://doi.org/10.9781/ijimai.2018.02.008>.

Love, C. S. (2023).

"Just the facts ma'am": Moral and ethical considerations for artificial intelligence in medicine and its potential to impact patient autonomy and hope.

Linacre Quarterly, 90(4), 375–394.

<https://doi.org/10.1177/00243639231162431>

Luo, Y., & Hu, N. (2023).

Can ChatGPT be a new educational tool in medicine?

Medicina clinica, 161(8), 363-364.

<https://doi.org/10.1016/j.medcli.2023.05.018>

▶ **Martín Ríos, M. del P., Villegas Delgado, C., Domínguez Barragán, M. L., & Rodríguez Lainz, J. L. (2023).** *La Tecnología y la inteligencia artificial al servicio del proceso.* Editorial Colex, S.L.

▶ **McLennan, S., Fiske, A., Tigard, D., Müller, R., Haddadin, S., & Buyx, A. (2022).** *Embedded ethics: A proposal for integrating ethics into the development of medical AI.* BMC Medical Ethics, 2310. <https://doi.org/1186/s12910-022-00746-3>

▶ **Mejía Trejo, J. (2024).** *Inteligencia Artificial. Fundamentos de Ingeniería de Prompts con ChatGPT como Innovación impulsora de la Creatividad.* Academia Mexicana de Investigación y Docencia en Innovación AMIDI.

▶ **Méndez-Martínez, J. A., & José Ángel Méndez-Martínez. (2023).** *Investigación, Innovación y Transferencia Del Conocimiento.* Dykinson, S.L.

▶ **Mitchell, M., & Rodríguez Tapia, M. L. (2024).** *Inteligencia artificial: guía para seres pensantes.* Capitán Swing.

▶ **Nurock, V. (2020).** *¿Puede prestar cuidados la inteligencia artificial?* Cuadernos de Relaciones Laborales, 38(2). <https://doi.org/10.5209/crla.70880>

▶ **Ocampo Ponce, M. (2024).** *Algunos fundamentos de la filosofía realista de santo Tomás de Aquino para una valoración de la inteligencia artificial (IA).* Revista Perseitas, 12, 72–92. <https://doi.org/10.21501/23461780.4724>

▶ **Paladino, M. S. (2023).** *Inteligencia artificial en medicina: Reflexiones éticas desde el pensamiento de Edmund Pellegrino.* Cuadernos de Bioética. 34(110), 25-35. <https://doi.org/10.30444/CB.140>

▶ **Palmer, A., & Schwan, D. (2022).** *Beneficent dehumanization: Employing artificial intelligence and carebots to mitigate shame-induced barriers to medical care.* Bioethics, 36(2), 187–193. <https://doi.org/10.1111/bioe.12986>

▶ **Peñas García, L. (2024).** *La ética en la inteligencia artificial generativa.* Labor Hospitalaria: Organización y Pastoral de la Salud, 55(338), 75–81. <https://www.laborhospitalaria.com/la-etica-en-la-inteligencia-artificial-generativa/>

▶ **Pérez-Ugena, M. (2024).** *La inteligencia artificial: Definición, regulación y riesgos para los derechos fundamentales.* Estudios de Deusto: Revista de Derecho Público, 72(1), 307-337. <https://doi.org/10.18543/ed.3108>

▶ **Puertas Prats, E., & Cano García, M. E. (2024).** *¿Puede la Inteligencia Artificial proporcionar un feedback más sostenible?* Digital Education Review, 45, 50-58.

LH n.339

▶ **Rahimzadeh, V. (2024).**

Personal but necessarily predictive? developing a bioethics research agenda for AI-enabled decision-making tools.
The American Journal of Bioethics, 24(7), 29–31.

<https://doi.org/10.1080/15265161.2024.2353031>

▶ **Ray, P. P. (2023).**

Generative artificial intelligence (AI) and medical ethics: A symbiotic dance for the future.

Journal of Oral and Maxillofacial Surgery, 81(12), 1457–1459.

<https://doi.org/10.1016/j.joms.2023.09.015>

▶ **Rebollo Delgado, L. (2023).**

Inteligencia artificial y derechos fundamentales.
Editorial DykinsonRíos,

▶ **A. J. (2024).**

Más allá de las visiones Apocalípticas: El debate sobre la inteligencia artificial.
Debates IESA, 28(2), 16–18.

▶ **Rocha Espíndola, M. A., Sansó-Rubert Pascual, D., Rodríguez Dos Santos, N. (2023).**

Inteligencia artificial y derecho: Reflexiones jurídicas para el debate sobre su desarrollo y aplicación.
Dykinson.

▶ **Rocha, M. L. & Pereira, R. S. (2020).**

Inteligência artificial & direito.
Almedina.

▶ **Roco-Videla, Á, Caviedes-Olmos, M., Aguilera-Eguía, R. A., & Olguín-Barraza, M. (2024).**

La inteligencia artificial y su implicación

en el desarrollo de revisiones sistematizadas. Usos y limitaciones.

Nutricion Hospitalaria, 41(1), 264–265.
<https://doi.org/10.20960/nh.04889>

▶ **Santamaría Vázquez, M. (2024).**

Formación y especialización en atención temprana: uso de recursos tecnológicos y de inteligencia artificial.

Editorial Universidad de Burgos.

▶ **Sauerbrei, A., Hallowell, N., & Kerasidou, A. (2022).**

Algorithmic ethics: A technically sweet solution to a non-problem.

American Journal of Bioethics, 22(7), 28–30.

<https://doi.org/10.1080/15265161.2022.2075050>

▶ **Savulescu, J., Giubilini, A., Vandersluis, R., & Mishra, A. (2024).**

Ethics of artificial intelligence in medicine.
Singapore medical journal, 65(3), 150–158.
<https://doi.org/10.4103/singaporemedj.SMJ-2023-279>

▶ **Serafim, M. C., Bertoncini, A. L. C., Ames, M. C., & Pansera, D. (2024).**

Inteligencia artificial (de)generativa: Sobre la imposibilidad de que un sistema de IA tenga una experiencia moral.
Scripta Theologica, 56(2), 467–502.
<https://doi.org/10.15581/006.56.2.467-502>

▶ **Seungmi, P., & Insun, J. (2023).**

The impact of nursing students' biomedical and artificial intelligence ethical awareness, ethical values, and professional self-concept on their ethical decision-making confidence.

Journal of Korean Academic Society of Nursing Education, 29(4), 371–380.

<https://doi.org/10.5977/jkas-ne.2023.29.4.371>

Shaw, J., Ali, J., Atuire, C. A., Cheah, P. Y., Español, A. G., Gichoya, J. W., Hunt, A., Jjingo, D., Littler, K., Paolotti, D., & Vayena, E. (2024). *Research ethics and artificial intelligence for global health: Perspectives from the global forum on bioethics in research*. BMC Medical Ethics, 25(1), 46. <https://doi.org/10.1186/s12910-024-01044-w>

Silva-Peñafiel, G. E., Castillo-Parra, B. F., Tixi-Gallegos, K. G., & Urgiles-Rodríguez, B. E. *Revolución de la Inteligencia Artificial en la Educación Superior*. (2024). Editorial Grupo AEA.

Sisk, B. A., Antes, A. L., & DuBois, J. M. (2024). *An overarching framework for the ethics of artificial intelligence in pediatrics*. JAMA Pediatrics, 178(3), 213–214. <https://doi.org/10.1001/jamapediatrics.2023.5761>

Smith, H., Birchley, G., & Ives, J. (2024). *Artificial intelligence in clinical decision-making: Rethinking personal moral responsibility*. Bioethics, 38(1), 78–86. <https://doi.org/10.1111/bioe.13222>

Spector-Bagdady, K., Rahimzadeh, V., Jaffe, K., & Moreno, J. (2022). *Promoting ethical deployment of artificial intelligence and machine learning in healthcare*. American Journal of Bioethics, 22(5), 4–7. <https://doi.org/10.1080/15265161.2022.2059206>

Sung J. J. Y. (2024). *Introduction to artificial intelligence in medicine*. Singapore medical journal, 65(3), 132.

<https://doi.org/10.4103/singaporemedj.SMJ-2024-060>

Swiecki, Z., Khosravi, H., Chen, G., Martinez-Maldonado, R., Lodge, J. M., Milligan, S., Selwyn, N., & Gašević, D. (2022). *Assessment in the age of artificial intelligence*. Computers and Education: Artificial Intelligence, 3, 100075.

Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). *Developing evaluative judgement: Enabling students to make decisions about the quality of work*. Higher Education, 76(3), 467–481.

Tal, A., Elyoseph, Z., Haber, Y., Angert, T., Gur, T., Simon, T., & Asman, O. (2023). *The artificial third: Utilizing ChatGPT in mental health*. American Journal of Bioethics, 23(10), 74–77. <https://doi.org/10.1080/15265161.2023.2250297>

Tegmark, M. (2018). *Life 3.0: Being human in the age of artificial intelligence*. Penguin Books.

Tejedor, S., Pulido, C., Hita, E., Robledo, K., & Salaverría, R. (2023). *La Inteligencia artificial en el periodismo: mapping de conceptos, casos y recomendaciones*. Editorial UOC.

Ufarte-Ruiz, M., Borden, S., & Codina, L. (2024). *Introduction. The revolution driven by artificial intelligence continues: A journey through the media landscape*.

LH n.339

Communication & Society, 37(3), 197–199.
<https://doi.org/10.15581/003.37.3.197-199>

▶ **Urgelés Puértolas, D. (2023).**
Implicaciones a corto plazo de la inteligencia artificial en la atención sanitaria.
Labor Hospitalaria: Organización y Pastoral de la Salud, 55(336-337), 37–43.

▶ **Vallespín Pérez, D. (2023).**
Inteligencia artificial y proceso: Eficiencia vs. Garantías.
Juruá Editorial.

▶ **Vásquez, B., Moreno-Lacalle, R., Soriano, G. P., Juntasoopeepun, P., Locsin, R. C., & Evangelista, L. S. (2023).**
Technological machines and artificial intelligence in nursing practice.
Nursing & health sciences, 25(3), 474-481.
<https://doi.org/10.1111/nhs.13029>

▶ **Villas Olmeda, M., & Camacho Ibáñez, J. (2022).**
Manual de ética aplicada en inteligencia artificial.
Anaya Multimedia.

▶ **Viniegra-Velázquez L. (2024).**
El progreso en medicina y la inteligencia artificial.
Boletín médico del Hospital Infantil de México, 81(3), 121-131.
<https://doi.org/10.24875/BMHIM.24000007>

▶ **Vitola-Quintero, M., Ballestas-Campo, N., Pérez-Cerro, J., & Forbes-Santiago, R. (2024).**
Implicaciones Éticas, Sociales y Ambientales de la Inteligencia Artificial para el Desarrollo Sostenible: Una Revisión de la Literatura.
Revista Científica Anfibios, 7(1), 72-81.
<https://doi.org/10.37979/afb.2024v7n1.148>

▶ **Yu, K. H., Healey, E., Leong, T. Y., Kohane, I. S., & Manrai, A. K. (2024).**
Medical Artificial Intelligence and Human Values.
The New England journal of medicine, 390(20), 1895-1904.
<https://doi.org/10.1056/NEJMra2214183>

LH

HUMANIZACIÓN, PASTORAL Y ÉTICA DE LA SALUD Y SOCIAL
www.laborhospitalaria.com

