

# Inteligencia artificial: un puente hacia el futuro de la medicina

Rui Nunes<sup>1</sup>, Sofia B Nunes<sup>1</sup>

1. Faculdade de Medicina da Universidade do Porto, Porto/Portugal.

## Resumen

La inteligencia artificial se ha convertido en una herramienta esencial en diversos campos de la medicina, posibilitando nuevos patrones de innovación, como medicina de precisión, neurociencias, genómica y proteómica. Este artículo busca sintetizar los principales desafíos éticos surgidos de la aplicación de la inteligencia artificial en la medicina, con el objetivo de promover una implementación responsable de estas tecnologías digitales, siempre al servicio de las personas y la humanidad. Además, se proponen un conjunto de recomendaciones para el desarrollo de una inteligencia artificial confiable en el contexto médico. Metodológicamente, se realizó un análisis exhaustivo de los principales instrumentos éticos y jurídicos internacionales orientados a la regulación de la inteligencia artificial. Se concluye que una gobernanza ética eficaz es fundamental para preservar la agencia humana en la aplicación de los sistemas de inteligencia artificial y garantizar total transparencia en los procesos de investigación y desarrollo.

**Palabras clave:** Inteligencia artificial. Ética médica. Regulación de la inteligencia artificial. Transparencia en la investigación. Bioética. Tecnología y salud.

## Resumo

### Inteligência artificial: uma ponte para o futuro da medicina

A inteligência artificial tornou-se ferramenta essencial em diversos domínios da medicina, possibilitando novos padrões de inovação, como medicina de precisão, neurociências, genômica e proteômica, entre outros. Este artigo busca sintetizar os principais desafios éticos suscitados pela aplicação da inteligência artificial na medicina, a fim de promover uma implementação responsável dessas novas tecnologias digitais, sempre a serviço das pessoas e da humanidade. Além disso, propõe-se um conjunto de recomendações para o desenvolvimento de uma inteligência artificial confiável no contexto médico. Do ponto de vista metodológico, realizou-se análise aprofundada dos principais instrumentos éticos e jurídicos internacionais voltados à regulação da inteligência artificial. Conclui-se que a exigência de uma governança ética eficaz torna imprescindível garantir a preservação da agência humana na aplicação dos sistemas de inteligência artificial, bem como assegurar total transparência nos processos de pesquisa e desenvolvimento dessas tecnologias.

**Palavras-chave:** Inteligência artificial. Ética médica. Regulação da inteligência artificial. Transparência na pesquisa. Bioética. Tecnologia e saúde.

## Abstract

### Artificial intelligence: a bridge to the future of medicine

Artificial intelligence has become an essential tool across various fields of medicine, enabling new patterns of innovation such as precision medicine, neurosciences, genomics, and proteomics. This article aims to synthesize the main ethical challenges arising from the application of artificial intelligence in medicine, in order to promote responsible implementation of these digital technologies, always serving people and humanity. Additionally, it proposes a set of recommendations for developing trustworthy artificial intelligence in the medical context. Methodologically, an in-depth analysis was conducted of the main international ethical and legal instruments regulating artificial intelligence. It is concluded that effective ethical governance is essential to preserve human agency in the application of artificial intelligence systems and to ensure full transparency in research and development processes.

**Keywords:** Artificial intelligence. Medical ethics. Regulation of artificial intelligence. Transparency in research. Bioethics. Technology and health.

Los autores declaran que no existe ningún conflicto de interés.

La inteligencia artificial (IA) está cambiando drásticamente la vida de las personas en todo el planeta, influyendo en aspectos fundamentales de nuestra vida social en ámbitos como la economía, el sistema financiero, las artes creativas, la educación e incluso la salud pública o la prestación de cuidados de salud. También la investigación científica y el desarrollo tecnológico ya están hoy controlados en gran medida por la IA, lo que permite nuevos estándares de innovación, como la proteómica, lo que implica serios desafíos éticos<sup>1</sup>.

Numerosas instituciones internacionales — como la Organización Mundial de la Salud (OMS)<sup>2</sup>, la Comisión Europea<sup>3</sup>, el Parlamento Europeo<sup>4</sup>, la Office of Science and Technology Policy (Estados Unidos)<sup>5</sup>, el Nuffield Council on Bioethics<sup>6</sup> y el Consejo Nacional de Ética de las Ciencias de la Vida de Portugal<sup>7</sup>— expresó importantes reservas sobre la aplicación de la IA, sugiriendo la necesidad de una supervisión efectiva en su desarrollo, al igual que control humano en su aplicación. También el Consejo Federal de Medicina (CFM) ha mostrado especial preocupación por este tema y ha creado un grupo de trabajo para tal efecto.

Este artículo pretende resumir los principales desafíos éticos que plantea la IA en medicina para posibilitar la implementación responsable de nuevas tecnologías digitales que estén siempre al servicio de la humanidad. También busca promover un conjunto de recomendaciones para una IA confiable en medicina. Para ello, y desde un punto de vista metodológico, los autores realizaron un análisis detallado de los principales instrumentos éticos y legales internacionales sobre regulación de la IA.

### Inteligencia artificial en la investigación y en la medicina

El desarrollo de la inteligencia artificial —definida como un sistema diseñado por un ser humano que, dado un objetivo complejo, actúa en las dimensiones física y digital, percibiendo su entorno mediante la adquisición de datos, interpretando los datos obtenidos de forma estructurada o no, razonando desde el conocimiento o procesando la información derivada de estos datos y decidiendo el mejor curso de acción para lograr este objetivo<sup>8</sup>— es parte de una amplia transición digital que está afectando a nuestra sociedad. Este fenómeno se

caracteriza por una rápida expansión de diferentes tecnologías de información y comunicación (TIC), un desarrollo sin precedentes de la computación cuántica, la implementación de la robótica en los más diversos dominios de actividad o la generación de metadatos, *big data* y *lake data*, sin control y supervisión efectivos.

Los sistemas de IA tienen características distintivas que implican un cambio de paradigma, no se trata solo de una nueva herramienta digital. De hecho, la IA tiene la capacidad de aprendizaje y de autoaprendizaje, de forma automática (*machine learning*), profunda (*deep learning*) o mediante la asociación entre diferentes conceptos. Otra característica distintiva de la IA es el *networking*, es decir, la capacidad de trabajar en red interactiva: la actividad de los robots autónomos, por ejemplo, es independiente, pero también está interconectada. El concepto mismo de “robot social” se refiere no solo a la interacción con humanos, sino también a la conectividad entre diferentes sistemas de IA.

Por otra parte, los robots autónomos se mueven en el espacio físico humano. De esta manera, la IA no se limita espacialmente a una computadora, sino que puede moverse e incluso contactar con cualquier persona mediante el uso competente del lenguaje humano. Este es un enorme desafío para la humanidad, dado que la capacidad de dominar el lenguaje puede dar lugar a la obtención de aspectos esenciales de la cultura y de la civilización humanas, incluso antes de alcanzar la singularidad<sup>9</sup>. La singularidad se entiende como una etapa en la que la inteligencia artificial, intencionalmente o no, adquiere la capacidad de controlar el desarrollo colectivo de la humanidad sin que nuestra especie logre comprender los mecanismos de actividad y funcionamiento de los sistemas de IA<sup>10</sup>. Por lo tanto, es esencial comprender cómo evoluciona la IA<sup>11</sup>.

Más especulativa es la posibilidad de que la IA desarrolle la sintiencia<sup>12</sup> (la capacidad de tener experiencias subjetivas) o incluso desarrolle una personalidad cibernética<sup>13</sup>. Basada en *software* que actúa en el mundo virtual (asistentes de voz, *software* para análisis de imágenes, motores de búsqueda, sistemas de reconocimiento facial, etc.), o incorporada en dispositivos de *hardware* (robots sociales, vehículos autónomos, drones, etc.), la IA parece no tener límites tecnológicos para su desarrollo. Es decir, tiene el potencial de revolucionar la

forma en que enseñamos, aprendemos, investigamos y aplicamos el conocimiento científico.

En la docencia y en la investigación, la IA generativa —que tiene ChatGPT, DeepSeek, Claude, Gemini y DALL-E como buenos ejemplos— ha dado lugar a una innovación científica y pedagógica disruptiva debido a la agilidad con la que se produce y la velocidad con la que se implementa. Si se utiliza correctamente, con rectitud y profesionalismo, puede animar a los estudiantes a adquirir habilidades técnicas con mayor facilidad <sup>14</sup>.

La IA generativa permite la implementación de un modelo de *active learning* más personalizado y atractivo en todos los dominios científicos, contribuyendo rápidamente al desarrollo de contenidos pedagógicos útiles para apoyar a los estudiantes (resúmenes, *flashcards*, preguntas sobre contenidos, etc.) y a los docentes (preparación de materiales para uso en planes de clase, informes, diapositivas, preguntas para discusión y preguntas de examen). También puede ser una ayuda importante en la enunciación de preguntas científicas que incentiven la reflexión y la investigación sobre un tema determinado, o incluso posibilitar la utilización de material bibliográfico proporcionado por los docentes en lugar de la simple consulta en internet. Es una herramienta excelente para la traducción de textos y la redacción de trabajos, además permite desarrollar nuevas aplicaciones, con un potencial inagotable <sup>15</sup>.

Sin embargo, la IA generativa plantea serias reservas que deben ser abordadas, especialmente en el ámbito de la rectitud científica y académica. Dado que recurre a bases de datos primarias que fluyen libremente en internet (algunos datos pueden exponenciarse de forma segmentada y, por lo tanto, los resultados pueden estar sesgados), es posible cuestionar si es plagio utilizar un texto generativo previamente publicado que recurre a la creación intelectual de terceros. Se trata de una circunstancia que suscita naturales reservas éticas y tiene implicaciones en el usufructo de los derechos de autor o en la generación de patentes comerciales, especialmente en medicina. En consecuencia, se necesita un nuevo marco legal y regulatorio para encuadrar el uso de la IA generativa en el ámbito académico, en los centros de investigación y en la traducción para las empresas, teniendo en cuenta los aspectos éticos siguientes:

1. Protección de datos y privacidad individual.
2. Derechos de autor
3. Prevención del plagio.
4. Obtención de referencias fidedignas.
5. Garantía de verificabilidad, fiabilidad y veracidad de la información.
6. Lucha contra la desinformación.
7. Control del flujo de información.
8. Garantía de usuarios incluidos digitalmente.

Por otra parte, la creatividad científica, fundamental para una evolución consistente de la medicina, puede verse limitada con el uso sistemático de la inteligencia artificial, porque el sesgo introducido por los bancos de datos primarios que alimentan la IA y la IA generativa limita la creatividad intelectual que es distintiva del espíritu científico. Si bien actualmente resulta difícil para un científico pensar “fuera de la caja” debido a las restricciones administrativas sistemáticas creadas por las políticas científicas institucionales, será aún más complejo en el futuro alejarse del algoritmo de convergencia entre la IA y la evidencia científica sistematizada por metaanálisis. A esto se suma el hecho de que el azar en la ciencia (*serendipity*) — responsable de increíbles beneficios científicos en la medicina (descubrimiento de la penicilina, de la heparina o de los rayos X)— seguramente seguirá siendo residual.

El enorme potencial de la inteligencia artificial conlleva riesgos, lo que exige extrema precaución en su uso. Uno de los principales riesgos es la falta de explicabilidad de la IA, es decir, la incapacidad de comprender cómo se procesa la IA y cómo se logran determinados resultados. La falta de reproducibilidad (la capacidad de otra investigación de mostrar el mismo comportamiento bajo las mismas condiciones, con el mismo método, con la evidencia científica correspondiente) y de replicabilidad (la capacidad de que una investigación independiente llegue a conclusiones similares, aunque no idénticas, cuando hay diferencias en la muestra, los procedimientos de investigación y los métodos de análisis de datos) deben valorarse cuidadosamente, de modo que la confianza hacia la ciencia médica <sup>16</sup> y sus aplicaciones no sea cuestionada <sup>17</sup>.

Por diversas razones, existe una crisis de confianza en la ciencia médica, de la cual la crisis de reproducibilidad es solo una faceta <sup>18</sup>. Otros factores, como la peligrosa disminución de los

niveles de libertad en la planificación, realización y análisis de los estudios científicos<sup>19</sup>, la mercantilización progresiva de la ciencia, la exigencia de criterios, por parte del *peer review*, que limitan objetivamente la creatividad intelectual, pueden verse exacerbadas en el futuro por la IA, por lo que es esencial contar con soluciones creativas para hacer frente éticamente a tales fenómenos.

En medicina, la IA plantea un desafío ético y social considerable. Puede ser extremadamente útil para promover nuevas modalidades de diagnóstico y tratamiento, además de prevenir enfermedades potencialmente mortales. Además, puede proporcionar a los médicos un análisis más preciso y detallado, ayudando en el diagnóstico y tratamiento de muchas enfermedades<sup>20</sup>. En la radiología, por ejemplo, la IA es extremadamente útil, una vez que ayuda a detectar imágenes sospechosas y a mejorar las imágenes, o incluso en el cribado<sup>21</sup>. La radiología es una de las áreas en las que la IA tendrá su impacto a corto plazo, y se anticipa una mayor accesibilidad a este tipo de servicios, especialmente si se combina con la telemedicina y con otras formas de e-salud<sup>22</sup>. Sin embargo, esta agilidad en la interpretación de la imagen puede llevar a generalizar su uso e interpretación por profesionales sanitarios no especializados en el área, lo que implica una reducida capacidad para analizar críticamente los resultados y detectar errores por parte de la IA.

Todas las áreas de la medicina se verán afectadas por la IA, como la dermatología<sup>23</sup>, la detección y el tratamiento del cáncer de mama<sup>24</sup>, las ciencias cardiovasculares<sup>25</sup>, la producción de antibióticos<sup>26</sup> e incluso la nutrición clínica. En este ámbito, la IA puede ayudar a analizar datos complejos, a sugerir intervenciones nutricionales personalizadas, a evaluar las necesidades nutricionales, a prevenir enfermedades o a supervisar la ingesta dietética, orientando a los pacientes hacia opciones alimentarias más saludables<sup>27</sup>.

En los cuidados paliativos, la IA tiene un enorme potencial para apoyar a los médicos en la toma de decisiones e identificar a los pacientes con alto riesgo de mortalidad o de tratamiento inadecuado. Además, las herramientas de IA pueden facilitar algunos aspectos centrales de los cuidados paliativos, como dirigir la entrevista clínica para promover la planificación anticipada de los cuidados o ajustar el plan de tratamiento a las necesidades y preferencias de los pacientes<sup>28</sup>. En los cuidados

continuos, la IA puede ayudar a los cuidadores a apoyar a los adultos mayores o ser útil para monitorear a los pacientes en tiempo real, a veces a largas distancias, como la telemedicina. En la medicina, la evolución ha sido sorprendente, aunque la falta de explicabilidad, dando lugar al conocido problema “caja negra” (el usuario no logra entender el problema con el dispositivo o cómo funciona), si se trata de un problema complejo.

La transformación progresiva a el “médico digital” —un contexto en el que el médico puede asumir el papel de copiloto de la IA en la práctica clínica— requiere una redefinición del marco de responsabilidades y una nueva forma de responsabilidad profesional y organizacional. Cabe recordar, sin embargo, que las aplicaciones modernas de la IA en la medicina clínica, por medio de la transcripción de consultas médicas<sup>29</sup>, ya tienen la capacidad de realizar un diagnóstico diferencial, de hacer sugerencias de métodos diagnósticos complementarios y de producir recomendaciones de tratamiento. Esta función de “*scribe and coach*” conlleva la posibilidad de distorsionar la relación médico-paciente, afectando aspectos esenciales de esta relación, como la intuición médica basada en la experiencia o la empatía y la compasión tan necesarias en la relación entre personas. Por otra parte, cabe preguntarse si la orientación terapéutica resultante de los *prompts* y sus respuestas respectivas serían recomendadas de acuerdo con las *leges artis*, y qué mecanismos efectivos existen para prevenir daños e interurrencias indeseables.

Se debe profundizar siempre en el carácter ético de la relación con el paciente, y este es el mejor capital que la medicina tiene para ofrecer en una sociedad profundamente tecnológica<sup>30</sup>. Es importante asegurar que la relación clínica con el paciente no se vea desvirtuada por una separación entre los profesionales de la salud y los pacientes, y es necesario elaborar medidas para profundizar la relación de confianza en la práctica clínica y en la investigación con seres humanos.

En síntesis, en la interfaz entre medicina, biomedicina, ciencia e investigación, la IA plantea serios desafíos que pueden convertirse en importantes oportunidades de desarrollo colectivo. Tendrá un impacto extraordinario en la gestión sanitaria, tanto en la gestión operativa como en la regulación transnacional del acceso a la salud, por ejemplo, mediante la generalización de la *health wallet*,

billetera digital de datos de salud, una tecnología que reunirá no solo los datos recopilados en la historia clínica electrónica, sino todos los demás datos de salud almacenados, a título de ejemplo:

1. Sistemas digitales en el sector privado y social.
2. Telemedicina y otras formas de e-salud.
3. Aplicaciones móviles de salud.
4. Soluciones de atención virtual.
5. Aplicaciones de monitorización remota.
6. Dispositivos electrónicos diseñados para usar en el cuerpo humano (relojes inteligentes, joyas inteligentes, gafas digitales, estetoscopios con *bluetooth*, rastreadores de *fitness*, etc.).

En cualquier caso, esta sorprendente evolución observada en los últimos años plantea importantes reservas éticas que es necesario profundizar y aclarar.

### Inteligencia artificial, ética y regulación

La implementación de la inteligencia artificial, especialmente de la IA general (*artificial general intelligence*), plantea cuestiones éticas fundamentales, tanto para la sociedad como para la medicina, la ciencia y la investigación. Estas cuestiones están en constante evolución, a medida que las emergentes aplicaciones de IA plantean problemas nuevos, antes no previstos. Aun así, la implementación de la IA debe prestar especial atención al respeto de valores éticos fundamentales como la autonomía humana, la no discriminación, la protección de datos y de la privacidad individual, la justicia y los derechos de las poblaciones vulnerables, como los niños, los adultos mayores y las personas con discapacidad. También implica promover la inclusión digital y combatir la desinformación que puede causar la exclusión digital, con un impacto innegable en el mercado laboral, incluida la medicina y otras profesiones sanitarias.

Cabe asegurar la supervisión y el control humano constante durante el desarrollo y el uso de la IA, pero este esfuerzo por sí solo no es suficiente. Es necesario asegurar que las empresas que realizan investigaciones con y en inteligencia artificial promuevan la transparencia en sus procedimientos e incorporen siempre una perspectiva ética en el diseño de dichos sistemas. Solo entonces la IA será verdaderamente confiable y contribuirá a un desarrollo humano sostenible. De acuerdo con el

Independent High-Level Expert Group on Artificial Intelligence<sup>31</sup>, es esencial asegurar en el desarrollo y aplicación de la IA siete criterios éticos esenciales:

1. Agencia y supervisión humanas.
2. Robustez técnica y seguridad.
3. Privacidad y buena gobernanza de datos.
4. Transparencia.
5. Diversidad, no discriminación y justicia.
6. Bienestar social y ambiental.
7. *Accountability* (responsabilidad).

Otra cuestión importante que plantea el uso de la IA es la evolución del marco de valores sociales, considerando que este puede verse moldeado por el sesgo de la propia tecnología. Se trata de una cuestión crucial a largo plazo, ya que las fuentes primarias que alimentan la IA generativa están fuertemente influenciadas por principios como la doctrina de los derechos humanos, el respeto a la dignidad humana, la protección de la identidad individual, la igualdad y la no discriminación, así como la preservación del medio ambiente y de la biodiversidad, valores fundacionales de las democracias pluralistas y pilares de la ética médica moderna. Sin embargo, nada impide que, en un futuro más o menos próximo, nuevas ideologías alimenten la IA generativa, formando corrientes de opinión menos comprometidas con estos valores y sus referencias éticas, y más centradas en el control de los ciudadanos. Además, puede que la IA se utilice para modificar corrientes de pensamiento previamente consagradas<sup>32</sup>. Actualmente, la IA parece apoyarse en un núcleo sólido de principios éticos<sup>33</sup>, lo que es especialmente relevante en áreas sensibles desde el punto de vista humano, como la medicina.

Por otra parte, la libertad de expresión puede quedar seriamente condicionada, favoreciendo la difusión masiva de desinformación y misinformación, con un impacto significativo en el sector de la salud. Un ejemplo reciente fue la difusión de información falsa durante las campañas de vacunación universal, que tuvo consecuencias para la salud pública. Todos los esfuerzos deben dirigirse a crear una sociedad civil vibrante, ilustrada y bien informada sobre hechos científicos validados por agencias de reconocida credibilidad, lo que implica aumentar los niveles generales de literacia, especialmente en lo que respecta a cómo interactuar con diferentes redes sociales y con la propia IA<sup>34</sup>. También pueden considerarse nuevas formas de

cooperación social, como la “inteligencia colectiva” aplicada a la toma de decisiones públicas<sup>35</sup>.

La OMS reconoce que la IA tiene un enorme potencial para ampliar el acceso a los cuidados de salud<sup>36</sup>. Con la expansión del uso de sistemas digitales, se esperan ganancias significativas en equidad en el acceso a servicios de calidad adecuada, considerando que la salud es un derecho universal<sup>37</sup>. Es necesario preservar la solidaridad, pilar de la vida social y de la construcción del Estado de bienestar social, asegurando que la IA no ponga en riesgo estos derechos ni contribuya a la creación de nuevas formas de exclusión social o de restricción del acceso a los sistemas de salud. La justicia, la igualdad y la no discriminación deben garantizarse a nivel nacional e internacional. En otras palabras, los beneficios esperados deben ser compartidos por la sociedad en su conjunto, y no solo por grupos específicos. La adopción de la IA debe respetar un principio humanitario común y es deseable que contribuya a una mayor justicia global y a la reducción de las desigualdades entre el Norte y el Sur.

Compartir los resultados de investigaciones sobre IA contribuirá no solo a aumentar la productividad a escala mundial, sino también a promover una mayor justicia social. Esto se debe a que las tecnologías digitales y la IA, si se implementan de forma adecuada, pueden desempeñar un papel decisivo en el logro de la igualdad como objetivo universal<sup>38</sup>, fomentando la creación y el fortalecimiento de comunidades inclusivas, seguras, resilientes y ambientalmente sostenibles<sup>39</sup>. Es esencial promover la igualdad de género, garantizando que mujeres y hombres tengan las mismas condiciones para ejercer plenamente sus derechos humanos, así como para contribuir y disfrutar de los beneficios del desarrollo económico, social, cultural y político. En ese sentido, el Foro Económico Mundial destaca la necesidad de invertir en investigación centrada en la mujer, incluida la recopilación de datos científicos sobre el género femenino, la ampliación del acceso de las mujeres a servicios de salud de buena calidad y la promoción del liderazgo femenino<sup>40</sup>.

La IA y la IA generativa plantean importantes desafíos éticos para la medicina y otras ciencias de la vida, y requieren una atención especial en el sector sanitario. A la luz de lo anterior, se presentan a continuación algunas recomendaciones para la implementación ética y responsable de la IA:

1. Promover el respeto a la libertad individual y prevenir el daño a los seres humanos: uno de los valores fundamentales de las sociedades civilizadas es el ejercicio de la libertad individual y de la ciudadanía responsable<sup>41</sup>. En medicina, el ejercicio de la libertad ética individual ha transformado la relación médico-paciente y ha llevado, por ejemplo, a la implementación generalizada del consentimiento informado y de las directivas anticipadas de voluntad, como el testamento vital<sup>42</sup>. La IA puede y debe contribuir al ejercicio de la libertad y de la autodeterminación de la persona, lo que implica asegurar una literacia digital adecuada para la población en general y para los pacientes en particular, dado que muchas herramientas utilizadas actualmente —como la historia clínica electrónica procesada por IA, o el uso de IA generativa para interpretar el consentimiento informado<sup>43</sup>— requiere una inclusión digital sustancial. Para lograr estos objetivos, es importante involucrar a los *stakeholders* a lo largo de todo el ciclo de vida de los sistemas de IA, así como redefinir el sistema de responsabilidad civil, incluida la previsión de seguro específico para la actividad. Es esencial que la implementación de los sistemas de IA —desde el diseño y la investigación hasta la implementación y la operativización— esté siempre bajo control humano. Se deben crear mecanismos sólidos de gobernanza y supervisión de la IA, como el Reglamento sobre Inteligencia Artificial de la Unión Europea<sup>44</sup>. Se requiere especial atención al llamado “riesgo existencial” (*X-Risk*) para la humanidad y la civilización<sup>45</sup>, resultante del desarrollo de la personalidad cibernética o de la singularidad tecnológica<sup>46</sup>.
2. Protección de la privacidad individual, salvaguarda de los datos personales y promoción de una buena gobernanza de los datos y metadatos: la privacidad individual es uno de los derechos más protegidos, ya que la intimidad de la vida privada forma parte de la esfera más identitaria y exclusiva de la persona humana. Directamente relacionado con este derecho está la protección de los datos personales, otro símbolo destacado del Estado democrático de derecho<sup>47</sup>. En el ámbito de la salud, la protección de la privacidad y de los datos personales tiene especial importancia, ya que es esencial para el libre desarrollo de

- la personalidad. En la ciencia y en la investigación, se percibe un enorme apetito por los datos personales de salud, que son sensibles y privados por naturaleza; por lo tanto, la IA y las tecnologías relacionadas —como la computación cuántica— deben desarrollarse bajo el paradigma del respeto a estos valores<sup>48</sup>. Investigadores, aunque bien intencionados, no deben permitir que el *ethos* de la ciencia se anteponga a los derechos individuales, bajo pena de ampliar el alcance del innovador derecho al olvido, es decir, la posibilidad de que un paciente solicite la eliminación de sus datos personales de los sistemas digitales<sup>49</sup>. El uso de datos sintéticos generados por IA puede ser una solución para eludir la limitación de la anonimización y la seudonimización (desidentificación) para garantizar la protección total de los datos personales sensibles<sup>50</sup>, lo que requiere la adopción de una estrategia mundial de ciberseguridad y protección de datos<sup>51</sup>.
3. Garantía de integridad científica y prevención de conflictos de intereses: la ciencia, especialmente la medicina, ya ha superado todas las barreras tecnológicas, planteándose, sin embargo, la cuestión de sus límites éticos y sociales. La integridad científica y la ética en la investigación biomédica se aplican a todas las etapas del desarrollo científico, desde la concepción del proyecto hasta el uso comercial. El uso de IA generativa, como ChatGPT<sup>52</sup>, requiere la redefinición de las reglas de consulta a estudios previamente publicados, dejando clara la autoría original; también demanda repensar los criterios de originalidad, las medidas de combate al plagio, las normas editoriales y de supervisión de revistas científicas, además de reforzar mecanismos que aseguren la replicabilidad y la reproducibilidad de la investigación como prueba de su valor científico<sup>53</sup>. Deben implementarse mecanismos modernos para verificar los conflictos de interés de los investigadores, asegurando que el *ethos* científico no sea desvirtuado por las nuevas tecnologías. Asimismo, es esencial diseñar los sistemas de IA para que respeten estos valores y principios científicos universales<sup>54</sup>, para evitar la *turbocharging misinformation* y el sesgo algorítmico<sup>55</sup>, y que se evalúe la creación de un código de conducta específico para el uso de IA<sup>56</sup>.
  4. Protección de grupos vulnerables, como niños, adultos mayores, personas con discapacidad, migrantes y enfermos en general: es un principio civilizatorio que la IA contribuya a la inclusión, y no a la exclusión. La concepción contemporánea de un modelo de bienestar social que asegure la igualdad de oportunidades para todos, independientemente de la “lotería de la vida”, debe encontrar en la IA un socio ideal para lograr plenamente este objetivo. Para ello, es necesario reestructurar las cadenas de valor en torno a las nuevas tecnologías digitales, con el fin de promover mejoras en la calidad de vida y en el bienestar social. Si bien los sistemas de IA ofrecen beneficios innegables, también presentan riesgos que pueden afectar negativamente a las poblaciones más vulnerables y que son difíciles de predecir, identificar o medir. En la salud mental, por ejemplo, se estima que la realidad virtual podría tener un impacto particularmente relevante, produciendo resultados concretos para los pacientes<sup>57</sup>. Por lo tanto, es urgente implementar una estrategia de alfabetización en IA<sup>58</sup> que contemple las nuevas tecnologías de salud digital, realizadas de forma participativa y positiva, para que todos puedan beneficiarse de ella, especialmente en el acceso a la salud y a los avances científicos<sup>59</sup>.
  5. Garantía de impacto positivo de la IA en la salud mundial y en la promoción de la One Health: tanto en el acceso a los sistemas de salud como en situaciones de emergencia de salud pública, la IA se considera una herramienta indispensable para la gestión de la salud individual y colectiva. Una nueva ola de medicina personalizada y de precisión encuentra en la IA el recurso ideal para su rápida implementación. A su vez, en salud pública, la IA ha demostrado una gran eficacia, en China y Canadá, en el rastreo de contactos, identificando a todas las personas expuestas a gotículas respiratorias o a secreciones de casos de COVID-19. Este rastreo permitió estratificar riesgos de exposición e implementar medidas como el aislamiento profiláctico y la vigilancia epidemiológica, para prevenir la propagación de la enfermedad infecciosa. Respuestas de este tipo adquieren aún mayor importancia a escala mundial, en el contexto de la salud mundial, en el que se requiere una estrategia supranacional y coordinada. Para asegurar un rastreo epidemiológico eficiente, con beneficios concretos para

la salud mundial de las poblaciones, es esencial estandarizar los procedimientos y los sistemas de clasificación. La codificación es esencial para permitir la comparación de estrategias y resultados, como lo ejemplifica la Clasificación Internacional de Intervenciones de Salud<sup>60</sup>. De esta manera, se promueve la One Health, un concepto que reconoce la profunda interconexión entre la salud humana, la salud animal y el medio ambiente que nos rodea<sup>61</sup>.

6. Reconocer la inteligencia artificial fiable como el 18° Objetivo de Desarrollo Sostenible (ODS): dado el impacto transversal e interconectado de la IA en todos los ámbitos de actividad, los 17 ODS actualmente vigentes se verán, de alguna manera, influenciados por esta tecnología<sup>62</sup>. La Agenda 2030 abarca varias dimensiones del desarrollo sostenible, con el propósito de promover la paz, la justicia y la creación de instituciones eficaces. En este contexto, establecer límites claros para la aplicación de la IA surge como una medida inevitable, especialmente relevante en las áreas de la medicina y de las ciencias de la vida y la salud.

### Consideraciones finales

Para que la IA sea considerada confiable en la medicina, en la salud, en la investigación y en la sociedad en general, algunos pasos parecen inevitables. En primer lugar, es esencial desarrollar y utilizar sistemas de IA que respeten la libertad y la autodeterminación de la persona humana, para evitar posibles daños y asegurar justicia en

su uso, especialmente mediante un control humano efectivo. Se debe prestar especial atención a los grupos más vulnerables, como los niños, los adultos mayores y las personas con discapacidad, así como aquellos en riesgo de exclusión social debido a los cambios en las relaciones laborales promovidos por la IA o por el llamado “dividendo generacional”, resultante de la adquisición natural de habilidades digitales por parte de las generaciones más jóvenes en comparación con las anteriores.

Por otra parte, para asegurar el pleno desarrollo humano y un crecimiento económico sostenible y especializado, es necesario estimular la innovación y la adquisición de habilidades digitales, incluida la profundización en la explicabilidad de la IA. Esto requiere un nuevo enfoque de la formación académica transgeneracional y la promoción de la educación universal de la IA en las escuelas, para que los jóvenes adquieran rápidamente habilidades transversales en este ámbito. Las facultades de medicina, a su vez, deben adaptarse a esta realidad y promover, desde ahora, la educación para el uso responsable de la IA en la formación médica.

Se puede concluir que, dado que la IA es una tecnología internacional en rápida expansión, asegurar estos valores éticos requiere redefinir los mecanismos de gobernanza global para asegurar no una prohibición, sino una regulación efectiva e integral. La necesidad de una gobernanza ética eficaz hace fundamental preservar la agencia humana en la aplicación de sistemas de IA (*human in the loop*), así como asegurar total transparencia y una rigurosa *accountability* en la investigación con y sobre inteligencia artificial.

*El autor Rui Nunes ejerce el cargo de editor científico honorario de la Revista Bioética. El presente artículo es parte del homenaje de la revista al 80 aniversario del Consejo Federal de Medicina.*

### Referencias

1. Stahl BC. Artificial intelligence for a better future. An ecosystem perspective on the ethics of AI and emerging digital technologies [Internet]. Cham: Springer; 2021 [acceso 29 jul 2025]. Disponible: <https://tinyurl.com/ftu75btx>
2. World Health Organization. Ethics and governance of artificial intelligence for health. Guidance on large multimodal models [Internet]. Geneva: WHO; 2024 [acceso 29 jul 2025]. Disponible: <https://tinyurl.com/z8cwb4nf>
3. European Commission. European health data space [Internet]. 2024 [acceso 29 jul 2025]. Disponible: <https://tinyurl.com/5n6nbbt2>

4. European Parliament. Artificial intelligence in healthcare [Internet]. 2022 [acesso 29 jul 2025]. Disponível: <https://tinyurl.com/5n97dsvk>
5. Office of Science and Technology Policy. Blueprint for an AI Bill of Rights: making automated systems for the American people [Internet]. 2022 [acesso 29 jul 2025]. Disponível: <https://tinyurl.com/2cua5m8v>
6. Nuffield Council on Bioethics. Artificial intelligence (AI) in healthcare and research. London: Nuffield Council on Bioethics; 2018.
7. Nunes R, Ricou M, Godinho I, Neves MP, coordenadores. Livro branco sobre inteligência artificial (IA): inquietações sociais, propostas éticas e orientações políticas [Internet]. Lisboa: Conselho Nacional de Ética para as Ciências da Vida; 2024 [acesso 29 jul 2025]. Disponível: <https://tinyurl.com/fzavv5ma>
8. European Commission. A definition of AI: main capabilities and scientific disciplines. Brussels: European Commission; 2018.
9. Harari YN. Nexus: a brief history of information networks from the stone age to AI. New York: Random House; 2018.
10. Ananthaswamy A. How close is AI to human-level intelligence?. *Nature* [Internet]. 2024 [acesso 29 jul 2025];636(5):22-5. Disponível: <https://tinyurl.com/3mwefdh>
11. Jones N. How should we test AI for human level intelligence?: Open AI's o3 electrifies quest. *Nature* [Internet]. 2025 [acesso 29 jul 2025];637(8047):774-5. DOI: 10.1038/d41586-025-00110-6
12. Butlin P, Long R, Elmoznino E, Bengio Y, Birch J, Constant A, Deane G *et al.* Consciousness in artificial intelligence: insights from the science of consciousness (Preprint). *Arxiv* [Internet]. 2023 [acesso 29 jul 2025]. Disponível: <https://tinyurl.com/5n8hnjfy>
13. Lenharo M. What should we do if AI becomes conscious? These scientists say it's time for a plan. *Nature* [Internet]. 2024 [acesso 29 jul 2025];636:533-4. DOI: 10.1038/d41586-024-04023-8
14. Lee P, Bubeck S, Petro J. Benefits, limits, and risks of GPT-4 as an AI chatbot for medicine. *N Engl J Med* [Internet]. 2023 [acesso 29 jul 2025];388(13):1233-9. DOI: 10.1056/NEJMSr2214184
15. Haug CJ, Drazen J. Artificial intelligence and machine learning in clinical medicine. *N Engl J Med* [Internet]. 2023 [acesso 29 jul 2025];388(13):1201-8. DOI: 10.1056/NEJMra2302038
16. Ioannidis J. Why replication has more scientific value than original discovery. *Behav Brain Sci* [Internet]. 2018 [acesso 29 jul 2025];41:e137. DOI: 10.1017/S0140525X18000729
17. National Academies of Sciences. Reproducibility and replicability in science. Washington: National Academies Press; 2019.
18. Baker M. Is there a reproducibility crisis?. *Nature* [Internet]. 2016 [acesso 29 jul 2025];533(7604):452-4. Disponível: <https://tinyurl.com/2hc44u9n>
19. Wicherts J, Veldkamp C, Augusteijn H, Bakker M, van Aert RCM, van Assen MALM. Degrees of freedom in planning, running, analyzing, and reporting psychological studies: a checklist to avoid p-Hacking. *Front Psychol* [Internet]. 2016 [acesso 29 jul 2025];7:1832. DOI: 10.3389/fpsyg.2016.01832
20. Ahuja AS. The impact of artificial intelligence in medicine on the future role of the physician. *PeerJ* [Internet]. 2019 [acesso 29 jul 2025];7:e7702. DOI: 10.7717/peerj.7702
21. Rajpurkar P, Lungren MP. The current and future state of AI interpretation of medical images. *New Eng J Med* [Internet]. 2023 [acesso 29 jul 2025];388(21):1981-90. DOI: 10.1056/NEJMra2301725
22. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJ. Artificial intelligence in radiology. *Nat Rev Cancer* [Internet]. 2018 [acesso 29 jul 2025];18:500-10. Disponível: <https://tinyurl.com/56sdzn6m>
23. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM *et al.* Dermatologist-level classification of skin cancer with deep neural networks. *Nature* [Internet]. 2017 [acesso 29 jul 2025];542:115-8. DOI: 10.1038/nature21056
24. McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H *et al.* International evaluation of an AI system for breast cancer screening. *Nature* [Internet]. 2020 [acesso 29 jul 2025];577:89-94. Disponível: <https://tinyurl.com/whe4pfbj>
25. Zargazadeh A, Javanshir E, Ghaffari A, Mosharkesh E, Anari B. Artificial intelligence in cardiovascular medicine: an updated review of the literature. *J Cardiovasc Thorac Res* [Internet]. 2023 [acesso 29 jul 2025];15(4):204-9. DOI: 10.34172/jcvtr.2023.33031

26. Marchant J. Powerful antibiotics discovered using AI. *Nature* [Internet]. 2020 [acceso 29 jul 2025]. DOI: 10.1038/d41586-020-00018-3
27. Bond A, Mccay K, Lal S. Artificial intelligence & clinical nutrition: what the future might have in store. *Clin Nutr* [Internet]. 2023 [acceso 29 jul 2025];57:542-9. DOI: 10.1016/j.clnesp.2023.07.082
28. Peruselli C, De Panfilis L, Gobber G, Melo M, Tanzi S. Artificial intelligence and palliative care: opportunities and limitations. *Recent Prog Med* [Internet]. 2020 [acceso 29 jul 2025];111(11):639-45. DOI: 10.1701/3474.34564
29. Yadav GS, Longhurst CA. Will AI make the electronic health record more efficient for clinicians?. *N Eng J Med AI* [Internet]. 2025 [acceso 29 jul 2025];2(3). DOI: 10.1056/AIe2500020
30. Mittelstadt B. The impact of artificial intelligence on the doctor-patient relationship [Internet]. Strasbourg: Council of Europe; 2021 [acceso 29 jul 2025]. Disponible: <https://bit.ly/4mcgv9Z>
31. European Commission. Ethics guidelines for trustworthy AI. Brussels: European Commission; 2019.
32. Mann PS, Earp BD, Møller N, Vynn S, Savulescu J. AUTOGEN: a personalized large language model for academic enhancement. Ethics and proof of principle. *Am J Bioeth* [Internet]. 2023 [acceso 29 jul 2025];23(10):28-41. DOI: 10.1080/15265161.2023.2233356
33. Nuñez M. Anthropic just analyzed 700.000 Claude conversations and found its AI has a moral code of its own. *VentureBeat* [Internet]. 2025 [acceso 29 jul 2025]. Disponible: <https://tinyurl.com/ykf5s9by>
34. Hub France IA. ChatGPT: usages, impacts et recommandations – note de synthèse [Internet]. Paris: Hub France IA; 2023 [acceso 29 jul 2025]. Disponible: <https://tinyurl.com/yvpnef7r>
35. Ryan M, Gambrell D, Noveck BS. Using collective intelligence to solve public problems [Internet]. New York: The Govlab; 2020 [acceso 29 jul 2025]. Disponible: <https://tinyurl.com/bdtsmjw7>
36. World Health Organization. Global diffusion of eHealth: making universal health coverage achievable. Report of the third global survey on eHealth [Internet]. Geneva: WHO; 2016 [acceso 29 jul 2025]. Disponible: <https://tinyurl.com/mry9f6vb>
37. Nunes R. Healthcare as a universal human right: sustainability in global health. New York: Routledge; 2022.
38. Venkatapuram S. Health justice. Cambridge: Polity Press; 2011.
39. Ren S. How much water does AI consume? The public deserves to know [Internet]. OECD AI Policy Observatory; 2023 [acceso 29 jul 2025]. Disponible: <https://tinyurl.com/2rjyfpnh>
40. World Economic Forum. Closing the women's health gap: a \$1 trillion opportunity to improve lives and economies [Internet]. Geneva: WEF; 2024 [acceso 29 jul 2025]. Disponible: <https://tinyurl.com/52h3hkcc>
41. Nunes R. Ensaaios em bioética [Internet]. Brasília: CFM; 2017 [acceso 29 jul 2025]. Disponible: <https://tinyurl.com/2se5p2yp>
42. Nunes R. Diretivas antecipadas de vontade [Internet]. Brasília: CFM; 2016 [acceso 29 jul 2025]. Disponible: <https://tinyurl.com/354nv2jm>
43. Mirza FN, Tang OY, Connolly ID, Abdulrazeq HA, Lim RK, Roye GD et al. Using ChatGPT to facilitate truly informed medical consent. *N Eng J Med AI* [Internet]. 2024 [acceso 29 jul 2025];1(2). DOI: 10.1056/Alcs2300145
44. European Union. Proposal for a regulation of the European parliament and of the council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts [Internet]. Brussels: European Union; 2023. Disponible: <https://tinyurl.com/yczhtmnj>
45. Jecker NS, Atuire CA, Bélisle-Pipon JC, Ravitsky V, Ho A. AI and the falling sky: interrogating X-Risk. *J Med Ethics* [Internet]. 2024 [acceso 29 jul 2025];50(12):811-7. DOI: 10.1136/jme-2023-109702
46. Penrose R, Severino E, Scardigli F, Testoni I, Vitiello G, D'Ariano G et al. Artificial intelligence versus natural intelligence. Cham: Springer; 2022.
47. European Union. General data protection regulation. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data [Internet]. Brussels: European Union; 2016 [acceso 29 jul 2025]. Disponible: <https://tinyurl.com/y93u93bh>
48. European Commission. Study on health data, digital health and artificial intelligence in healthcare [Internet]. 2022 [acceso 29 jul 2025]. Disponible: <https://tinyurl.com/2rn8snf5>

49. Correia M, Rego G, Nunes R. Gender transition: is there a right to be forgotten? *Health Care Anal* [Internet]. 2021 [acceso 29 jul 2025];29(4):283-300. DOI: 10.1007/s10728-021-00433-1
50. Kokosi T, Harron K. Synthetic data in medical research. *BMJ Med* [Internet]. 2022 [acceso 29 jul 2025];26:1(1):e000167. DOI: 10.1136/bmjmed-2022-000167
51. World Health Organization. Global strategy on digital health 2020-2025 [Internet]. Geneva: WHO; 2021 [acceso 29 jul 2025]. Disponible: <https://tinyurl.com/54z2rpew>
52. Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepaño C *et al.* Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. *MedRxiv* [Internet]. 2022 [acceso 29 jul 2025]. DOI: 10.1101/2022.12.19.22283643
53. Bianchini S, Müller M, Pelletier P. Artificial intelligence in science: an emerging general method of invention. *Res Policy* [Internet]. 2022 [acceso 29 jul 2025];51(10):104604. DOI: 10.1016/j.respol.2022.104604
54. Constantinescu M, Crisp R. Can robotic AI systems be virtuous and why does this matter?. *Int J Soc Robot* [Internet]. 2022 [acceso 29 jul 2025];14(6):1547-57. DOI: 10.1007/s12369-022-00887-w
55. Chen J. Who's training whom? *N Eng J Med AI* [Internet]. 2024 [acceso 29 jul 2025];1(5). DOI: 10.1056/Alp2400006
56. Daley K. Two arguments against humanfriendly AI. *AI Ethics* [Internet]. 2021 [acceso 29 jul 2025];1:435-44. DOI: 10.1007/s43681-021-00051-6
57. Freeman D, Reeve S, Robinson A, Ehlers A, Clark D, Spanlang B *et al.* Virtual reality in the assessment, understanding, and treatment of mental health disorders. *Psychol Med* [Internet]. 2017 [acceso 29 jul 2025];47(14):2393-400. DOI: 10.1017/S003329171700040X
58. Organization for Economic Cooperation and Development. Empowering learners for the age of AI: an AI literacy framework for primary and secondary education. Paris: OECD; 2025.
59. Ratwani RM, Classen D, Longhurst C. The compelling need for shared responsibility of AI oversight: lessons from Health IT Certification. *JAMA* [Internet]. 2024 [acceso 29 jul 2025];332(10):787-88. DOI: 10.1001/jama.2024.12630
60. Wübbeler M, Geis S, Stojanovic J, Elliott L, Gutierrez-Ibarluzea I, Lenoir-Wijnkoop I. Coding public health interventions for health technology assessments: a pilot experience with WHO's International Classification of Health Interventions (ICHI). *Front Public Health* [Internet]. 2021 [acceso 29 jul 2025];9:620637. DOI: 10.3389/fpubh.2021.620637
61. Vithanage M, Prasad MNV. One health: human, animal, and environment triad. Hoboken: John Wiley & Sons; 2023.
62. Nunes R, Nunes SB. Reliable artificial intelligence: the 18th sustainable development goal. *J Ethics Legal Technol* [Internet]. 2024 [acceso 29 jul 2025];6(2):5-20. DOI: 10.14658/pupj-JELT-2024-2-2

Rui Nunes - Doctor - [ruinunes@med.up.pt](mailto:ruinunes@med.up.pt)

 0000-0002-1377-9899

Sofia B Nunes - Estudiante de doctorado - [asnunes@med.up.pt](mailto:asnunes@med.up.pt)

 0000-0002-7182-9890

#### Correspondencia

Alameda Prof Hernâni Monteiro, 4200-319. Porto, Portugal.

#### Participación de los autores

Los autores participaron igualmente en la elaboración de este artículo.

Editora responsable: Dilza Teresinha Ambrós Ribeiro

Recibido: 4.8.2025

Revisado: 10.8.2025

Aprobado: 14.8.2025