

# Inteligencia artificial: desafíos éticos y futuros

Jhadson Silva Leonel<sup>1</sup>, Camila Ferreira Silva Leonel<sup>2</sup>, Jonas Byk<sup>1</sup>, Sylvania da Conceição Furtado<sup>1</sup>

1. Universidade Federal do Amazonas, Manaus/AM, Brasil. 2. Faculdade Estácio do Amazonas, Manaus/AM, Brasil.

## Resumen

Esta revisión narrativa de la literatura aborda la evolución y la aplicación de la inteligencia artificial en la salud, con énfasis en avances como la interpretación de imágenes para diagnósticos y el uso de robots en hospitales. Se discuten cuestiones éticas, incluida la privacidad de los datos y el sesgo algorítmico, enfatizando los principios éticos y las regulaciones, como la Ley General de Protección de Datos, en Brasil. Se ejemplifica el desafío del sesgo, resaltando la necesidad de eliminar la discriminación en los sistemas, y se aborda la complejidad de la responsabilidad y la transparencia en la medicina, centrándose en la importancia de comprender los algoritmos para evitar errores médicos. Se destaca la relevancia del debate acerca del potencial positivo de la inteligencia artificial, reforzando que es necesario enfrentar dilemas éticos para asegurar el uso ético y responsable de la tecnología en el futuro.

**Palabras clave:** Inteligencia artificial. Ética. Medicina.

## Resumo

### Inteligência artificial: desafios éticos e futuros

Esta revisão narrativa da literatura aborda a evolução e a aplicação da inteligência artificial na saúde, com ênfase em avanços como interpretação de imagens para diagnósticos e uso de robôs em hospitais. Questões éticas, incluindo privacidade de dados e vies algorítmico, são discutidas, enfatizando princípios éticos e regulamentações, como a Lei Geral de Proteção de Dados, no Brasil. O desafio do vies é exemplificado, ressaltando a necessidade de eliminar discriminação nos sistemas, e aborda-se a complexidade da responsabilidade e transparência na medicina, com foco na importância de compreender algoritmos para evitar erros médicos. Ressalta-se a relevância do debate acerca do potencial positivo da inteligência artificial, reforçando que é preciso enfrentar dilemas éticos para garantir o uso ético e responsável da tecnologia no futuro.

**Palavras-chave:** Inteligência artificial. Ética. Medicina.

## Abstract

### Artificial intelligence: ethical and future challenges

This narrative literature review addresses the evolution and use of artificial intelligence in health care, with a focus on advancements such as image interpretation in diagnostics and the use of robots in hospitals. Ethical issues are discussed, including data privacy and algorithmic bias, emphasizing ethical principles and regulations such as the Brazilian General Data Protection Law. The challenge of bias is illustrated, highlighting the need to eliminate discrimination in systems, and the complexity of responsibility and transparency in medicine is addressed, with a focus on the importance of understanding algorithms to avoid medical errors. The relevance of the debate on the positive potential of artificial intelligence is underscored, stressing the need to tackle ethical dilemmas to ensure the ethical and responsible use of technology in the future.

**Keywords:** Artificial intelligence. Ethics. Medicine.

Los autores declaran que no existe ningún conflicto de interés.

La inteligencia artificial (IA) ha desempeñado un papel cada vez más importante en los últimos años. Las máquinas ya no se limitan a ejecutar tareas físicas, sino que también desempeñan funciones intelectuales que requieren lo que se considera inteligencia. El origen de este concepto se remonta a la década de 1950, cuando el matemático británico Alan Turing propuso la famosa Prueba de Turing en su artículo "Computing Machinery and Intelligence"<sup>1</sup>. A pesar de esto, durante el verano de 1956 se produjo un hito decisivo con la Conferencia del Dartmouth College, que es ampliamente reconocida como el inicio de la IA como campo de estudio<sup>2</sup>.

Inicialmente, la IA se aplicaba principalmente a la resolución de problemas del mundo real mediante la programación del conocimiento de expertos en programas informáticos. Estos programas, llamados sistemas expertos o sistemas basados en conocimiento, se desarrollaron con base en entrevistas a expertos en determinadas áreas. Sin embargo, hubo limitaciones, como la subjetividad y la falta de cooperación de los expertos<sup>3</sup>.

En la actualidad, existe un gran entusiasmo con relación a los beneficios potenciales de la IA, una vez que las máquinas están aprendiendo a conducir automóviles independientes, y los traductores automáticos se están volviendo cada vez más precisos. Además, la IA está presente en tareas cotidianas como leer correos electrónicos, lavar ropa y recomendar películas en plataformas de streaming<sup>4</sup>.

En el ámbito de la salud, la IA se ha aplicado en varias investigaciones académicas con resultados prometedores. Se desarrollan algoritmos para interpretar imágenes de radiografía<sup>5</sup>, mamografía<sup>6</sup>, tomografía computarizada<sup>7</sup>, resonancia magnética<sup>8</sup> e identificar enfermedades como el Alzheimer<sup>9</sup>. Además, se ha utilizado en la patología<sup>10</sup> para identificar lesiones cancerosas<sup>11</sup>, en la interpretación de imágenes de la retina<sup>12</sup> y en la detección de arritmias<sup>13</sup> y otras condiciones médicas.

Los robots equipados con IA también se utilizan en hospitales y clínicas, como el robot creado por la empresa iFlytek, que aprobó el examen nacional de licencia médica nacional en China<sup>14</sup>. Otro ejemplo es el software brasileño Laura, que identifica por medio de IA<sup>15</sup> el riesgo de infección hospitalaria. Estas aplicaciones han demostrado el potencial de la IA en el ámbito de la salud.

El rápido desarrollo de tecnologías para procesar y almacenar datos ha impulsado el crecimiento de la IA. Y, a medida que los problemas se vuelven más complejos y la cantidad de datos aumenta, es necesario desarrollar herramientas computacionales avanzadas y personalizadas basadas en el aprendizaje automático, que dependan cada vez menos de la intervención humana. Sin embargo, este desarrollo va acompañado de preocupaciones, principalmente con relación a la ética y al impacto en la sociedad, considerando cuestiones legales y de responsabilidad, para garantizar que la IA sea beneficiosa y justa.

Este artículo tiene como objetivo realizar una revisión narrativa de la literatura con un enfoque exploratorio mediante el análisis de artículos con información sobre el uso de la IA en los servicios de salud y sus implicaciones éticas.

### Aspectos éticos

Una corriente tradicional de la filosofía sostiene que la ética es exclusiva de los seres humanos, y que las cosas en sí mismas no son necesariamente buenas o malas. *La tecnología es amoral*, comentó Bill Gates en una entrevista. Añadió *nos corresponde a nosotros reflexionar profundamente sobre las nuevas tecnologías y sobre cómo deben o no utilizarse*, como menciona Garattoni<sup>16</sup>. Sin embargo, los desafíos que presenta la IA reavivan la cuestión de la exclusividad humana en la ética.

Desde el principio, la interacción entre humanos y máquinas ha generado debates. En 1950, el escritor y bioquímico estadounidense Isaac Asimov, en su obra *Yo, robot*<sup>17</sup>, abordó la complejidad del tema y creó un clásico al enunciar las tres leyes de la robótica: 1) un robot no puede dañar a un ser humano o, por omisión, permitir que un ser humano sufra algún mal; 2) los robots deben obedecer las órdenes de los humanos, salvo cuando dichas órdenes entren en conflicto con la primera ley; y 3) un robot debe proteger su propia existencia siempre que esto no entre en conflicto con las leyes anteriores. Posteriormente, añadió una cuarta ley, conocida como Ley Cero<sup>17</sup>, que establece que un robot no puede causar daño a la humanidad o, por omisión, permitir que la humanidad sufra algún mal.

Recientemente, cinco principios éticos destacan en las principales iniciativas de regulación y en

las discusiones relacionadas con la IA: beneficencia, no maleficencia, autonomía, justicia y explicabilidad<sup>18</sup>. Estos principios éticos, definidos como valores o pautas generales que guían el comportamiento humano, sirven como base para evaluar la moralidad de las acciones y proporcionan un marco general para la toma de decisiones sobre cuestiones éticas.

## Problemas relacionados con la ética algorítmica

### Privacidad de los datos

Con el avance de la tecnología, las personas están cada vez más conectadas y producen una cantidad cada vez mayor de datos, generados a partir de diversas fuentes, como teléfonos móviles, automóviles, transacciones con tarjetas de crédito y elecciones de contenido en los servicios de *streaming* y dispositivos de internet de las cosas. Esta información es almacenada por varias empresas, que crean bases de datos con la información de los usuarios.

Debido a la limitación humana, es prácticamente imposible realizar un análisis exhaustivo de estos datos, una vez que requeriría mucho tiempo para una evaluación completa y la formulación de conclusiones estructuradas. Sin embargo, con el auge de la IA y el aumento de la capacidad de procesamiento, estos datos pueden analizarse de diversas formas en un corto espacio de tiempo.

Además, una preocupación cada vez más pertinente en la actualidad es el destino de la información almacenada en grandes bases de datos. Grandes empresas, como Meta (anteriormente Facebook), se han enfrentado a demandas legales por compartir datos sin el debido consentimiento.

La filtración de datos de millones de usuarios de Facebook, obtenida por medio de los algoritmos de la empresa Cambridge Analytica, permitió acceder a información confidencial sobre el comportamiento de 50 millones de usuarios de la red social, que sirvió para orientar las estrategias de campaña. Este caso es un ejemplo notable de la discusión sobre vigilancia, privacidad y algoritmos<sup>19</sup>.

Cada día, las personas se ven inundadas de anuncios inmediatamente después de realizar una búsqueda en internet, ya sea por interés de compra o por mera curiosidad. Datos relacionados con la

salud, la alimentación, los hábitos de ejercicio físico e incluso los lugares frecuentados por usuarios de teléfonos inteligentes y otros dispositivos pueden, gracias a su capacidad avanzada de procesamiento, revelar tendencias de salud e información que antes se desconocía. Sin embargo, diferentes empresas pueden tener dicha información, y las personas a menudo desconocen cómo y quién la utilizará.

Debido a la creciente preocupación por la seguridad de los datos, entidades como la Organización Mundial de la Salud (OMS) elaboraron en el 2021 un informe mundial que contiene seis principios esenciales para asegurar que la IA funcione en beneficio del interés público en todos los países<sup>20</sup>. Además, la Organización para la Cooperación y el Desarrollo Económico (OCDE) ha elaborado un conjunto de directrices intergubernamentales para guiar a los Estados en la regulación de la IA, manteniendo su innovación, confiabilidad, seguridad y respeto por los derechos humanos y los valores democráticos.

Treinta y ocho países miembros se han adherido a las Recomendaciones del Consejo de la OCDE sobre IA, un documento que establece principios que seguir internamente<sup>21</sup>. Varios otros países también están desarrollando su propia legislación para garantizar una mayor seguridad y transparencia en el uso de información privada por parte de las empresas.

En la Unión Europea, por ejemplo, se promulgó la Resolución del Parlamento Europeo del 16 de febrero del 2017, que contiene recomendaciones sobre disposiciones de derecho civil relacionadas con la robótica. La Resolución busca establecer reglas coordinadas entre los Estados miembros para garantizar la seguridad en el desarrollo de la IA, así como la protección de los derechos fundamentales de las personas y empresas, sin inhibir la inversión, la innovación y la utilización<sup>22</sup>.

En Brasil, Ley 13.709/2018<sup>23</sup>, conocida como Ley General de Protección de Datos (LGPD), se creó con el objetivo de proteger los derechos fundamentales a la libertad y, más concretamente, a la privacidad. Además, el Senado Federal instituyó la Comisión de Juristas Responsable de Subvencionar la Elaboración de un Sustitutivo sobre Inteligencia Artificial en Brasil (CJSUBIA), que trabaja en la elaboración de la regulación legal brasileña relacionada con el uso de IA.

También hay algunos proyectos de ley (PL) en curso para regular la IA, a saber: el PL 5051/2019<sup>24</sup>,

PL 21/2020<sup>25</sup>, PL 872/2021<sup>26</sup> y el PL 759/2023<sup>27</sup>. La importancia de este tema se evidencia en el número de PL en curso, lo que destaca la necesidad de una regulación legal en este ámbito en crecimiento.

### Sesgo en la inteligencia artificial

El uso creciente de la IA acarrea una gran preocupación sobre el sesgo que puede estar presente o ser amplificado por sistemas que utilizan esta tecnología. Para comprender este fenómeno, es fundamental analizar el proceso de aprendizaje de las máquinas, que puede darse mediante un paso a paso predefinido, conocido como algoritmo, algo similar a un estudio realizado con libros o con la orientación de profesores.

Además, está el enfoque del aprendizaje automático, que no enseña directamente, sino que proporciona instrucciones sobre cómo aprender con base en ejemplos y datos, brindando autonomía al proceso de aprendizaje. En este contexto, los datos desempeñan un papel fundamental, y, cuanto más datos confiables estén disponibles para el entrenamiento, mejor será el modelo resultante.

Sin embargo, los algoritmos no están libres de valores y suposiciones, ya que los parámetros operativos son establecidos por los desarrolladores y configurados por los usuarios que buscan resultados específicos, lo que puede priorizar algunos valores e intereses en detrimento de otros. Incluso en los parámetros establecidos, no se garantiza automáticamente un comportamiento éticamente aceptable<sup>28,29</sup>.

Entre varios ejemplos notorios de sesgo, podemos mencionar un concurso de belleza realizado en el 2016, que utilizó un jurado compuesto por robots generados por IA y en el que se inscribieron miles de personas de varios países, que enviaron sus fotografías para su evaluación<sup>30</sup>. El resultado llamó la atención porque, de los 44 ganadores, solo uno tenía la piel oscura, lo que generó un acalorado debate sobre cómo el sistema podría perpetuar sesgos, produciendo resultados no intencionados, a veces distorsionados e incluso racistas.

El análisis reveló que la base de entrenamiento del sistema estaba compuesta por imágenes de actores y actrices de Hollywood, que eran predominantemente personas blancas. No hubo intención por parte de los desarrolladores de favorecer a un grupo étnico, pero se constató negligencia al verificar los sesgos en la base de entrenamiento<sup>31</sup>.

Un ejemplo relevante de sesgo relacionado con la salud ocurrió con un programa utilizado por una gran compañía de seguros estadounidense. El objetivo era identificar a los usuarios que necesitaban cuidados adicionales, permitiendo a un equipo multidisciplinario realizar un seguimiento preventivo para evitar tratamientos costosos, como ingresos en unidades de cuidados intensivos (UCI). El sistema se basó en la frecuencia con la que los usuarios buscaban atención médica y en los costos de consultas y hospitalizaciones; para evitar sesgo racial, los datos no contenían información sobre la raza de los usuarios<sup>32</sup>.

Sin embargo, el análisis reveló que la mayoría de los usuarios identificados como necesitados de atención adicional eran blancos. Investigaciones adicionales revelaron que los usuarios negros, debido a restricciones financieras, usaban menos el plan de salud, lo que resultaba en menos datos disponibles para que el sistema identificara patrones de gravedad en su atención. Esto pone de relieve cómo los sesgos y prejuicios pueden ocultarse en los datos, dificultando su detección<sup>32</sup>.

Otros ejemplos de sesgo en los resultados de la IA incluyen anuncios de empleo para puestos muy bien remunerados, con menos probabilidad de que se presenten a mujeres<sup>33</sup>; búsquedas de nombres con sonoridad claramente negra, que tienen mayor probabilidad de desencadenar anuncios relacionados con registros de arrestos<sup>34</sup>; y búsquedas de imágenes para profesiones como CEO, que devuelven menos imágenes de mujeres<sup>35</sup>. Además, los sistemas de reconocimiento facial, ampliamente utilizados en la aplicación de la ley, tienen un desempeño inferior a la hora de reconocer los rostros de mujeres de cualquier color y de individuos negros<sup>36</sup>.

Eliminar los sesgos y asegurar la imparcialidad en los sistemas basados en IA se ha convertido en un desafío cada vez mayor, y es crucial reconocer que los datos no son neutrales, ya que reflejan decisiones humanas que pueden ser prejuiciosas; por lo tanto, las bases de datos de entrenamiento deben ser confiables y libres de sesgos. Empresas como Google e IBM se esfuerzan por desarrollar métodos técnicos que permitan evaluar y mitigar los sesgos en sus conjuntos de datos. Además, los organismos reguladores y los gobiernos desarrollan políticas y regulaciones específicas para abordar los sesgos en la IA, reconociendo la necesidad de transparencia y responsabilidad en este campo en constante evolución<sup>37,38</sup>.

## Responsabilidad y transparencia en la inteligencia artificial

El actual avance tecnológico desempeña un papel fundamental en el campo de la medicina, con el creciente uso de equipos inteligentes basados en sistemas de IA. Estas innovaciones son cruciales en la medicina diagnóstica, de modo que la interpretación de los resultados, tradicionalmente realizada por médicos, cuenta ahora con la ayuda de algoritmos sumamente especializados y automatizados.

Según lo que establece el Código de Ética Médica (CEM), se prohíbe a los médicos *dejar de usar todos los medios disponibles de promoción de la salud y de prevención, diagnóstico y tratamiento de enfermedades, científicamente reconocidos y a su alcance, en favor del paciente*<sup>39</sup>. Esto resalta que la negligencia en el uso de los recursos disponibles constituye una clara violación del CEM. Sin embargo, la falta de transparencia y la complejidad de los algoritmos pueden dificultar la comprensión por parte de los profesionales de la salud con relación a las conclusiones del sistema.

Esto puede hacer que los médicos duden en adoptar sistemas de IA que no comprenden completamente, lo que dificulta la integración efectiva en la práctica clínica. Además, la falta de comprensión puede aumentar el riesgo de errores, ya que los profesionales son incapaces de identificar decisiones incorrectas, lo que torna compleja la atribución de responsabilidad en casos de errores médicos, generando desafíos legales y éticos.

Según la Ley 3.268/1957<sup>40</sup>, que regula los consejos de medicina, y con la Ley 12.842/2013<sup>41</sup>, que reafirma la anterior y se conoce como Ley del Acto Médico, el diagnóstico de enfermedades es prerrogativa exclusiva de los médicos debidamente habilitados y registrados en los Consejos Regionales de Medicina (CRM). Por lo tanto, los médicos no pueden eximirse de la responsabilidad del diagnóstico de enfermedades, ya que esto es un aspecto fundamental de su profesión.

Asignar responsabilidades a los médicos ante la complejidad de los sistemas de IA puede resultar en una distribución injusta de responsabilidades. Esto se debe a la compleja realidad que implica la responsabilidad médica con relación a los parámetros utilizados tradicionalmente<sup>42,43</sup>.

Los juristas argumentan que el análisis debe tener en cuenta si el daño causado por el sistema de IA fue resultado de negligencia, impericia o mala

conducta por parte del usuario, y que los médicos no deben ser responsabilizados por los daños causados por agentes autónomos. Esto se debe al hecho de que los sistemas inteligentes tienen autonomía y acceso a los datos, lo que los hace capaces de tomar decisiones independientes<sup>42,43</sup>.

Esta complejidad que implica la identificación y atribución de responsabilidades por los daños resultantes de actividades algorítmicas es un desafío multisectorial. Por lo tanto, detectar el daño, determinar su causa y, lo más importante, identificar quién debe ser responsabilizado por dichos daños son tareas complejas.

Diversos actores humanos están involucrados, incluidos programadores, proyectistas, fabricantes, diseñadores y usuarios, y la cuestión ética de la responsabilidad está estrechamente vinculada al tema de la transparencia. Solo algoritmos no opacos permiten identificar los responsables del proceso y los posibles errores. En este contexto, es esencial un esfuerzo continuo para garantizar que los sistemas de IA en la medicina sean transparentes, comprensibles y estén sujetos a regulaciones que consideren los desafíos éticos y legales que surgen con la evolución tecnológica, para que sigan desempeñando un papel cada vez más significativo.

## Consideraciones finales

El avance de la IA ha transformado el mundo, permitiendo que las máquinas desempeñen tareas complejas y racionales, que antes se consideraban privilegios humanos. Aunque tiene sus raíces en las décadas de 1950 y 1960, su impacto se ha vuelto especialmente notable en las últimas décadas, a medida que los avances tecnológicos han permitido el desarrollo de sistemas cada vez más sofisticados.

Hoy en día, esta tecnología impregna la vida cotidiana y abarca desde la industria del automóvil hasta el diagnóstico médico de precisión. Así, los hospitales y clínicas cuentan con sistemas autónomos para mejorar la salud y el bienestar de los pacientes.

Sin embargo, el rápido crecimiento no está exento de desafíos, ya que la creciente complejidad de los problemas y la cantidad masiva de datos generados requieren herramientas computacionales más

autónomas. En este escenario, el aprendizaje automático desempeña un papel central, permitiendo que los sistemas adquieran conocimiento y tomen decisiones con una mínima intervención humana.

Si bien promete beneficios significativos, la tecnología también plantea preocupaciones éticas, como la privacidad, el sesgo algorítmico y la

responsabilidad, por lo que, a medida que avanza la IA, la sociedad necesita promover discusiones y regulaciones sensatas que garanticen su uso ético y justo. Los beneficios son innegables, pero su potencial solo se aprovechará plenamente si los sujetos involucrados están preparados para afrontar los desafíos y adoptar enfoques responsables.

## Referencias

1. Turing AM. Computing machinery and intelligence (1950). In: Lewis HR, editor. Ideas that created the future. Boston: MIT Press; 2021. p. 147-64.
2. Patel VL, Shortliffe EH, Stefanelli M, Szolovits P, Berthold MR, Bellazzi R, Abu-Hanna A. The coming of age of artificial intelligence in medicine. *Artif Intell Med* [Internet]. 2009 [acceso 22 maio 2024];46(1):5-17. DOI: 10.1016/j.artmed.2008.07.017
3. Mendes RD. Inteligência artificial: sistemas especialistas no gerenciamento da informação. *Ciênc Inf* [Internet]. 1997 [acceso 22 maio 2024];26(1):39-45. DOI: 10.1590/s0100-19651997000100006
4. Faceli K, Lorena AC, Gama J, de Almeida TA, Carva ACPLF. Inteligência artificial: uma abordagem de aprendizado de máquina. Rio de Janeiro: LTC; 2021.
5. Nam JG, Park S, Hwang EJ, Lee JH, Jin K-N, Lim KY *et al.* Development and validation of deep learning-based automatic detection algorithm for malignant pulmonary nodules on chest radiographs. *Radiology* [Internet]. 2019 [acceso 22 maio 2024];290(1):218-28. DOI: 10.1148/radiol.2018180237
6. Wu N, Phang J, Park J, Shen Y, Huang Z, Zorin M *et al.* Deep neural networks improve radiologists' performance in breast cancer screening. *IEEE Trans Med Imaging* [Internet]. 2020 [acceso 22 maio 2024];39(4):1184-94. DOI: 10.1109/TMI.2019.2945514
7. Shadmi R, Mazo V, Bregman-Amitai O, Elnekave E. Fully-convolutional deep-learning based system for coronary calcium score prediction from non-contrast chest CT [Internet]. In: Proceedings of 2018 IEEE 15th International Symposium on Biomedical Imaging; 4-7 abr 2018; Washington. Washington: IEEE; 2018 [acceso 22 maio 2024]. p. 24-28. DOI: 10.1109/ISBI.2018.8363515
8. Crimi A, Bakas S, editors. Brainlesion: glioma, multiple sclerosis, stroke and traumatic brain injuries: 7th International Workshop, BrainLes 2021, held in conjunction with MICCAI 2021, virtual event, September 27, 2021: revised selected papers [Internet]. Part I. Cham: Springer International Publishing; 2022 [acceso 22 maio 2024]. DOI: 10.1007/978-3-031-08999-2
9. Ding Y, Sohn JH, Kawczynski MG, Trivedi H, Harnish R, Jenkins NW *et al.* A deep learning model to predict a diagnosis of Alzheimer disease by using 18F-FDG PET of the brain. *Radiology* [Internet]. 2019 [acceso 22 maio 2024];290(2):456-64. DOI: 10.1148/radiol.2018180958
10. Chang HY, Jung CK, Woo JI, Lee S, Cho J, Kim SW, Kwak TY. Artificial intelligence in pathology. *J Pathol Transl Med* [Internet]. 2019 [acceso 22 maio 2024];53(1):1-12. DOI: 10.4132/jptm.2018.12.16
11. Brinker TJ. Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task. *Eur J Cancer* [Internet]. 2019 [acceso 22 maio 2024];113:47-5. DOI: 10.1016/j.ejca.2019.04.001
12. Fauw J, Ledsam JR, Romera-Paredes B, Nikolov S, Tomasev N, Blackwell S *et al.* Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat Med* [Internet]. 2018 [acceso 22 maio 2024];24(9):1342-50. DOI: 10.1038/s41591-018-0107-6
13. Hannun AY Rajpurkar P, Haghpanahi M, Tison GH, Bourn C, Turakhia MP, Ng AY. Publisher correction: cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nat Med* [Internet]. 2019 [acceso 22 maio 2024];25(3):530. DOI: 10.1038/s41591-019-0359-

14. Saracco R. Congrats Xiaoyi: you are now a medical doctor. IEEE Future Directions [Internet]. 2017 [acesso 22 maio 2024]. Disponível: <https://bit.ly/4bH91G7>
15. Instituto Laura Fressatto. Quem somos [Internet]. 2019 [acesso 22 maio 2024]. Disponível: <https://institutolaura.org/quem-somos>
16. Garattoni B. Bill Gates indica 5 livros para você ler neste final de ano. Superinteressante [Internet]. 2016 [acesso 22 maio 2024]. Disponível: <https://bit.ly/461cwGd>
17. Asimov I. Eu, robô. São Paulo: Aleph; 2014.
18. Floridi L, Cowls J. A unified framework of five principles for AI in society. Harv Data Sci Rev [Internet]. 2019 [acesso 22 maio 2024];1:2-15. DOI: 10.1162/99608f92.8cd550d1
19. Confessore N. Cambridge Analytica and Facebook: the scandal and the fallout so far. New York Times [Internet]. 4 abr 2018 [acesso 22 maio 2024]. Disponível: <https://nyti.ms/3W5Xmeg>
20. World Health Organization. Ethics and governance of artificial intelligence for health: WHO guidance [Internet]. Geneva: World Health Organization; 2021 [acesso 22 maio 2024]. Disponível: <https://bit.ly/3XXDAEg>
21. Organisation for Economic Co-operation and Development. Recommendation of the Council on Artificial Intelligence. OECD Legal Instruments [Internet]. 2019 [acesso 22 maio 2024]. Disponível: <https://bit.ly/4bKBb35>
22. Parlamento Europeu. Resolução do Parlamento Europeu de 16 de fevereiro de 2017, que contém recomendações à Comissão sobre disposições de Direito Civil sobre Robótica [Internet]. 2017 [acesso 22 maio 2024]. Disponível: <https://bit.ly/3S8tSen>
23. Brasil. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). Diário Oficial da União [Internet]. Brasília, nº 157, p. 59-64, 15 ago 2018 [acesso 22 maio 2024]. Seção 1. Disponível: <https://bit.ly/3VVwttm>
24. Brasil. Projeto de Lei nº 5.051, de 2019. Estabelece os princípios para o uso da inteligência artificial no Brasil. Senado Federal [Internet]. 2019 [acesso 22 maio 2024]. Disponível: <https://bit.ly/460iLdr>
25. Brasil. Projeto de Lei nº 21, de 2020. Estabelece princípios, direitos e deveres para uso de inteligência artificial no Brasil, e dá outras providências. Câmara dos Deputados [Internet]. 2020 [acesso 22 maio 2024]. Disponível: <https://bit.ly/4cHC39S>
26. Brasil. Projeto de Lei nº 872, de 12 de março de 2021. Dispõe sobre o uso da inteligência artificial. Senado Federal [Internet]. 2021 [acesso 22 maio 2024]. Disponível: <https://bit.ly/3zH4J4i>
27. Brasil. Projeto de Lei nº 759, de 2023. Regulamento os sistemas de inteligência artificial e dá outras providências. Câmara dos Deputados [Internet]. 2023 [acesso 22 maio 2024]. Disponível: <https://bit.ly/4d3PctP>
28. Kavoori AP. Review: Lisa Nakamura, Cybertypes: race, ethnicity and identity on the internet (New York&London: Routledge, 2002), 169pp. ISBN 0415938368/0415938376 (pbk). Converg Int J Res New Media Technol [Internet]; 2002 [acesso 22 maio 2024];9(1):123-4. DOI: 10.1177/135485650300900108
29. Brey P, Soraker JH. Philosophy of computing and information technology. In: Meijers A, editor. Handbook philosophy of technology and engineering sciences [Internet]. Amsterdam: Elsevier; 2009 [acesso 22 maio 2024]. p. 1341-407. DOI: 10.1016/B978-0-444-51667-1.50051-3
30. Beauty.AI [Internet]. [s.d.] [acesso 22 maio 2024]. Disponível: <https://beauty.ai>
31. Pearson J. Why an AI-judged beauty contest picked nearly all white winners. VICE [Internet]. 2016 [acesso 22 maio 2024]. Disponível: <https://tny.im/X4xZ0>
32. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science [Internet]. 2019 [acesso 22 maio 2024]; 366(6464):447-53. DOI: 10.1126/science.aax2342
33. Datta A, Tschantz MC, Datta A. Automated experiments on ad privacy settings: a tale of opacity, choice, and discrimination. Proc Priv Enhancing Technol [Internet]. 2015 [acesso 22 maio 2024]; 2015(1):92-112. DOI: 10.1515/popets-2015-0007
34. Sweeney L. Discrimination in online ad delivery. Commun ACM [Internet]. 2013 [acesso 22 maio 2024];56(5):44-54. DOI: 10.1145/2447976.2447990

35. Kay M, Matuszek C, Munson SA. Unequal representation and gender stereotypes in image search results for occupations. In: Begole B, Kim J, editors. CHI'15: proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems [Internet]. New York: Association for Computing Machinery; 2015 [acceso 22 maio 2024]. p. 3819-28. DOI: 10.1145/2702123.2702520
36. Klare BF, Burge MJ, Klontz JC, Vorder Bruegge RW, Jain AK. Face recognition performance: role of demographic information. IEEE Trans Inf Forensics Secur [Internet]. 2012 [acceso 22 maio 2024];7(6):1789-801. DOI: 10.1109/tifs.2012.2214212
37. Wexler J, Pushkarna M, Bolukbasi T, Wattenberg M, Viegas F, Wilson J. The what-if tool: interactive probing of machine learning models. arXiv:1907.04135 [Preprint]. 2019 [acceso 22 maio 2024]. DOI: 10.48550/arXiv.1907.04135
38. Bellamy RKE, Dey K, Hind M, Hoffman SC, Houde S, Kannan K *et al.* AI fairness 360: an extensible toolkit for detecting, understanding, and mitigating unwanted algorithmic bias. arXiv:1810.01943 [Preprint]. 2018 [acceso 22 maio 2024]. DOI: 10.48550/arXiv.1810.01943
39. Conselho Federal de Medicina. Código de Ética Médica. Resolução CFM nº 2.217, de 27 de setembro de 2018, modificada pelas Resoluções CFM nº 2.222/2018 e 2.226/2019 [Internet]. Brasília: CFM; 2019 [acceso 22 maio 2024]. Disponível: <https://tny.im/S9McE>
40. Brasil. Lei nº 3.268, de 30 de setembro de 1957. Dispõe sobre os Conselhos de Medicina, e dá outras providências. Diário Oficial da União [Internet]. Rio de Janeiro, p. 23013, 1 out 1957 [acceso 22 maio 2024]. Disponível: <https://tny.im/s1AW8>
41. Brasil. Lei nº 12.842, de 10 de julho de 2013. Dispõe sobre o exercício da medicina. Diário Oficial da União [Internet]. Brasília, p. 1º, 11 jul 2013 [acceso 22 maio 2024]. Disponível: <https://tny.im/DwLsM>
42. Dadalto L, Pimentel W. Responsabilidade civil do médico no uso da inteligência artificial. Revista Iberc [Internet]. 2019 [acceso 22 maio 2024];2(3):1-21. DOI: 10.37963/iberc.v2i3.86
43. Roberto E. Responsabilidade civil pelo uso de sistemas de inteligência artificial: em busca de um novo paradigma. Internet e Sociedade [Internet]. 2020 [acceso 22 maio 2024];1(1):121. Disponível: <https://tny.im/L46K>

Jhadson Silva Leonel – Especialista – [jhadsonqueiroz@hotmail.com](mailto:jhadsonqueiroz@hotmail.com)

 0009-0006-3009-0213

Camila Ferreira Silva Leonel – Magíster – [camilanutri.cf@gmail.com](mailto:camilanutri.cf@gmail.com)

 0000-0002-0535-4743

Jonas Byk – Doctor – [jonas.byk@hotmail.com](mailto:jonas.byk@hotmail.com)

 0000-0001-5854-4316

Silvania da Conceição Furtado – Doctora – [silvaniafurtado@ufam.edu.br](mailto:silvaniafurtado@ufam.edu.br)

 0000-0003-0065-3119

#### Correspondencia

Jhadson Silva Leonel – Rua Professor Andrade Bezerra, Rua 2, ap. 204, bloco H, Residencial Eliza Miranda CEP 69077-767. Manaus/AM, Brasil.

#### Participación de los autores

Jhadson Silva Leonel realizó la recopilación y el análisis de datos y participó en el diseño de la investigación y en la redacción del artículo. Camila Ferreira Silva Leonel colaboró en el formato, corrección y redacción del artículo. Jonas Byk participó en el diseño de la investigación y colaboró en el análisis de datos y en la redacción del artículo. Silvania da Conceição Furtado participó como orientadora de la asignatura para la redacción del artículo.

**Recibido:** 4.1.2024

**Revisado:** 13.6.2024

**Aprobado:** 11.7.2024