

# Inteligência artificial: desafios éticos e 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.

## Resumo

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.

## Resumen

### Inteligencia artificial: desafíos éticos y futuros

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.

## 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.

Declararam não haver conflito de interesse.

A inteligência artificial (IA) tem desempenhado papel cada vez mais importante nos últimos anos. Máquinas não se limitam mais a executar tarefas físicas, mas também desempenham funções intelectuais que exigem o que se considera inteligência. O início desse conceito remonta à década de 1950, quando o matemático britânico Alan Turing propôs o famoso Teste de Turing em seu artigo "Computing machinery and intelligence"<sup>1</sup>. Apesar disso, um marco decisivo ocorreu durante o verão de 1956, com a realização da Dartmouth College Conference, que é amplamente reconhecida como o início da IA como campo de estudo<sup>2</sup>.

Inicialmente, a IA foi aplicada principalmente na solução de problemas do mundo real por meio da programação do conhecimento de especialistas em programas de computador. Esses programas, chamados de sistemas especialistas ou sistemas baseados em conhecimento, foram desenvolvidos com base em entrevistas com especialistas em determinadas áreas. No entanto, havia limitações, como subjetividade e falta de cooperação dos especialistas<sup>3</sup>.

Atualmente é grande o entusiasmo em relação aos potenciais benefícios da IA, de forma que máquinas estão aprendendo a dirigir carros independentes e tradutores automáticos estão se tornando cada vez mais precisos. Além disso, a IA está presente em tarefas cotidianas como ler e-mails, lavar roupas e recomendar filmes em plataformas de *streaming*<sup>4</sup>.

Na área da saúde, a IA tem sido aplicada em várias pesquisas acadêmicas com resultados promissores. Algoritmos estão sendo desenvolvidos para interpretar imagens de radiografia<sup>5</sup>, mamografia<sup>6</sup>, tomografia computadorizada<sup>7</sup>, ressonância magnética<sup>8</sup> e identificar doenças como Alzheimer<sup>9</sup>. Além disso, tem sido utilizada na patologia<sup>10</sup> para identificação de lesões cancerígenas<sup>11</sup>, na interpretação de imagens da retina<sup>12</sup> e na detecção de arritmias<sup>13</sup> e outras condições médicas.

Robôs equipados com IA também estão sendo usados em hospitais e clínicas, como o robô criado pela empresa iFlytek, que passou no exame nacional para licenciamento de médicos na China<sup>14</sup>. Outro exemplo é o *software* brasileiro Laura, que identifica risco de infecção hospitalar por meio da IA<sup>15</sup>. Essas aplicações demonstraram o potencial da IA na área da saúde.

O rápido desenvolvimento de tecnologias para processamento e armazenamento de dados tem impulsionado o crescimento da IA. E, à medida que os problemas se tornam mais complexos e a quantidade de dados aumenta, é necessário desenvolver ferramentas computacionais avançadas e personalizadas, baseadas no aprendizado de máquina, que dependem cada vez menos da intervenção humana. No entanto, esse desenvolvimento vem acompanhado de preocupações, principalmente em relação à ética e ao impacto na sociedade, considerando questões legais e de responsabilidade, para garantir que a IA seja benéfica e justa.

Este artigo tem como objetivo realizar uma revisão narrativa da literatura com abordagem exploratória por meio de análise de artigos com informações acerca do uso da IA em serviços de saúde e suas implicações éticas.

## Aspectos éticos

Uma corrente tradicional da filosofia sustenta que a ética é exclusiva dos seres humanos, e que as coisas em si mesmas não são necessariamente boas ou más. *A tecnologia é amoral*, comentou Bill Gates em uma entrevista. Ele acrescentou que *cabe a nós refletir profundamente sobre as novas tecnologias e como devem ou não ser utilizadas*, conforme refere Garattoni<sup>16</sup>. No entanto, os desafios apresentados pela IA reacendem a questão da exclusividade humana na ética.

Desde o início, a interação entre seres humanos e máquinas tem gerado debates. Em 1950, o escritor e bioquímico norte-americano Isaac Asimov, em sua obra *Eu, robô*<sup>17</sup>, abordou a complexidade do tema e criou um clássico ao enunciar as três leis da robótica: 1) um robô não pode ferir um ser humano ou, por omissão, permitir que um ser humano sofra algum mal; 2) os robôs devem obedecer às ordens dos seres humanos, exceto quando tais ordens entrarem em conflito com a primeira lei; e 3) um robô deve proteger sua própria existência, desde que isso não entre em conflito com as leis anteriores. Mais tarde, ele acrescentou uma quarta lei, conhecida como Lei Zero<sup>17</sup>, que estabelece que um robô não pode causar mal à humanidade ou, por omissão, permitir que a humanidade sofra algum mal.

Recentemente, cinco princípios éticos se destacam nas principais iniciativas de regulamentação e nas discussões relacionadas a IA: beneficência, não maleficência, autonomia, justiça e explicabilidade<sup>18</sup>. Esses princípios éticos, definidos como valores ou diretrizes gerais que orientam o comportamento humano, servem como base para avaliar a moralidade das ações e fornecem uma estrutura geral para a tomada de decisões em questões éticas.

## Problemas relacionados à ética algorítmica

### Privacidade dos dados

Com o avanço da tecnologia, as pessoas estão cada vez mais conectadas e produzem uma quantidade crescente de dados, gerados de diversas fontes, como celulares, carros, transações com cartões de crédito, escolhas de conteúdo em serviços de *streaming* e dispositivos de internet. Essas informações são armazenadas por diversas empresas, que criam bancos de dados com informações dos usuários.

Devido à limitação humana, é praticamente impossível realizar uma análise abrangente desses dados, uma vez que isso demandaria muito tempo para uma avaliação completa e a formulação de conclusões estruturadas. No entanto, com a ascensão da IA e o aumento da capacidade de processamento, pode-se analisar esses dados de diversas formas em um curto espaço de tempo.

Além disso, uma preocupação cada vez mais pertinente nos dias de hoje é o destino das informações armazenadas em grandes bancos de dados. Grandes empresas, como a Meta (antigo Facebook), enfrentaram processos legais devido ao compartilhamento de dados sem o devido consentimento.

O vazamento de dados de milhões de usuários do Facebook, obtido por meio dos algoritmos da empresa Cambridge Analytica, permitiu o acesso a informações confidenciais sobre o comportamento de 50 milhões de usuários da rede social, que foram utilizadas para orientar estratégias de campanha. Esse caso é um exemplo notável da discussão sobre vigilância, privacidade e algoritmos<sup>19</sup>.

Diariamente, as pessoas são inundadas por anúncios imediatamente após realizar pesquisas

na internet, seja por interesses de compra, seja por mera curiosidade. Dados relacionados a saúde, alimentação, hábitos de exercícios físicos e até mesmo locais frequentados por usuários de *smartphones* e outros aparelhos podem, graças à capacidade avançada de processamento, revelar tendências e informações sobre saúde que antes eram desconhecidas. No entanto, diferentes empresas podem ter tais informações, e frequentemente as pessoas desconhecem como e por quem elas serão utilizadas.

Devido à crescente preocupação com a segurança dos dados, entidades como a Organização Mundial da Saúde (OMS) elaboraram um relatório global em 2021 contendo seis princípios essenciais para assegurar que a IA funcione em prol do interesse público em todos os países<sup>20</sup>. Além disso, a Organização para a Cooperação e Desenvolvimento Econômico (OCDE) produziu um conjunto de diretrizes intergovernamentais para orientar os Estados na regulamentação da IA, mantendo sua inovação, confiabilidade, segurança e respeito pelos direitos humanos e valores democráticos.

Trinta e oito países-membros aderiram à Recomendações do Conselho da OCDE sobre IA, documento que estabelece princípios a serem seguidos internamente<sup>21</sup>. Vários outros países também estão desenvolvendo legislações próprias para garantir maior segurança e transparência no uso de informações privadas por empresas.

Na União Europeia, por exemplo, foi promulgada a Resolução do Parlamento Europeu de 16 de fevereiro de 2017, que contém recomendações sobre disposições de direito civil relacionadas à robótica. A Resolução busca estabelecer regras coordenadas entre os Estados-membros para garantir a segurança no desenvolvimento da IA, bem como a proteção dos direitos fundamentais das pessoas e das empresas, sem inibir o investimento, a inovação e a utilização<sup>22</sup>.

No Brasil, a Lei 13.709/2018<sup>23</sup>, conhecida como Lei Geral de Proteção de Dados (LGPD), foi criada com o objetivo de proteger os direitos fundamentais à liberdade e, mais especificamente, à privacidade. Além disso, o Senado Federal instituiu a Comissão de Juristas Responsável por Subsidiar Elaboração de Substitutivo sobre Inteligência Artificial no Brasil (CJSUBIA), que está trabalhando na elaboração da regulamentação legal brasileira relacionada ao uso de IA.

Há também alguns projetos de lei (PL) em andamento para regulamentar a IA, a saber: o PL 5051/2019<sup>24</sup>, o PL 21/2020<sup>25</sup>, o PL 872/2021<sup>26</sup> e o PL 759/2023<sup>27</sup>. A importância desse tema é evidenciada pela quantidade de PL em andamento, destacando a necessidade de regulamentação legal nessa área em crescimento.

### Viés na inteligência artificial

O uso crescente da IA acarreta grande preocupação com o viés que pode estar presente ou ser amplificado por sistemas que usam essa tecnologia. Para compreender esse fenômeno, é fundamental analisar o processo de aprendizado das máquinas, que pode se dar por meio de um passo a passo predefinido, conhecido como algoritmo, algo semelhante a estudo realizado com livros ou orientação de professores.

Além disso, há a abordagem do aprendizado de máquina, que não ensina diretamente, mas fornece instruções sobre como aprender com base em exemplos e dados, conferindo autonomia ao processo de aprendizado. Nesse contexto, os dados desempenham papel fundamental, e, quanto mais dados confiáveis estiverem disponíveis para o treinamento, melhor será o modelo resultante.

Entretanto, algoritmos não estão isentos de valores e pressupostos, pois os parâmetros operacionais são estabelecidos por desenvolvedores e configurados por usuários que buscam resultados específicos, o que pode priorizar alguns valores e interesses em detrimento de outros. Mesmo dentro dos parâmetros estabelecidos, não se garante automaticamente um comportamento eticamente aceitável<sup>28,29</sup>.

Entre diversos exemplos notórios de viés, pode-se citar um concurso de beleza realizado em 2016, que utilizou um júri composto por robôs gerados por IA e teve inscrição de milhares de pessoas de vários países, que enviaram suas fotos para avaliação<sup>30</sup>. O resultado chamou a atenção, pois, dos 44 vencedores, apenas um tinha pele escura, o que gerou um debate acalorado sobre como o sistema poderia perpetuar vieses, produzindo resultados não intencionais, por vezes distorcidos e até mesmo racistas.

A análise revelou que a base de treinamento do sistema era composta por imagens de atores e atrizes de Hollywood, que predominantemente eram

pessoas brancas. Não houve intenção dos desenvolvedores de favorecer um grupo étnico, mas se constatou negligência na verificação de vieses na base de treinamento<sup>31</sup>.

Um exemplo relevante de viés relacionado à saúde ocorreu com um programa usado por uma grande seguradora americana. O objetivo era identificar usuários que necessitavam de cuidados adicionais, permitindo que uma equipe multidisciplinar os acompanhasse preventivamente para evitar tratamentos caros, como internações em unidades de terapia intensiva (UTI). O sistema se baseou na frequência com que os usuários buscavam atendimento médico e nos custos de consultas e internações; para evitar viés racial, os dados não continham informações sobre a raça dos usuários<sup>32</sup>.

No entanto, a análise revelou que a maioria dos usuários identificados como necessitados de cuidados adicionais eram brancos. A investigação posterior revelou que os usuários negros, devido a restrições financeiras, usavam menos o plano de saúde, resultando em menos dados disponíveis para o sistema identificar padrões de gravidade em seu atendimento. Isso destaca como os vieses e preconceitos podem se ocultar nos dados, tornando sua detecção desafiadora<sup>32</sup>.

Outros exemplos de viés nos resultados da IA incluem anúncios de emprego para posições altamente remuneradas, que apresentam menor probabilidade de ser apresentados a mulheres<sup>33</sup>; pesquisas de nomes com sonoridade distintamente negra, que têm maior probabilidade de desencadear anúncios relacionados a registros de prisões<sup>34</sup>; e pesquisas de imagens para profissões como CEO, que produzem menos imagens de mulheres<sup>35</sup>. Além disso, sistemas de reconhecimento facial, amplamente usados na aplicação da lei, têm desempenho inferior no reconhecimento de rostos de mulheres de qualquer cor e de indivíduos negros<sup>36</sup>.

Eliminar vieses e garantir a imparcialidade em sistemas baseados em IA tem se tornado um desafio cada vez maior, sendo crucial reconhecer que os dados não são neutros, pois refletem decisões humanas que podem ser preconceituosas; portanto, bancos de dados de treinamento devem ser confiáveis e isentos de viés. Empresas como Google e IBM estão se esforçando para desenvolver métodos técnicos que possibilitem a avaliação e a mitigação de vieses em seus conjuntos de

dados. Além disso, órgãos reguladores e governos desenvolvem políticas e regulamentações específicas para lidar com o viés na IA, reconhecendo a necessidade de transparência e responsabilidade nesse campo em constante evolução<sup>37,38</sup>.

### Responsabilidade e transparência na inteligência artificial

O atual avanço tecnológico está desempenhando papel fundamental na área da medicina, com a crescente utilização de equipamentos inteligentes baseados em sistemas de IA. Essas inovações são cruciais na medicina diagnóstica, de forma que a interpretação de resultados, tradicionalmente realizada por médicos, agora conta com o auxílio de algoritmos altamente especializados e automatizados.

Conforme estabelecido no Código de Ética Médica (CEM), é vedado ao médico *deixar de usar todos os meios disponíveis de promoção de saúde e de prevenção, diagnóstico e tratamento de doenças, cientificamente reconhecidos e a seu alcance, em favor do paciente*<sup>39</sup>. Isso ressalta que a negligência no uso dos recursos disponíveis constitui violação clara do CEM. No entanto, a falta de transparência e a complexidade dos algoritmos podem dificultar a compreensão dos profissionais de saúde em relação às conclusões do sistema.

Isso pode levar os médicos a hesitarem em adotar sistemas de IA que não compreendem completamente, o que prejudica a integração eficaz na prática clínica. Além disso, a falta de compreensão pode aumentar o risco de erros, uma vez que os profissionais não conseguem identificar decisões incorretas, o que torna complexa a atribuição de responsabilidade em casos de erros médicos, gerando desafios legais e éticos.

De acordo com a Lei 3.268/1957<sup>40</sup>, que regulamenta os conselhos de medicina, e com a Lei 12.842/2013<sup>41</sup>, que reafirma a anterior e é conhecida como Lei do Ato Médico, o diagnóstico de doenças é prerrogativa exclusiva de médicos devidamente habilitados e registrados nos Conselhos Regionais de Medicina (CRM). Portanto, médicos não podem se eximir da responsabilidade pelo diagnóstico de doenças, uma vez que isso é parte fundamental de sua profissão.

A atribuição de responsabilidade a médicos diante da complexidade dos sistemas de IA pode

resultar em distribuição injusta de responsabilidades. Isso ocorre devido à realidade complexa que envolve a responsabilização médica em relação aos parâmetros tradicionalmente utilizados<sup>42,43</sup>.

Juristas argumentam que a análise deve considerar se o dano causado pelo sistema de IA foi resultado de negligência, imperícia ou má conduta por parte do usuário, e que médicos não devem ser responsabilizados por danos causados por agentes autônomos. Isso se deve ao fato de que sistemas inteligentes têm autonomia e acesso a dados, o que os torna capazes de tomar decisões independentes<sup>42,43</sup>.

Essa complexidade envolvendo a identificação e atribuição de responsabilidades por danos decorrentes de atividades algorítmicas é um desafio multissetorial. Assim, detectar o dano, determinar sua causa e, principalmente, identificar quem deve ser responsabilizado por tais danos são tarefas complexas.

Diversos atores humanos estão envolvidos, incluindo programadores, projetistas, fabricantes, *designers* e usuários, e a questão ética da responsabilidade está intimamente ligada ao tema da transparência. Somente algoritmos não opacos permitem identificar os responsáveis pelo processo e os possíveis erros. Nesse contexto, é essencial um esforço contínuo para garantir que sistemas de IA na medicina sejam transparentes, compreensíveis e sujeitos a regulamentações que considerem os desafios éticos e legais que surgem com a evolução tecnológica, a fim de que continuem a desempenhar papel cada vez mais significativo.

### Considerações finais

O avanço da IA transformou o mundo, permitindo que máquinas desempenhem tarefas complexas e racionais, anteriormente consideradas privilégios humanos. Embora tenha suas raízes nas décadas de 1950 e 1960, seu impacto se tornou especialmente notável nas últimas décadas, à medida que avanços tecnológicos possibilitaram o desenvolvimento de sistemas cada vez mais sofisticados.

Hoje, essa tecnologia permeia o cotidiano e envolve desde a indústria automobilística até o diagnóstico médico de precisão. Assim, hospitais

e clínicas contam com sistemas autônomos para melhorar a saúde e o bem-estar dos pacientes.

No entanto, o rápido crescimento não vem isento de desafios, pois o aumento da complexidade dos problemas e a quantidade massiva de dados gerados exigem que as ferramentas computacionais se tornem mais autônomas. Nesse cenário, o aprendizado de máquina tem papel central, permitindo que sistemas adquiram conhecimento e tomem decisões com mínima intervenção humana.

Embora prometa benefícios significativos, a tecnologia também levanta preocupações éticas, como privacidade, viés algorítmico e responsabilidade, de maneira que, à medida que a IA avança, a sociedade precisa promover discussões e regulamentações sensatas que assegurem seu uso ético e justo. Os benefícios são inegáveis, mas seu potencial só será plenamente realizado se os sujeitos envolvidos estiverem preparados para enfrentar os desafios e adotar abordagens responsáveis.

## Referências

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 [acesso 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 [acesso 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 [acesso 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 [acesso 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 [acesso 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 [acesso 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 [acesso 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 [acesso 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 [acesso 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 [acesso 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 [acesso 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/3VVwtm>
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 [acesso 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 [acesso 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 [acesso 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 [acesso 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 [acesso 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 [acesso 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 [acesso 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 [acesso 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 [acesso 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 – Mestre – [camilanutri.cf@gmail.com](mailto:camilanutri.cf@gmail.com)

 0000-0002-0535-4743

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

 0000-0001-5854-4316

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

 0000-0003-0065-3119

#### Correspondência

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

#### Participação dos autores

Jhadson Silva Leonel realizou a coleta e a análise de dados e participou da concepção da pesquisa e da redação do artigo. Camila Ferreira Silva Leonel colaborou na formatação, correção e redação do artigo. Jonas Byk participou da concepção da pesquisa e colaborou na análise de dados e na redação do artigo. Silvania da Conceição Furtado participou como orientadora na disciplina para escrita do artigo.

Recebido: 4.1.2024

Revisado: 13.6.2024

Aprovado: 11.7.2024