

Bioethics and artificial intelligence: a current overview of the literature

Érica Antunes Naves

UnitedHealth Group Brasil, São Paulo/SP, Brasil.

Abstract

Artificial intelligence refers to the performance, by computer devices, of intellectual processes characteristic of human beings, such as reasoning, discovering meanings, generalizing or learning from experience. Artificial intelligence occurs when computer programs perform action for which they were not explicitly programmed. Although a well-defined concept, its complex performance poses various bioethical conflicts and questions, often clarified only when they emerge. Despite the regulations put in place during the field's development, these are constantly undergoing adaptations thus justifying further studies on the subject.

Keywords: Bioethics. Artificial intelligence. Technology control, biomedical.

Resumo

Bioética e inteligência artificial: panorama atual da literatura

O termo inteligência artificial refere-se à realização, por dispositivos computacionais, de processos intelectuais característicos dos seres humanos, como raciocinar, descobrir significados, generalizar ou aprender com experiências. A atuação da inteligência artificial ocorre quando programas computacionais realizam ações para as quais não foram explicitamente programados. Apesar de o conceito ser bem definido, o desempenho dessa tecnologia é muito complexo, de modo que a bioética encontra diversos conflitos e questões relacionadas a ela, muitas vezes esclarecidas apenas no momento em que surgem. Embora regulamentações tenham sido instituídas ao longo do desenvolvimento da área, ela constantemente passa por adaptações, o que justifica novos estudos sobre o tema.

Palavras-chave: Bioética. Inteligência artificial. Controle da tecnologia biomédica.

Resumen

Bioética e inteligencia artificial: panorama actual de la literatura

El término inteligencia artificial se refiere a sistemas informáticos capaces de realizar procesos intelectuales característicos de los seres humanos, como razonar, descubrir significados, generalizar o aprender de las experiencias. La actuación de la inteligencia artificial se produce cuando los programas informáticos realizan acciones para las cuales no fueron explícitamente programados. Aunque el concepto está bien definido, la actuación de esta tecnología es muy compleja, por lo que la bioética se encuentra ante diversos conflictos y cuestiones relacionadas con ella, que muchas veces solo pueden aclararse cuando surgen. Aunque a lo largo de su desarrollo se vienen estableciendo normativas, este campo sufre constantes adaptaciones, lo que justifica la realización de nuevos estudios sobre el tema.

Palabras clave: Bioética. Inteligencia artificial. Control de la tecnología biomédica.

The author declares no conflict of interest.

The term artificial intelligence (AI) refers to the performance by computing devices of typical human intellectual processes, such as reasoning, discovering meaning, generalizing or learning from experience¹. This occurs when computer programs carry out actions for which they were not explicitly programmed, and can be described as the use of such devices to perform tasks that previously required human cognition².

The American Medical Association prefers the term “augmented intelligence” to “artificial intelligence” to emphasize the role of computers in furthering the improvement of medical skills rather than replacing them³. The integration of AI into medical practice has grown significantly in recent years and continues to grow². Therefore, the discussion on ethical and legal principles in this field, including the development of specific regulation, is a constant and extremely relevant concern.

Method

A search was carried out in the PubMed database using the descriptors “bioethics” and “artificial intelligence,” considering scientific papers published between 2018 and 2022. The analysis also included opinions from leading societies in the field. This was followed by a critical comparison based on different criteria according to the type of study, discussion of the subject and conclusions.

Discussion

All the works analyzed herein define the concept of AI and agree that it is a technology which emulates intellectual processes typical of humans to achieve goals without programming a specific action¹⁻⁴.

There is broad agreement that any device showing intelligence develops reasoning, devises strategies, solves puzzles, makes judgments under uncertainty, demonstrates knowledge (including common sense), plans, learns, communicates in natural language and combines all these skills in the pursuit of common goals¹.

Currently, examples of areas of AI activity include understanding human speech, competing at the highest level in strategic gaming systems, driving autonomous cars, planning intelligent routing in content delivery networks and military simulations¹.

In the medical field, Bali and collaborators¹ present examples such as IBM’s Watson for Oncology, which prescribed medication for the treatment of cancer patients with equal or higher efficiency than human specialists. The Hanover project, developed by Microsoft in Oregon, USA, analyzed medical research to enable personalized cancer treatment.

Moreover, the United Kingdom’s National Health Service (NHS) uses Google’s DeepMind platform to detect health risks by analyzing data from mobile applications and medical imagery collected from patients. Another example is the Stanford radiology algorithm, which detected pneumonia better than human radiologists while matching expert ophthalmologists in making referrals for patients with diabetic retinopathy¹.

Vearrier and collaborators² agree that AI has shown substantial potential benefits for both physicians and patients, transforming therapeutic relationships from the traditional physician-patient dyad into a triad involving physician, patient and machine. Nevertheless, they believe that new AI technologies require careful supervision, legal standards, patient safeguards and provider education, with doctors acknowledging the limits and risks of AI alongside its potential benefits.

With so many AI algorithms in use and under development, in all fields and especially in medicine, there is a growing concern to debate the ethical and legal aspects involved, including regulation. In the past, such regulatory changes were based on philosophical concepts and the need to respond to sociocultural demands. Currently, the challenge is to integrate a number of disruptive technologies, such as AI.

Although it provides numerous advantages for involving an emerging reality and unpredictable results, which are not risk-free, this field generates uncertainty and precaution, requiring bioethical regulation⁴. Sánchez López and collaborators⁴

propose three basic principles to regulate stakeholders: 1) respect for investigation; 2) justice; and 3) transparency.

Despite their promise of delivering results, AI algorithms are under investigation for inconsistent performance, especially in minority communities, which can lead to suboptimal clinical decisions and adverse patient outcomes. To date, such concerns have focused on medical negligence and stress that the responsibility of physicians for AI use is inextricably linked to that of those other actors⁵.

In an in-depth study of the liability of this issue, Mahila and collaborators⁵ reflect that the assignment of accountability determines whether patients obtain compensation, and from whom. Additionally, they extend the assessment by questioning whether potentially useful algorithms will be put into practice, as increasing accountability for the use or development of algorithms may discourage health system developers and leaders from introducing them into clinical practice. They stress that the larger AI accountability ecosystem and its role in ensuring safe execution and innovation in clinical care must be closely examined.

In an editorial, Sher, Sharp and Wright⁶ are optimistic about the future prospects of AI for health care and consider the importance of understanding strengths, limitations, opportunities, ethical challenges and risks. Equally important is the need to become familiar with the listed tools to critically analyze AI applications in health care, in order to distinguish between concrete improvements and mistakes.

Lazarus and collaborators⁷ expand the analysis of AI limitations for health education, specifically anatomy, describing strains between the promises and dangers of integrating this technology into this field. They end by providing practical recommendations for a well-considered approach when working together with AI, which serve as a guiding framework aimed at developing a more nuanced and balanced approach to the role of AI in health education.

The scientific and ethical aspects of this issue justified the organization of a meeting to

discuss the matter, in March 2017, in Barcelona, attended by different European experts in AI, computing and communication, among other areas. The debate resulted in the *Barcelona Declaration for the Proper Development and Usage of Artificial Intelligence in Europe*⁸, which contains the following principles and values: prudence, reliability, accountability, responsibility, constrained autonomy and human role, detailed below.

- Prudence: The leap forward in this area has been caused by a maturation of AI technologies, increased computing power and data storage capacity, the availability of internet-based delivery platforms and a greater willingness of many economic actors to experiment with the technology on their own;
- Reliability: All artificial systems used in society have to undergo tests to determine their reliability and security;
- Accountability: when an AI system makes a decision, people affected by those decisions should be able to get an explanation why the decision is made in terms they can understand and they should be able to challenge the decision with reasoned arguments;
- Responsibility: There is growing worry about AI chat-bots and other types of automated messaging systems operating on the internet and in social media which are designed for the manipulation of political opinion, disinformation by spreading false facts, extortion or other forms of malicious activity that are dangerous to individuals and destabilizing to society;
- Constrained autonomy: AI systems have not only the capacity to make decisions. When they are embedded in physical systems, such as self-driving cars, they have the potential to act on their decisions in the real world, which raises issues of safety and the possibility of autonomous AI overtaking human intelligence at some point;
- Human role: The current undeniable enthusiasm for AI sometimes gives the impression that human intelligence will no longer be needed, leading some companies to fire employees and replace them with AI systems. This is a serious

mistake, as all AI systems critically depend on human intelligence.

Tai⁹ points out that despite all the positive promises offered by AI, human experts are still essential and necessary to design, program and operate AI in order to avoid unpredictable errors. In this sense, he cites Beth Kindig, a technology analyst from San Francisco with more than a decade of experience in public and private technology companies.

Kindig published a free newsletter indicating that while AI holds potential promise for better medical diagnosis, human experts are still needed to resolve impasses and mitigate errors. Therefore, surveillance of AI tasks, carried out by the healthcare professional known as physician, cannot be overlooked⁹.

Final considerations


Like all emerging technologies, AI and robotics require the prior implementation of ethical operating standards that guarantee security and applicability in a field as sensitive as health care³. AI systems have the potential to radically transform clinical care and, even if moving at a slower pace, the legal system cannot remain stagnant regarding this innovation.

To fully realize the benefits of AI, the legal system must balance responsibility to promote innovation, safety and accelerated adoption of these algorithms⁵. The relatively unstable state of AI and its potential accountability provide an opportunity to develop a new model that accommodates medical progress and guides stakeholders on how best to respond to this disruptive innovation.

References

1. Bali J, Garg R, Bali TR. Artificial intelligence (AI) in healthcare and biomedical research: why a strong computational/AI bioethics framework is required? What is artificial intelligence? *Indian J Ophthalmol* [Internet]. 2018 [acesso 29 nov 2023];67(1):9-12. DOI: 10.4103/ijo.IJO_1292_18
2. Vearrier L, Derse AR, Basford JB, Larkin GL, Moskop JC. Artificial Intelligence in emergency medicine: benefits, risks, and recommendations. *J Emerg Med* [Internet]. 2022 [acesso 29 nov 2023];62(4):492-9. DOI: 10.1016/j.jemermed.2022.01.001
3. Crigger E, Khoury C. Making policy on augmented intelligence in health care. *AMA J Ethics* [Internet]. 2019 [acesso 29 nov 2023];21(2):E188-91. DOI: 10.1001/amajethics.2019.188.
4. Sánchez López JD, Cambil Martín J, Villegas Calvo M, Luque Martínez F. Artificial intelligence and robotics: reflections about the need of a new bioethics framework implementation. *J Healthc Qual Res* [Internet]. 2021 [acesso 29 nov 2023];36(2):113-4. DOI: 10.1016/j.jhqr.2019.07.009
5. Maliha G, Gerke S, Cohen IG, Parikh RB. Artificial intelligence and liability in medicine: balancing safety and innovation. *Milbank Q* [Internet]. 2021 [acesso 29 nov 2023];99(3):629-47. DOI: 10.1111/1468-0009.12504
6. Sher T, Sharp R, Wright RS. Algorithms and bioethics. *Mayo Clin Proc* [Internet]. 2020 [acesso 29 nov 2023];95(5):843-4. DOI: 10.1016/j.mayocp.2020.03.020
7. Lazarus MD, Truong M, Douglas P, Selwyn N. Artificial intelligence and clinical anatomical education: promises and perils. *Anat Sci Educ* [Internet]. 2022 [acesso 29 nov 2023]. DOI: 10.1002/ase.2221
8. Barcelona Declaration for the Proper Development and Usage of Artificial Intelligence in Europe. [Propõe diretrizes para um “código de conduta” dos profissionais de IA atualmente em discussão]. *Biocat* [Internet]. 2017 [acesso 29 nov 2023]. Disponível: <https://bit.ly/3w6SYCv>
9. Tai MCT. The impact of artificial intelligence on human society and bioethics. *Tzu Chi Med J* [Internet]. 2020 [acesso 29 nov 2023];32(4):339-43. DOI: 10.4103/tcmj.tcmj_71_20

Érica Antunes Naves – Graduate (specialist) – eanaves@hotmail.com

 0000-0003-3257-4941

Correspondence

Rua Conselheiro Brotero, 1486, Santa Cecília CEP 01232-010. São Paulo/SP, Brasil.

Received: 4.18.2023

Revised: 11.30.2023

Approved: 1.5.2024