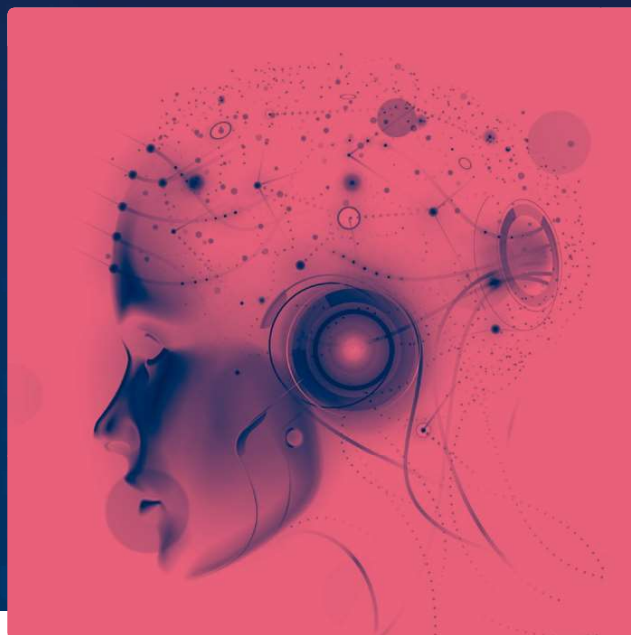


Ethical challenges of AI in the health sector

ROBERT SROKA



- **Bioethics was a response to the technological progress, which came ahead of legal regulations and often violated ethical norms, but opened the door to better solutions that saved life and health**
- **AI in the health sector holds great promise but also faces ethical challenges which call for a combination of expertise in AI ethics, medical ethics, and bioethics in order to be solved**

We can already see artificial intelligence systems play an increasingly important role in the health sector, in terms of system-wide operations, medical facility management, and patient treatment alike. AI can surely bring a range of benefits to the health sector, but whether it becomes socially acceptable will depend on both its technical effectiveness and 'ethicality'. This is why the issue of the ethics of artificial intelligence systems in the health sector is so important. At the same time, it is extremely challenging. First, because it directly affects the fundamental values of human life and health. Second, it requires an interdisciplinary combination of at least three ethical areas – artificial intelligence ethics, medical ethics, and bioethics. The latter seems to be particularly relevant here, not only because of the object of research, but mainly because of the method of reconciling ethical positions worked out over the last decades. In many cases, due to the very nature of artificial intelligence, it is reasonable to make use of procedures utilized in bioethics to evaluate the ethicality of AI.

A short article does not offer enough room to discuss in detail the enormity of the ethical challenges underlying the use of AI systems in the health sector. Therefore, I will only briefly mention some selected aspects that are important and worth further analysis. I will start from discussing the main ethical issues concerning AI and the most important documents prepared by international and national institutions that set the framework for the ethics of artificial intelligence. Then, I will indicate the key areas of medical ethics and bioethics that should be borne in mind when dealing with ethical dilemmas related to the use of AI in the health sector.

SELECTED ETHICAL CHALLENGES

The intensified development of AI research over the past decade or so has resulted in solutions that are already significantly affecting what we think and how we build social relationships. In many cases, it is an influence that we are not aware of, if only in terms of making purchasing decisions,

i.e. mainly in business, but also in education, science, culture, and means of communication – with particular emphasis on social media. This influence is often so strong that it touches upon the core ethical values of human rights founded on the category of dignity. To better understand the ethical challenges of AI in the health sector, it is reasonable to look at the most important ethical issues related to the utilization of AI systems in general.

One of the most widely used methods in AI systems considered in broad terms is machine learning. Its variation, supervised machine learning, which involves refining statistical models based on large data sets, includes tasks and solutions to those tasks. Based on training data, a statistical model “learns” the relationships between tasks and solutions. These techniques, particularly linear regressions, have long been used in medicine. A certain novelty in machine learning is the use of neural networks, which make it possible to e.g. create new solutions based on existing data. The problem is that in order to train a neural network, you need a large number of examples with a large variety that would make it possible to detect the existing dependencies. If this condition is not met, erroneous results are produced. AI systems therefore depend on the quality of the input data. Their reliability, in turn, depends on whether we know on what basis the trained models that use new data make decisions. If a model has learned to recognize erroneous relationships using training data, it will make errors that can lead to exposing patients to various risks in medicine when using new data. An example of such discriminatory error was an AI system used in U.S. hospitals. It was to qualify patients for specialized treatment programs, taking into account the annual cost of treating an individual. In deciding to make use of this data set, the system’s designers failed to consider the fact that in the United States, the spending on the treatment of black people is statistically lower than in the case of white people suffering equally from the same disease. In consequence, the system replicated the pattern resulting from the existing economic inequality by classifying black people as “less sick” than they actually were¹. Thus, the very quality of the data and the reliability of the results are ethical categories. It also turns out that trained models are often wrong in their assessment of real cases. Therefore, the accuracy of the decisions made is another ethical category – not just a praxeological or economic one. To summarize this one of the most important issues in AI ethics, I will use the following analogy: just as the brain needs oxygen, artificial intelligence needs data. If the air mixture is poisoned or inappropriate, it poisons the brain or causes delusions leading to a distorted perception of the surrounding world. The same is true for AI systems: poor quality data distorts the results of artificial intelligence systems. When such system are widely used in the health sector, this can lead to negative effects on a large scale.

Another broad spectrum of ethical challenges occurring particularly in the health sector is the matter of data acqui-

sition and the related issues of privacy, data security, and how data are used.

CODES OF ETHICS FOR AI SYSTEMS

Due to the rapid development, the profound impact on many different areas of life, and the wide reach of artificial intelligence systems, individual countries and international organizations have committed themselves to develop ethical principles for AI.

→ **It is important to notice that a race of sorts is underway to impose an ethical narrative for artificial intelligence. This is done by countries such as: Canada, France, Australia, the UK, Japan, China, but also India. International organizations are also involved. The efforts made in this respect by the OECD, UNESCO, and the European Union deserve particular attention.**

In 2019, the Organisation for Economic Co-operation and Development (OECD) published the first intergovernmental standards for the development of artificial intelligence – “*Recommendation of the Council on Artificial Intelligence*”. In the paper, the OECD stresses that artificial intelligence should benefit people and the planet by driving inclusive growth, sustainable development, and well-being. The OECD notes that AI systems should be designed to take into account the rule of law, human rights, democratic values, and diversity. They should also include adequate safeguards to allow human intervention where necessary to ensure a just society. The paper also addresses the issue of accountability by indicating, in line with the above principles, that organizations and individuals developing, implementing, or operating artificial intelligence systems should be held accountable for the proper functioning of such systems.

UNESCO’s work on AI ethics has a global impact too. The ethical proposal of this organization is based on values such as the dignity of each person, human rights, and fundamental freedoms. Its document emphasizes the basic principle of no harm – be it physical, economic, social, political, or psychological harm – in any phase of the “life cycle” of artificial intelligence. When discussing the application of AI in the health sector, it seems to be reasonable to mention that in promoting and developing the ethical principles for AI, UNESCO works closely with the International Bioethics Committee and the Intergovernmental Bioethics Committee.

Another important initiative addressing AI ethics is the Ethics Guidelines for Trustworthy AI, published by the European Commission and prepared by the High-Level Expert Group on Artificial Intelligence. The authors of the document were primarily guided by an approach in which the human being stands in the center of ethical reflection (*Human-Centric AI*). According to this proposition, artificial intelligence should

first and foremost be lawful, ethical (meaning consistent with ethical principles and values), and robust (meaning reliable, socially and technically 'durable').

Meanwhile, in April 2021, a Proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence was presented. The document is the result of legislative work aimed at maintaining the EU's technological leadership while ensuring that Europeans can benefit from new technologies developed and operated in accordance with the European values, fundamental rights, and essential principles. The purpose of this act is to propose an ecosystem for building trustworthy artificial intelligence. An important thing is that enhancing human well-being has been identified as the ultimate goal for artificial intelligence. The main ethical framework for AI has been set: safety, legal compliance, and compliance with respect for fundamental rights and EU values. Among the set of fundamental rights, emphasis is put on ensuring human dignity, respect for private life, and protection of personal data, non-discrimination, and equality between men and women. These regulations aim also to prevent a chilling effect on the rights to freedom of expression and freedom of assembly. Therefore, a ban on some particularly harmful practices using AI that go against EU values has been proposed. A methodology for evaluating artificial intelligence systems is also in place to identify those that involve high risk and pose a significant threat to human health and safety or fundamental rights. Special evaluation and surveillance procedures have been provided for this category of AI systems.

At the end of 2020, Poland adopted its own policy for the development of artificial intelligence, which repeatedly stresses the importance of ethics. The document makes it clear that it is crucial for the AI solutions created to always serve the human being, putting human dignity and rights first. It also emphasizes Poland's ambition to be an active participant in the global debate on the ethics of artificial intelligence.

AI ETHICS IN THE HEALTH SECTOR

The ethical reflection on the application of artificial intelligence systems in the medical sector cannot be limited to the current state of knowledge on the ethics of AI. It should take into account the perspective of medical ethics and bioethics. Developers of artificial intelligence systems for the health sector should take into account the considerations that shape the ethical framework of the physician's standard practice, which include principles such as patient welfare, patient autonomy, and social justice. The principle of utmost priority of patient welfare must not be overridden by business considerations, administrative requirements, or the use of unproven or non-transparent technologies. Also, the principle of patient autonomy, i.e. the possibility to make conscious decisions regarding treatment, may be disturbed

by new technologies in medicine, based on artificial intelligence systems whose creators did not foresee the possibility of materialization of the principle of patient autonomy at the stage of development. The third important principle in medical ethics is the principle of social justice, which refers to the fair distribution of resources and the elimination of discrimination in health care. As shown by the many examples of artificial intelligence systems already analyzed, the problem of AI algorithm bias (*algorithmic bias*) is one of the most serious challenges to face, which may lead to dangerous cases of discrimination. Another important challenge is the problem of observing professional secrecy, which, with artificial intelligence systems that "thrive" on data, is certainly a big issue.

AI is yet another important development in the technology that has spawned a new field of applied ethics. This field is bioethics, understood as formulating judgments and norms concerning biomedical interventions connected with the beginning of human life, its duration, and death². The interest in bioethics stems from the need to reflect on life and death in the conditions of modern technical civilization, as well as the need to adopt rules that enable us to make decisions in extreme and disputable situations. The discussions and disputes revolving around bioethics address the problem of the relationship between the progress of civilization and morality. So far, the advances in science and technology have made it ahead the existing legal norms and violated moral standards. In the past, however, the disparity between practice and moral principles did not have the far-reaching consequences that we can see today. Applying the existing technical capabilities beyond ethical control raises new problems that once did not exist and that need to be resolved – especially in the field of medicine³. This is why the subject matter of ethics of artificial intelligence systems is so important and, in my opinion, should draw on the methodological achievements of bioethics. Moreover, when it comes to medical grounds, the ethics of artificial intelligence will have to face a major bioethical dispute that is still ongoing – and where the two main contenders are the value of life and the quality of life.

When thinking about the development of AI ethics in the health sector, there are several elements that should be kept in mind:

- AI systems will be as ethical as their creators are. Therefore, when it comes to the domain of ethics of artificial intelligence systems in the health sector, there should be a special focus on the developers of these systems.
- Anticipating consequences and reacting quickly are highly important aspects of AI ethics. We need new institutional mechanisms for ethical evaluation of AI systems, which should be: interdisciplinary, responsive, systematic, and causative. In the case of the health sector, it is reasonable to draw on the experience of bioethics committees whose main objective is to protect human dignity.

- As for AI ethics, we should be especially mindful of the difference between merely feeling that something is ethical and whether it actually is.
- In the domain of AI, like in the sphere of medicine, prevention is better than cure, which is why it will be so important to make sure that the new forms of AI are trained in so-called *sandboxes*, i.e. tested in isolation, before being applied in practice in the health sector.

ROBERT SROKA

References:

1. Obern A., Szymielewicz K., *Sztuczna inteligencja non-fiction*, Panoptikon Foundation, <https://panoptikon.org/sztuczna-inteligencja-non-fiction>
2. Chyrowicz B., *Etyka stosowana. Metody i problemy*, Towarzystwo Naukowe KUL, p. 45.
3. Ślęczek-Czakon D., *Problem wartości i jakości życia w sporach bioetycznych*, Wydawnictwo Uniwersytetu Śląskiego, Katowice 2004, p. 10.



Robert Sroka, PhD – ESG Director and Chief Ethics Officer responsible for developing and overseeing a comprehensive responsible investment policy, ESG areas and corporate governance at Abris Capital Partners. Member of the Department of Entrepreneurship and Ethics in Business at Kozminski University. He deals with business ethics, ethics of new technologies, and the environmental, employment, and social aspects to be considered in making investment decisions. He previously worked in the Forensic & Integrity Department and the Climate Change and Sustainability Services at EY. In 2009–2014, he headed the Working Group for Responsible Investment in the Prime Minister's Consultative-Advisory Body for CSR. He also chaired the Working Group for Ethics and Responsible Business Standards at the Ministry of Development. He has taken part in the initiatives of the Working Group for Artificial Intelligence at the Ministry of Digital Affairs. He is a member of the Working Group for Artificial Intelligence at the Chancellery of the Prime Minister of Poland. Author and editor of a number of books, reports, and guides on business ethics, including: "Undiscovered Leaders of Business Ethics" and "Ethics and Human Rights in Business. In Search of a Method".