Artificial Intelligence and Decision Making in Health: Risks and Opportunities

Márcia Santana Fernandes and José Roberto Goldim

Abstract The use of systems that include Artificial Intelligence (AI) imposes an assessment of the risks and opportunities associated with their incorporation in the health area. Different types of AI present multiple ethical, legal and social challenges. AI systems involved incorporated with new imaging and signal processing technologies. AI systems in the area of communication have made it possible to carry out previously non-existent interactions and facilitate access to data and information. The greatest concern involves the areas of planning, knowledge and reasoning, as AI systems are directly associated with the decision-making process. So, the central objective of this chapter is to reflect and suggest recommendations, with the foundation of the Complex Bioethics Model, about the decision-making process in health with AI support, considering risks and opportunities. The chapter is organized in two parts: (1) The decision-making processes in health and AI; (1.1) The health area the use of AI and decision-making processes: opportunities and risks to treat electronic health records (EHR) and (2) Complex Bioethics Model (CBM) and AI.

1 Introduction

Complexity, in the sense proposed by Edgar Morin, translates the moment we are living the so-called fourth Revolution. The *aversion to Manichaeism and the understanding that complexity is not everything, it is not the totality of reality, but it is the best that can, at the same time, open up to the intelligible and reveal the inexplicable*. The uncertainty of everyday life is an element of acceptance, or even its ambiguity. Artificial intelligence (AI), its potential uses and, in the same

M. S. Fernandes (✉) · J. R. Goldim
Laboratório de Pesquisa em Bioética e Ética na Ciência do Centro de Pesquisas do Hospital de Clínicas de Porto Alegre-LAPEBEC/HCPA DENS, Porto Alegre, Brazil
e-mail: msfernandes@hcpa.edu.br; http://lattes.cnpq.br/2132565174726788; jgoldim@hcpa.edu.br; http://lattes.cnpq.br/0485816067416121

H. Sousa Antunes et al. (eds.), *Multidisciplinary Perspectives on Artificial Intelligence and the Law*, Law, Governance and Technology Series 58,
https://doi.org/10.1007/978-3-031-41264-6_10

proportion, the legal, ethical and social challenges should be reflected in an *ambient* and *ambience* complex.

Technology and medicine has a long history of connection, but one of the milestones was the work of Lee Lusted, who in his article *Medical Electronics* (1955), reported on a series of large numbers of medical electronic devices developed in that time, indicating a rapid expansion of this field. He said at the time: *Electric phenomena in the human body had long been of interest, but the low signal amplitude made study difficult* (Lusted 1955). At the same period of time, Turing (Turing 1950) established the pillars for computer science and Artificial Intelligence (AI).

AI driven technologies impose an evaluation of the risks and opportunities associated with its incorporation in the life and living of human beings. The theme incorporates old-new questions, as Ulrick Beck pointed out in The Risk Society (1986),[1] and further developed in World at Risk (2007),[2] to the debate involving the impact of technology in the life of human beings: *how do we want to live? What is there of human in the human being, of natural in nature, that needs to be protected? (...) These old-new questions can be tossed back and forth between everyday life, politics and science. In the most advanced stage of the civilization process, they once again enjoy priority on the agenda - also or precisely at times when they are cloaked in the camouflage of mathematical formulas and methodological controversies* (Beck 2011, p. 34).

So, for some time, it has been possible to have a person-machine interaction by means of natural language systems (Chat-bot). On many occasions, there is no clear perception that this communication is being made with a machine and not with other people. In the 1970s, Jacques Monod already warned that it was increasingly difficult to establish the limit between the natural and the artificial (Monod, 1970). The simulation or substitution of real activities, increasingly similar to those performed by artificial mechanisms and systems, generates this ambiguity of perception.

The technological arrogance, according to Hans Jonas (Jonas 1994), causes these results to be understood as unquestionable. The infallibility of computers has been discussed since the beginning of their use, when they were still called "electronic brains". At that time there was already the proposition that the quality of the information generated was not unquestionable, but depended on the quality of the input data and the processes used. This became known by the acronym GIGO (Garbage In, Garbage Out). That is, if the data or systems are inadequate, the results generated will be compromised (The Hammond Times 1957).

These old-new questions have been at the heart of discussions involving AI and decision making. In this perspective, Floridi et al. (2018) in text, published in 2018, maintain that AI is a reality without return and for this reason it is necessary to form reflections towards an AI Society for Good (Good AI Society).

[1] Beck (2011). First edition in 1986.

[2] Beck (2009). First edition in 2007.

The opportunities and risks to protect the dignity of the human person and provide for their development should be permeated by the traditional principles of North American Bioethics—beneficence, non-maleficence, autonomy and justice (Beauchamp and Childress 1979), in addition to the principle of explicability.

In this analysis, in our view, two perspectives should be added: (1) European perspective, proposed by Peter Kemp and Jacob Dahl Rendtorff, use four other principles: Dignity; Autonomy, understood as Freedom; Integrity; and Vulnerability; because in this perspective, principles are not weighted, but there must be a coherence in their application (Kemp and Rendtorff 2008) and (2) an approach based by Complex Bioethics Model (CBM), in other words, bioethics understood as complex, shared and interdisciplinary reflection on the adequacy of actions involving life and living (Goldim 2006a, 2006b).

Life and living complement each other, they give the adequate dimension of each person. Life is described by the organic aspects, that is, by the biological characteristics. On the other hand, living refers to the relational aspects, the biography of each one (Agamben 1998). The ensemble of these characteristics is what gives the uniqueness of each person. The Complex Bioethical Model (CBM) (Goldim 2006b) embodies a perspective of a complex interdisciplinary field[3] of reflection on life and living.

It is precisely this desire to know and study population health and human health that marks scientific studies and establishes the foundations for research and experimentation. The need to respond to the challenges generated by epidemics, famine, wars, population growth and urban centres was the motivation for the chain to *the invention of science* (Wooton 2015).

Its central objective throughout time is to identify determinants of diseases and, more recently, of health at the population level. So, historically, the specific contribution of epidemiology has been the progressive constitution of a coherent set of methods and concepts, with the aim of assessing the determinants of health, where robust systems, like technologies driven by AI, are central to process and organize healthcare personal and sensitive data and information.

The processing of a lot of data in an efficient and precise way is fundamental for the development of scientific medicine, so computational tools for machine learning and mining large volumes of data, in an approach known as Big Data, and the joint evaluation of large volumes of data has allowed the establishment of new relationships, of new, previously unidentified understandings.

Another important development in this area is the increasing use of algorithms for decision making. These tools, increasingly improved and based on highly complex

[3] Bourdieu (1996, 2004). We use the expression field from Pierre Bourdieu. The *field* for Bourdieu is organized by principles such as economic capital and cultural capital, assuming struggles in social space, according to social positions. And composing the sense of *field* is the sense of *habitus*. Bourdieu's *habitus* can be understood as a system of dispositions, socially constituted, which establish the generating and unifying principle of the set of practices and ideologies characteristic of a group of agents. The habitus, from Bourdieu's perspective, produces the individual and makes him internalize the values and rules of belonging to society.

processes, have provided optimized solutions to countless problems, including modifying the decision making processes themselves. Most of these systems work in the quest to recognize patterns of similarity. This is the area that became known as Artificial Intelligence. Strictly speaking, artificial intelligence is not an intelligence in itself, but automated decision making processes. Pierre Levi makes a blunt criticism of the use of the expression "artificial intelligence", he does not recognize in these systems the possibility of generating new knowledge or of having an understanding of the world (Lévy 2022). Algorithms are made by people in the service of institutions, which have their belief systems and values, which end up directing the processing and interpretations.

Therefore, use of systems that include Artificial Intelligence (AI) imposes an assessment of the risks and opportunities associated with their incorporation in the health area: health care; experimental and clinical research and personalized medicine. AI systems involving areas of communication capable of performing parallel computations for data processing and knowledge representation (denominated artificial neural networks (ANN)); that have made possible to carry out previously non-existent interactions and facilitate access to data and information; technologies for detecting image, sound; performing heath assistance with robotics and areas of planning, knowledge and reasoning, when AI systems are directly associated with the decision-making process.

Ramesh et al. (2004), presented a literature review in 2004 on the use of the 'artificial intelligence' and 'neural networks (computer)' and an overview of different artificial intelligent techniques along with the review of important clinical applications. Their results show "the proficiency of artificial intelligent techniques has been explored in almost every field of medicine". The authors indicate areas of activity: clinical diagnosis; prognosis; ultrasound images; predict survival in patients; used for the administration of anaesthetics in the operating room; used form of evolutionary computation for medical applications in genetics e natural evolution, nominated 'Genetic Algorithms'. (Ramesh et al. 2004).

In this context, different types of AI present multiple ethical, legal and social challenges in the world, as pointed out by OCDE.[4] However, the diversity and vulnerability of social, economic and access to the Universal health coverage (UHC) that exists in South America made analises of health technologies AI-driven more complex. The standard to accomplish in terms of access of health are the UN's Sustainable Development Goals (SDGs), which, by 2030, a member state must guarantee: (1) access to health services for all people in need of health, independent of socio-economic characteristics, location, wealth or any other vulnerability; (2) financial protection, i.e. all people should be safe from financial risk when incurring health care expenses; (3) access to quality of health services, that is to say health care has to be effective in providing care and improving outcomes, while it is also

[4] OECD. Recommendation of the Council on Artificial Intelligence, OECD/Legal/0449, 2021. information, please consult the Compendium of OECD Legal Instruments at http://legalinstruments.oecd.org.

cost effective and sustainable, because access without quality can be considered an empty universal health coverage promise (OECD and The World Bank 2020).

The report *Health at a Glance: Latin America and the Caribbean 2020* compares key indicators for population health and health systems across the 33 Latin America and the Caribbean (LAC) countries. It presents comparable data on health status and its determinants, health care resources and activities, health expenditure and financing, and health care quality, along with selected health inequality indicator, including the pandemic COVID-19:

A main barrier for accessing such health services arise from out-of-pocket health expenditures, which in LAC represent on average 34% of total health spending, well above the 21% average in OECD countries. The high level of out-of-pocket expenditures in LAC are an indication of weaker health systems, lower levels of health services coverage and, overall, a worse baseline scenario to confront this pandemic when compared to most OECD countries (OECD and The World Bank 2020).

Particularly, in Brazil some parts of the country have more access to health and health technologies than others, despite the fact that Brazil has the biggest public health system in the World, the Unified Health System (*Sistema Único de Saúde—SUS)*, has as principle an universal access—an universal health coverage (UHC) to national and foreigners, that serves more than 190 million people, 80% of whom depend exclusively on it for any health care. The SUS is an achievement of the Brazilian people, guaranteed by the Federal Constitution of 1988, in Article 196, through Law No. 8.080/1990, that must be guaranteed and improved constantly.

So, the central objective of this chapter is to reflect and suggest recommendations, with the foundation of the Complex Bioethics Model, about the decision-making process in health with AI support, considering risks and opportunities. The chapter is organized in two parts: (1) The decision-making processes in health and AI; (1.1) The health area the use of AI and decision-making processes: opportunities and risks to treat electronic health records (EHR) and (2) Complex Bioethics Model (CBM) and AI.

The central assumption is to maintain balance and preserve the characteristics of humanity present in the act of deciding, taking into account ethical, legal and social aspects, cross by the principle of trust, when using AI systems in health. The justification for using personal and sensitive health data and information must be associated with actions on behalf of individuals and society, in terms of assistance, research involving human beings, whether of a sanitary, epidemiological, clinical or biobank nature.

We hope that our recommendations can contribute to the development of regulatory frameworks—ethical and legal—good practices and compliance for the use of AI in health. Our perspective is to analyze examples at national and international levels, focusing on the diversity and vulnerability of social, economic and access to health existing in South America and particularly in Brazil.

2 Decision-Making Processes in Health and AI

Decision making involving IA should consider Generative AI. The characteristics of the Generative AI systems are to establish connections, by means of new computational tools, based on large volumes of data, concepts and structures already built, to propose new contents, which are expressed in a language as natural as possible.

The term generative was created by Noam Chomsky, in 1965, when he proposed "generative grammar", or also called "generative grammar". This proposal sought to explain that language is established through skills and performances to produce sentences and what are the paths used to establish them (Chomsky 2015). This idea was based on the thought of Wilhelm von Humboldt, who, in 1836, said that language makes infinite use of finite means (Humboldt 1836).

The new Generative AI systems of, such as ChatGPT and Perplexity, constitute a new level of complexity in dealing with data, in the quest to extract information that can guide actions. These new systems not only label or classify pre-existing data, but generate new content, by aggregation and combination, from the available elements.

What are the ethical implications associated with Generative Artificial Intelligence? The texts prepared by ChatGPT and Perplexity respond:

> **ChatGPT:** Artificial Intelligence (AI) presents three main areas of ethical concern for society: privacy and surveillance, prejudice and discrimination, and philosophical challenges. AI can lead to unemployment as machines take over jobs that were previously done by humans. Inequality is also a concern, as AI can be used to perpetuate existing power structures. AI can also lead to algorithmic bias, where models incorporate and deploy human and social biases at scale.
>
> Privacy is a fundamental human right recognized in the UN Declaration of Human Rights, but AI apps can pose a real threat to this right. Misinformation is another ethical problem with AI, as machine learning models can easily generate incorrect text.
>
> To ensure the ethical use of AI, education and awareness about the ethics of AI must be increased. Examples of ethical AI include using AI to improve health outcomes or reduce energy consumption. It is important to remember that ethical issues with AI are not limited to negative impacts; there are also incidental benefits, such as enhanced analytical capabilities or increased efficiency in decision-making processes. (https://chat.openai.com/chat/c4bf8bd6-df96-4cec-b96d-dd54ec97ef2f; Feb 14th, 2023)
>
> **Perplexity**: Bias and discrimination: AI systems can perpetuate and amplify existing biases in society, leading to discriminatory decisions and outcomes. For example, facial recognition technology has been shown to have higher error rates for black people and women.
>
> Privacy and data protection: AI systems often rely on large amounts of personal data to make decisions, and there are concerns about how this data is collected, stored and used and who has access to it.
>
> Employment impact: AI has the potential to automate many jobs, leading to unemployment and social unrest. There is also a risk that AI will reinforce existing power imbalances and exacerbate economic inequality.
>
> Accountability and responsibility: when AI systems make decisions that have significant consequences for individuals and society, it can be difficult to determine who is responsible for those decisions and how they can be held accountable.

> Human rights: AI has the potential to infringe on human rights, such as freedom of expression and privacy, and there are concerns about the use of AI for surveillance and control by governments and other organizations.
>
> It is important to address these ethical implications through ongoing dialogue and debate and by establishing clear ethical guidelines and principles for the development and use of AI. This includes considering the potential impacts of AI on individuals and society and ensuring that AI is developed and used in a responsible and ethical manner. (https://www.perplexity.ai/?s=u&uuid=a6d1842e-bf56-4e0a-9fd4-2ff4f00a8e13; Feb 14th, 2023)

As we can see there are connections, from the ChatGPT and Perplexity answers, between ethical issues and concepts, such as privacy, surveillance and control, freedom of speech, prejudice and discrimination, were interpreted and new content arranged. All these issues already preexisted the use of generative AI, however they have been discussed in different levels of depth and application. Perhaps, the current challenge is to think about these issues in this new perspective.

Whenever an innovation is incorporated by society, discussions about its suitability arise. When the printed book was made available to society, when the first encyclopedias appeared, when the internet gave access to a volume of data never before imagined, there was questioning about the appropriateness of using these means of disseminating data, knowledge, and information. A good example is the current discussion of the impact of Generative AI on education, it is generating anxiety among schools, parents and teachers. This same discussion has occurred in other historical moments and incorporation of new technologies. There are already educational models that allow incorporating these situations brought by Generative AI in a creative way. Instead of copying or generating content, perhaps the challenge of education is to evaluate the quality of the information generated. It is to use this challenge to incorporate a critical and complex reflection in the different levels of life to establish security, transparency and trust in the use of AI-Generated.

The *Study on e-Health Interoperability of Heath Data and Artificial Intelligence for Health and care in the European Union—Final Study Report* (European Union 2021) points out the lack of trust in AI-driven decision support is hindering the wider adoption in heath, and also integrating new technologies into current clinical practice; research and personal medicine are indeed legal, ethical and social challenges. These challenges are increased by the necessary internationalization of the health area and the challenges of sharing data and information in order to achieve global health.

Recommendations have also been developed by countries and organizations, highlighting the recommendation proposed by the European Commission, in 2020, in the "White Paper—On Artificial Intelligence—A European approach to excellence and trust", with the purpose of establishing the political paths to seek the appropriate use of AI. In this document, the Commission recommends the establishment of standards and guidelines for investment in the area of AI, aiming at two central objectives: promoting the adoption of AI and addressing the risks associated with certain uses of this new technology (European Commission 2020). The Commission also established a High Level Expert Group that published Guidelines on trusted AI in April 2019, composed of seven key requirements: respect

for the dignity of the human person; robust technical and security systems; privacy and data management; transparency; respect for diversity, non-discrimination and equity; social and environmental well-being; and accountability.

The Common Digital Market is one of ten priorities of the European Union. In this context, the following decisions are taken: Decision No 922/2009 / EC of the European Parliament and of the Council of September 2009 on interoperability solutions for European public administrations (e-Health European Interoperability Framework) (European Union 2012) and Decision (EU) 2015/2240 of the European Parliament and the Council of 25 November 2015 establishing a program on interoperability solutions and common frameworks for public administrations, businesses and citizens (ISA program) as a means of modernizing the sector. European e-Health Interoperability Framework (ReEIF) (European Union 2015).

The European Union seeks to integrate the electronic medical records of European citizens, recognizing the weaknesses related to various aspects of data use, whether for security, privacy protection, ethical suitability, management, storage and disposal, and interoperability between state information systems to establish trustable structure of E-Health. These measures are part of the goal of creating a digital single market.[5]

Decision-making processes, particularly in the health area, are based on trust and the relationship of trust—which are necessarily identified with all those involved in this relationship. The relationships occur in all spheres, between the public administration and the administered; between private entities; between private entities and human beings and between human beings. The pre criteria for establishing the basis of trust, in situations involving IA, are not different, on the contrary should be intensified, because must be composed of concrete mechanisms to inform, account for the use, motivation, process and transparency of the criteria used in decision making.

The principle of trust lies at the basis of legal relations, whether these are public or private. In turn, the principle of the protection of trust is presented in the individual dimension, or in the subjective aspect of legal security. This principle depends on the exercise of trust, with concrete indication of the breach of expectations in law or clear demonstration of the requirements for its demonstration.

O'Neill understands that trust cannot be confused with the mere disclosure or transparency of information and accountability (O'Neill 2004). From the philosophical perspective, trust is a central element in human relations, whether interpersonal or between individuals and the state, involving trust in institutions and their representatives. However, this state of trust is not presented merely by the disclosure of data and information, but must be underpinned by an intelligible narrative.

In the juridical perspective, the principle of trust, says Martins-Costa, has the immediate scope to ensure expectations. In the case in question, the situation of

[5] The example and efforts made by the European Union to integrate regulatory, technical, ethical and social aspects in the area of digital health are important to reflect in the design of systems, similar or not, for other parts of the world.

trust is materialized between the individual and the public administration, when personal data are provided for precise purposes—as health care, research or social security (Martins-Costa 2015). It also presents itself in legal businesses, involving the provision of personal data in exchange for specific health services.

2.1 The Health Area the Use of AI and Decision-Making Processes: Opportunities and Risks to Treat Electronic Health Records (EHR)

Undoubtedly health care; research involving human beings or public policy design—data and information are central. In turn, the use of AI in this scenario relies and requires the data and information spent in electronic health records (EHR). Therefore, the treatment of health data and information, sensitive data, must be based on the principle of trust. So, studying some aspects related to the use of EHR, combined with AI technologies, is a good example to establish opportunities and risks of this technology in the health area.

Electronic medical records serve as a collective memory of the assistance provided to the patient. Thus, they must gather general and health data records, the description of relevant personal and family facts, collected by health professionals during the patient's anamnesis. It is this history that opens the record of the assistance activities. Besides this information, other information is added, either as a record or as annexes, such as diagnoses, under the form of reports, images or data, prognoses, care plans, exam results, consultations performed by different professionals, participation in research or notes that are relevant to the case, with the primary purpose of better assisting the patients (Fernandes and Goldim 2019).

2.2 The Opportunities

The EHR must be protected and guided by a relationship of trust, based on respect for the person. The respect for the person is expressed by the deontological duties of confidentiality, by the legal duties of personality and by the bioethical principles. The patient provides the information considered as relevant based on the trust placed in the professional who is attending him/her. From the professional's point of view, this information is always considered to be privileged.

The use of genetic data in care, such as those used in genetic counselling and Personalized Medicine, has introduced new data, which may generate information that affects not only the patient but also other people related to him or her. Thus, the concept of personal privacy expands to that of relational privacy. This increases the responsibility associated with the registration and future use of this information.

As well as, the data EHR may assist in research-related interventions. This may involve the use of medicines, cells and other biological products, the performance of surgical or diagnostic procedures, the use of devices, changes in the care process, preventive care, among other activities. In all of them the sharing of these data can generate new and useful information.

The development of clinical research and also of personalized medicine, extends the care with the protection of personal data in the area of health, as they involve the need to use data and information of patients, collected in an protected environment by the principle of trust. Likewise, this principle creates expectations in the research participant—on a personal level, when the results of the research can affect or be beneficial to him/her, or on the social and community level of collaborating with scientific development.

Besides, EHR have been used as qualified sources of information for the establishment of public policies and research. Public policies are essential to guarantee access to health. It is worth noting that, from the perspective of Law, the issue of access to health is dealt with in the context of fundamental and civil rights. The importance, for example, of epidemiological cause-effect relationships studies, which make it possible to establish public policies, protocols, guidelines or norms for the prevention and treatment of diseases and/or for health promotion, are unquestionably important and they change the course of human development.

Moreover, from the epidemiological approach, Evidence-Based Medicine (EBM) emerged, proposed by McMaster University, Canada, in the 1990s, to record and systematize clinical evidence and the epidemiological knowledge derived from it, to improve results in the diagnosis and treatment of diseases and health care. It is an attempt to guide patient-associated decision making at the individual level based on collective data (Evidence-Based Medicine Working Group 1992).

So, the need to systematize the collection, storage and use of health data and information is directly connected with the development of medicine and the global increase of knowledge in health, both in terms of individual patient care, population health and global health. And today we have at our fingertips and in constant use tools such as AI to do that.

2.3 The Risks

The protection of personal data and information contained in EHR should consider the new context generated in the Information Society, for risk prevention (Fernandes 2019). The constant development and incorporation of new information and communication technologies, the use of new data protection techniques, including AI, blockchain, the use of social media, the interconnection of integrated health systems, in addition to the sharing generated by the Big Data environment itself (Roehrs et al. 2019).

So, these expanded possibilities of interconnecting, storing and processing a large and complex volume of data and information originated from EHR, amplifies the

national and international concern, demonstrated in the literature, about the security and preservation of patient data and information contained in PEPs. Particularly, the respect and the adequate use for its purpose—in favor of the patients—are highlighted topics. The literature review carried out in 125 scientific articles, selected from a total of 5278 articles, in the PubMed and Scielo databases, indicates as recurrent themes information security when dealing with electronic records and access to medical records (Caballero 2018).

EHR presents the data and information in a structured way, however, as the medical history should be developed in a text contextualized in the patient's life and living, the qualitative or even quantitative analysis may be hampered. Also, personal data and information, especially in health, should be considered as distinct concepts. Information does not exist in isolation, it needs a receiver, someone to give meaning and significance to the data. Isolated data describes characteristics of something, someone, some fact or situation. However, it is the information that gives meaning to this data. The information acts on the data, it is the result of the analysis and interpretation of the data. In short, it is the organization, categorization and systematization of data for a specific perspective and purpose that generates information. These definitions are a relevant starting point for understanding the importance of data and information in the area of health (Fernandes 2019).

Likewise, it is important to consider various notions and concepts related to the environment of large volumes of data—Big Data—generated in EHR and in the health system (Kulynych and Greely 2017). Big Data is an expression used generically to indicate the grouping of data, information, databases, open internet networks and other accessible data that initially aimed to improve strategic planning, marketing and commercial business (Manyika et al. 2011). This context, marked by fluidity, uncertainty and fugacity of data and information, required multiple sources to seek to understand complex and broad phenomena that AI systems can help to interpret.

The new perspective generated by the Big Data phenomenon has stimulated scientific work in various areas of knowledge. As pointed out by Mittelstadt and Floridi in a literature review article of 2016, AI in the health area is already a reality, besides others that would be on a horizon possibilities and others that are still only potential. Examples of situations that are already a reality include those related to the activities of Biobanks, Public Health studies and hypothesis testing in the health area. Possible situations include the interconnection of equipment and applications for personal health; the existence of online profiles connected to medical records; the creation of social media in the health area and the online and offline connections of personal profiles via wifi. Finally, they indicated as potential situations the connections between online medical records with other sources of personal data, as well as the involuntary connections of these data, both online and offline, originating from personal profiles for health surveillance purposes (Mittelstadt and Floridi 2016).

Floridi says:

> Clearly, the future of AI lies not just in "small data" but also, or perhaps mainly, in its increasing ability to generate its own data. That would be a remarkable development, and one may expect significant efforts to be made in that direction. As well as, translated difficult tasks into complex tasks. (...)How is this translation achieved? By transforming the environment within which AI operates into an AI-friendly environment (Floridi 2019, 2020).

For this and other reasons, accurate risk impact analysis and preventive actions should be taken, in the normative, good practices, compliance and ethical spheres, mainly to avoid bias in decision making. Algorithmic bias is one of the fears, particularly in the processing of sensitive personal data, suche as health, genetic and biometric data. As well as, algorithmic bias that may negatively discriminate and/or cause harm to individuals or certain groups—e.g. organised by gender, sex, age, physical or mental health status and economically or socially vulnerable peaple or groups (e.g. prisoners and the poor).

Norori et al. (2021), in an article entitled *Addressing bias in big data and AI for health care: A call for open science* point out that *the f*uture, and we would say the present, research is needed to set standards for *AI in healthcare that enable transparency and data sharing, while at the same time preserving patients' privacy.*

The authors present the distinctions between *statistical bias* and *social bias* as the starting point of the analysis. Statistical bias being that *refers to cases in which the distribution of a given dataset is not reflecting the true distribution of the population* and in turn, social bias *refers to inequities that may result in suboptimal outcomes for given groups of the human population* Norori et al. (2021).

The authors point out some examples of AI algorithms that are biased by design, regarding sex, age and race. The bias can be observed in studies that discriminate against the female gender in favor of the male, including in pre-clinical research, when in experimental models using animals there is a predominance of males. As well as in research for the development of medicines, when the majority of participants are men without a methodological reason that justifies it. Also, they pointed out, by the example, studies in the area of sleep disorders, when young patients are in favour of older patients. Moreover, racial bias when algorithms, in the area of skin cancer, are programmed to identify images of light skin and not dark skin, even if black population has a higher mortality rate from melanoma cancer. Also in the area of negative discrimination by race, there are algorithms in the area of hospital costs that induce to determine that black patients are healthier than white patients and for this reason, these receive a better treatment Norori et al. (2021).

These examples are enough to demonstrate that fears and lack of confidence in AI driven decision making are not in vain, or even disproportionate—they are a reality that should be normatively and ethically avoided.

Negative discrimination, as we have pointed out, in Latin America is aggravated by the large number of people who do not have access to health technologies or are discriminated against in "broad daylight" because of their condition and economic deprivation, lack of education and lack of sanitary conditions—as can be seen in the slums and peripheries and which is evidenced in the COVID-19 pandemic—

ironically data that has also been evidenced with the help of AI driven technologies (OECD and The World Bank 2020; Norori et al. 2021).

So what to do? Where should we act—nationally and internationally? What parameters should we have as a starting point? To try to answer and/or reflect on these questions, we turn to our second point, the Complex Bioethics Model (CBM) and AI.

3 Complex Bioethics Model (CBM) and AI

We are at a historic stage in which digital immigrants and digital natives coexist (Zur and Walker 2019). Digital immigrants had the opportunity to live in a society where all decisions were made only by human beings. Digital natives, on the other hand, naturalize the decisions made by algorithms.

The naturalization of decisions made only by artificial intelligence can involve several important ethical issues, such as technological arrogance, the vision of certainty and the impartiality of algorithms.

By using algorithms, machines follow a pattern of predictable, pre-programmed steps. Even with the incorporation of associated machine learning processes, these decisions carry with them only the rational elements associated with the decision-making process. In some models, values, affective issues and even cultural traditions can be included as elements of this decision-making process (Weber 1978). However, these non-rational actions are considered as if they were rational by the computational model. The computer doesn't hesitate, humans hesitate (Han 2015).

The processes used in artificial intelligence are the result of programming. Programming does not tolerate ambiguity or uncertainty, which are always present in the real world. Even using fuzzy logic-based methodologies, strictly speaking, it is a programmed uncertainty.

There is a belief that human beings are fallible but machines are not. Any and all decision-making process using artificial intelligence is based on a set of assumptions established by human persons. Even when there are self-programming systems, the root of the process is based on choices made by people who planned and implemented them. There are different levels of complexity, but they converge to a root where there is the presence of non-rational characteristics of its developers.

From an ethical point of view, any and all human action, or resulting from it, must be evaluated for its adequacy (Vasques 2000). This assessment requires not only the consideration of the facts, but the whole set of circumstances. One of these circumstances is the historical dimension, it is the perspective of insertion of these activities over time. It is a critical need to understand the complexity of the problem being evaluated.

This apparent dichotomy between artificial and natural is increasingly tenuous (Monod 1970). It is increasingly important to have a complex perspective in

understanding situations that have increasingly presented themselves to human society.

Using a complex approach to Bioethics, it's possible to have ethical arguments using different theoretical framework, Any of them, based on virtues; intention and consent; principles; responsibility; human rights; consequences and alterity could be used in order to understand the human-computer systems relation.

Virtues can be used to justify the personal behavior adequacy involved in the design and application of decision-making systems. Prudence, temperance and justice are fundamental virtues to be considered in these situations. Systems must be based on practical reasoning, must use the resources involved properly and, above all, do not discriminate against any person or group of persons. Virtues presupposes a desire for humanity, which projects itself in time, which always has a historical perspective (Comte-Sponville 1996).

The intentions and consent associated with the action must be considered in evaluating the moral worth associated with an action (Abelard 1995). The intention of whoever designs or uses a system must be adequate, it must aim at the good of the people. On the other hand, the use of the system is only considered appropriate when it has the consent of the people affected by it. This combination of wills, of those who do the action and those who suffer the action, is fundamental.

Principle-based ethics should also guide the assessment of the appropriateness of using systems. The four-principle framework—Dignity, Freedom, Integrity and Vulnerability—can be very helpful in these assessments (Kemp 2005). Coherence in the application of these principles, understood as guiding human actions, must be sought. Dignity unites us to all people, it is what gives the character of humanity to all of us. Freedom is the possibility to choose, to make choices free from coercion. Integrity, understood in its physical, mental and social dimensions, must always be based on the search for its preservation. Vulnerability should be considered whenever there is any possibility that dignity, freedom or integrity could be compromised. In a risk society, we are all always vulnerable, in different degrees and situations.

The ethics of human rights is based on expectations of action. Human rights can be approached from an individual or collective perspective or even in a transpersonal way (Bobbio 1992). From the right to life and privacy to the right to solidarity or to have a preserved environment, rights are expressions of other people's actions towards me. Artificial intelligence systems may not have this multiple perspective present when making decisions. Sometimes one right is privileged and the others are not taken into consideration. It may happen that, by guaranteeing the right to privacy, a system will end up abandoning the dimension of solidarity (de Oliveira Ascensão 2009).

Consequentialist ethics is based on risks and benefits associated with actions directed to individuals or collectivities. The consequentialist decision-making process, from a micro or macro point of view, is based on the analysis of utility (risk versus benefit) (Singer 1993). The most important issue is to aim for a balance between these two perspectives, to establish a win-win strategy.

Ethics of responsibility focuses on actions. Both perspectives, whether retrospective or prospective, assess the repercussions of the actions carried out. The retrospective approach focuses on causes and the prospective on effects. The usual approach to responsibility is to see who did it and how the action was done. More recently, the focus has shifted from the cause to the action's repercussions. If scientists are responsible for the social consequences of science (Marcuse 2009), so too are the people who design artificial intelligence systems. In this perspective, a new imperative was established in response to human actions: "In your present options, include the future integrity of the human being among the objects of your will" (Jonas 2006). In other words, we shouldn't do everything that technique allows us to do (Ropohl 1981).

Finally, alterity is another theoretical approach to evaluate the ethical basis of artificial intelligence. Systems are built to be permanent, to have an identity, an immutability and to assume the totality of associated actions. This is the perspective of sameness. Otherness, on the other hand, assumes impermanence, singularity, mystery and infinity (Levinas 1961). Otherness opens us to the other and reaffirms us as people. This perspective allows establishing an ethical co-presence, a co-responsibility, a perspective that goes beyond the simple relationship to become an effective interaction (Levinas 1991). From the perspective of artificial intelligence, sameness prevails over otherness. In alterity perspective it's impossible to approach new technologies from a neutral point of view.

The Complex Bioethics approach allows the integration of these different theoretical perspectives in the search for arguments to reflect on the adequacy of the use of artificial intelligence technologies (Goldim 2006b). It's a good way to get a comprehensive perspective on proposals that are often seen only in their technical aspects.

Contemporary ethical discussion should be guided by reflection on the new "information regime", as characterized by Han (Han 2021). This is our challenge: to reflect on this new model of society, where relationships have changed enormously. AI, Generative AI and other species, is just one of the multiple challenges that need to be discussed and deepened.

4 Conclusion

Artificial intelligence is defined in the sense that machines can perform tasks similar to those performed by human beings (McCarthy et al. 2006). In the beginning of computing, computers were called "electronic brains". Later, computer metaphors were used to explain how the human brain works. One of the current risks is to carry out this inversion again, that is, to want to explain human intelligence using artificial intelligence models.

Another challenge in transposing human intelligence to artificial intelligence is recognizing that humans can fail, then machines can fail too. If this transposition occurs, it could be the realization that there would be a proposal for an "artificial

stupidity" (O'connell 2017) associated with "artificial intelligence", like human stupidity and intelligence. Ethics and Bioethics could help in the reflection of the adequacy about the limits and borders between natural or artificial and intelligence or stupidity.

More important than discussing punctual ethical aspects, it is fundamental to reflect on the broader aspects of the use of AI, such as:

(a) to define ethically appropriate standards to guide the responsible creation of content by these systems;
(b) to establish monitoring strategies for the data and information generated by the AI and Generative AI to verify the veracity;
(c) to create guidelines that allow continuous audits of the processes of these systems in order to prevent that their processes can be used for purposes contrary to the interests of people, societies and humanity (Gocklin 2023).

In the bioethical approach to new technologies, it is essential to associate the principle of precaution with the principle of hope (Patrão-Neves 2021). That is, precaution seeking to guarantee the life of each one and hope seeking to maintain everyone's living.[6]

References

Abelard P (1995) Ethical writings: "ethics" and "dialogue" between a philosopher, a Jew and a Christian. Hackett, Indianapolis

Agamben G (1998) Homo sacer: sovereign power and bare life. Stanford University, Stanford

Beauchamp TL, Childress JF (1979) Principles of biomedical ethics. Oxford, New York

Beck U (2009) World at risk. Translated by Ciaran Cronin. Policy Press, Cambridge

Beck U (2011) Sociedade de risco – rumo a uma modernidade. 2ª Edição em tradução em Portugues, traduzida por Sebastião Nascimento. Editora, São Paulo

Bobbio N (1992) A era dos direitos. Campus, Rio de Janeiro

[6] *See* generally, on the different applications of Machine Learning and AI, in this book A Oliveira and M A T Figueiredo—Artificial intelligence: historical context and state of the art; I Trancoso, N Mamede, B Martins, H S Pinto and R Ribeiro—The impact of language technologies in the legal domain; J Gonçalves-Sá and F L Pinheiro—Societal Implications of Recommendation Systems: A Technical Perspective; A T Freitas—Data-driven approaches in healthcare: challenges and emerging trends; M Correia and L Rodrigues—Security and Privacy; E Magrani and P G F Silva—The Ethical and Legal Challenges of Recommender Systems Driven by Artificial Intelligence; M Lanz and S Mijic—Risks associated with the use of natural language generation: Swiss civil liability law perspective; W Gravett—Judicial Decision-making in the Age of Artificial Intelligence; and D Durães, P M Freitas and P Novais—The Relevance of Deepfakes in the Administration of Criminal Justice. *See* also, on AI and Healthcare, in this book A T Freitas—Data-driven approaches in healthcare: challenges and emerging trends; M N Duffourc and D S Giovanniello—The Autonomous AI Physician: Medical Ethics and Legal Liability; and R Nogaroli and J L M Faleiros Júnior—Ethical challenges of artificial intelligence in medicine and the triple semantic dimensions of algorithmic opacity with its repercussions to patient consent and medical liability.

Bourdieu P (1996) Razões práticas: sobre a teoria da ação. Papirus, Campinas
Bourdieu P (2004) Os usos sociais da ciência: por uma sociologia clínica do campo científi-co. UNESP, São Paulo
Caballero LG (2018) Informação de pesquisa clínica e a interface com o aplicativo de gestão para hospitais universitários: desafios éticos e regulatórios. Dissertação de Mestrado orientada por Márcia Santana Fernandes; co-orientada por Rafael Leal Zimmer. Programa de Pós-Graduação Stricto Sensu, Mestrado Profissional em Pesquisa Clínica do Hospital de Clínicas de Porto Alegre, Porto Alegre
Chomsky N (2015) Aspects of the theory of syntax. The MIT Press, Cambridge
Comte-Sponville A (1996) Pequeno tratado das grandes virtudes. Martins Fontes, São Paulo
de Oliveira Ascensão J (2009) Pessoa, direitos fundamentais e direito da personalidade. Rev Fac Direito Univ Lisb 50:9–31
European Commission (2020) White paper – on artificial intelligence – a European approach to excel-lence and trust. https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en. Accessed 24 Oct 2021
European Union (2021) Study on ehealth, interoperability of health data and artificial intelligence for health and care in the European Union. Publications Office of the European Union, Luxembourg
Evidence-Based Medicine Working Group (1992) Evidence-based medicine. A new approach to teaching the practice of medicine. JAMA 268:2420–2425
Fernandes MS (2019) Privacidade, sociedade da informação e Big Data. In: Benetti G, Corrêa AR, Fernandes MS, Nitschke GM, Pargendler M, Varela LB (eds) Direito, cultura e método - leituras da obra de Judith Martins-Costa. GZ Editora, Rio de Janeiro, pp 182–210
Fernandes MS, Goldim JR (2019) A sistematização de dados e informações em saúde em um contexto de big data e blockchain. In: Lucca N, de Lima CRP, Simão A, Dezem RMMM (eds) Direito e Internet IV. Quartier Latin, São Paulo, pp 333–357
Floridi L (2019) What the near future of artificial intelligence could be. Philos Technol 32:1–15
Floridi L (2020) What the near future of artificial intelligence could be. In: Burr C, Milano S (eds) The 2019 yearbook of the digital ethics lab. Springer International Publishing, Cham, pp 127–142
Floridi L, Cowls J, Beltrametti M, Chatila R, Chazerand P, Dignum V, Luetge C, Madelin R, Pagallo U, Rossi F, Schafer B, Valcke P, Vayena E (2018) AI4People-an ethical framework for a good AI society: opportunities, risks, principles, and recommendations. Minds Mach 28:689–707
Gocklin B (2023) Guidelines for responsible content creation with generative AI [Internet]. Digital Transformation. pp 1–7. Available from: http://contently.com/2023/01/03/guidelines-for-responsible-content-creation-with-generative-ai/. Accessed 9 Feb 2023
Goldim JR (2006a) A avaliação de projetos de pesquisa: aspectos éticos, científicos, legais, regulatórios e éticos. Rev HCPA 26:83–86
Goldim JR (2006b) Bioética: origens e complexidade. Rev HCPA 26:86–92
Han BC (2015) Sociedade do cansaço. Vozes, Petrópolis
Han BC (2021) Infokratie: Digitalisierung und die Krise der Demokratie. Mathes & Seitz Verlag, Berlin
Humboldt W (1836) Über die Verschiedenheit des menschlichen Sprachbaues: und ihren Einfluss auf die geistige Entwicklung des Menschengeschlechts. Druckerei der Königlichen Akademie der Wissenschaften, Berlin
Jonas H (1994) Ética, medicina e técnica. Vega, Lisboa
Jonas H (2006) O princípio responsabilidade: ensaio de uma ética para a civilização tecnológica. Contraponto, Rio de Janeiro
Kemp P (2005) The globalization of the world. In: Kemp P (ed) Philosophical problems today: world and worlshood. Springer, Berlin, pp 15–28
Kemp P, Rendtorff JD (2008) The barcelona declaration. Towards an integrated approach to basic ethical principles. Synth Philos 23(2):239–251

Kulynych J, Greely HT (2017) Clinical genomics, big data, and electronic medical records: reconciling patient rights with research when privacy and science collide. J Law Biosci 4:94–132

Levinas E (1961) Totalité et infini. Martinus Nijhoff, La Haye

Levinas E (1991) Entre nous: essais sur le penser-à-l'autre. Grasser & Fasquelle, Paris

Lévy P (2022) Inteligência Artificial produz "monstruosidades éticas". Revista Fórum [Internet]; pp 1–7. Available from: https://revistaforum.com.br/ciencia-e-tecnologia/2022/12/27/inteligncia-artificial-produz-monstruosidades-eticas-diz-pierre-levy-129333.html

Lusted LB (1955) Medical electronics. N Engl J Med 252:580–585

Manyika J, Chui M, Brown B, Bughin J, Dobbs R, Roxburgh C, Byers AH (2011) Big Data: the next frontier for innovation, competition and productivity. McKinsey Global Institute, New York

Marcuse H (2009) A responsabilidade da Ciência. Sci Stud 7:159–164

Martins-Costa J (2015) A Boa-Fé no direito privado: critérios para sua aplicação. Marcial Pons, São Paulo

McCarthy J, Minsky ML, Rochester N, Shannon CE (2006) A proposal for the dartmouth summer research project on artificial intelligence. AI Mag 27:12

Mittelstadt BD, Floridi L (2016) The ethics of Big Data: current and foreseeable issues in biomedical contexts. Sci Eng Ethics 22:303–341

Monod J (1970) Le hasard et la nécessité. Seuil, Paris

Norori N, Hu Q, Aellen FM, Faraci FD, Tzovara A (2021) Addressing bias in big data and AI for health care: a call for open science. Patterns 2. http://creativecommons.org/licenses/by/4.0/. Accessed 2 Sept 2022

O'connell M (2017) Art as "artificial stupidity". PhD thesis, University of Sussex, Brighton

O'Neill O (2004) Accountability, trust and informed consent in medical practice and research. Clin Med 4:269–276

OECD, The World Bank (2020) Health at a glance: Latin America and the Caribbean 2020. OECD Publishing, Paris

Patrão-Neves MDC (2021) Política do medo: amigos ou inimigos? Privacidade e segurança em tempos de medo global. In: Política do medo ou o mundo de hoje entre privacidade e segurança. E-book -Universidade Católica Editora, Lisboa. Acessível em https://www.uceditora.ucp.pt/pt/colloquia/3089-colloquia-1.html

Ramesh AN, Kambhampati C, Monson JRT, Drew PJ (2004) Artificial intelligence in medicine. Ann R Coll Surg Engl 86:334–338

Roehrs A, da Costa CA, da Rosa Righi R, da Silva VF, Goldim JR, Schmidt DC (2019) Analyzing the performance of a blockchain-based personal health record implementation. J Biomed Inform 92:103140

Ropohl G (1981) Technik - ein problem der philosophie? Philos Nat 18:413

Singer P (1993) Practical ethics. New York, Cambridge

The Hammond Times (1957) Work with new electronic "brains" opens field for army math experts. The Hammond Times. Nov 10, p 65

Turing AM (1950) Computing machinery and intelligence. Mind LIX(236):433–460

União Europeia (2012) eHealth interoperability framawork study. https://publications.europa.eu/en/publication-detail/-/publication/2bc03deb-afd8-4c2a-a394-16c039232b1e/language-en/format-PDF/source-95947333. Accessed 2 Sept 2022

União Europeia (2015) Decisão (UE) 2015/2240 do parlamento Europeu e do conselho. https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:32015D2240&from=en. Accessed 2 Sept 2022

Vasques AS (2000) Ética. Civilização Brasileira, Rio de Janeiro

Weber M (1978) Economy and Society: an outline of interpretive sociology. University of California, Berkeley

Wooton D (2015) The invention of Science – a new history of the scientific revolution. Harper Collins Publishers, New York

Zur O, Walker A (2019) On digital immigrants and digital natives. Zur Institute, Sebastopol