Philosophy and Artificial Intelligence: Exploring the Nexus between New and Old Artificial Crossroads

[1]Dr. Maigona timothy dodo, [2]Gotish Lengji Joash & [3]Abubakar Yahaya

[1]dodotimothy686@gmail.com

[1&2]Faculty of Education,
Department of Educational Foundations,
Nasarawa State University, Keffi

&

[3]Department of Educational Management,
Federal University of Education, Kontagora

DOI: https://doi.org/10.5281/zenodo.21548190

Abstract

The work is philosophically inclined and therefore the researchers adopt kate Turabean referencing style in undertaking the study. The article discusses which philosophical tools are indispensable and fundamental for understanding the meaning of technology in modern times. The intersection of Philosophy and artificial intelligence (AI) presents a fascinating nexus of old and new, where ancient questions about human nature, consciousness, reality converge with cutting edge technological advancements. As artificial intelligence (AI) continues to transform industries and challenge traditional notions of intelligence, agency, and existence, philosophical inquiry becomes increasingly essential for understanding the implications of these developments. This explores the artificial crossroads where Philosophy and AI meets, examining the tensions and synergies between human thoughts and machine intelligence. By revisiting classical philosophical debates and engaging with contemporary AI research, we can better grasp the opportunities and challenges arising from this convergence. Succinctly, this exploration reveals that the nexus between Philosophy and Artificial Intelligence is not merely a meeting point but a dynamic interface that can inform and transform both fields, yielding a new understanding into human condition and the possibilities of artificial intelligence. Some recommendations were prescribed to solve this challenge between philosophy and Artificial intelligence which were, interdisciplinary approach, human –AI collaboration, and future of work and education.

Keywords: Philosophy and Artificial intelligence.

Introduction

The technological evolution and human experience over the last four decades has brought about an important debate between computer science, mathematics and philosophy, regarding human-machine interaction. This topic reached interdisciplinary importance at the beginning of the 1980s and reached its peak at the end of the 1990s. At this point in history, personal computers gain market share, reaching recurrent civilian use and leaving the specialist use zone, expanding the more intensive use of the connection between

universities, governments and military bodies. The last two decades of this century have been iconic for the world of technological evolution, and the radical changes brought to educational institutions and companies in the throes of transformation and innovation. The process of globalization of the economy and culture is marked by advances in information systems that accompany trends towards connectivity and global standardization through miniaturization of electronic systems with increased speed and capacity, portability and compatibility of devices and the limitless growth of digitized information. The training of professionals enters society in the so-called data age and reaches an irreversible level of connection between artificial intelligence and social learning. In this process, anew area of knowledge called big data has emerged, whose field of research focuses on how to process, analyze and obtain information from large, complex data sets, from recurring new sources. According to Magrani,[1] the history of the internet can be understood in three generations: the internet of machines, the internet of people and the internet of things.

The process of industrialization introduced, for the first time, intelligent systems capable of carrying out tasks performed by workers. Industrial robots have had an impact on the world of production and work, increasing profitability by improving product quality, reducing production costs, and replacing workers in some tasks. Automation software offers the possibility of performing simple administrative support tasks and is evolving into large-scale automation of more complex processes. In this context, debates are intensifying about the relationship between humanity and technology, and narratives about whether the impacts have been positive or negative. In the fields of exact sciences and nature, humanities and the arts, many questions have arisen about the essence of human beings, fueling the tension between technological evolution and human experience. Since the second decade of this century, the Internet of Things (IoT) and the 5G Internet have brought greater connectivity to all processes involving the use of technology and culture. They have ample data transfer capacity and extremely high connection speed standards with multiple users connected simultaneously.

The concept “Philosophy”

Western philosophy was born in Greece; the term ‘philosophy’ too has its roots in Greece and in Greek language. It is quite commonly known that *philosophia* etymologically means ‘love of wisdom’ (*Philia* + *Sophia*), but *sophia* had a much wider range of application than the modern English "wisdom." Wherever intelligence can be exercised—in practical affairs, in the mechanical arts, in business—there is room for *Sophia*. Herodotus used the verb *philosophein* in a context in which it means nothing more than the desire to find out. We can find a gradual growth in the meaning of philosophy, as we go through the history of thought.

According to a tradition, Pythagoras was the first to describe himself as a philosopher. He speaks of three classes of people, attending the festal games: those who seek fame by taking part in them; those who seek gain by plying their trade; and those who are content

to be spectators. Philosophers resemble the third class: spurning both fame and profit, they seek to arrive at the truth by contemplation. Pythagoras distinguished the *sophia* sought by the philosopher (knowledge based on contemplation) from the practical shrewdness of the businessman and the trained skills of the athlete. Plato points to Socrates as *the* philosopher. Plato gives a few characteristics of philosophical wisdom, such as ability to enter into critical discussion, having direct access to "true reality," knowledge of the purpose of life, etc. As evident from above, although philosophy is etymologically defined as 'love of wisdom', the meaning of wisdom is taken in a wider sense.

Oxford Dictionary[2]defines philosophy as "that department of knowledge which deals with ultimate reality or with the most general causes and principles of things." It is presumed here that science, inheriting the cosmological tradition, does not offer us the knowledge of ultimate reality; only philosophy can do this. Science can only tell us *how*, whereas philosophy can tell us *why*, things happen as they do. Although science too speaks about the *why* or the causes, the "general causes and principles" of the philosopher are "higher" and "more ultimate" than the causes and principles that science reveals to us. There are two very different forms of activity now go under the name of "philosophy": one is essentially rational and critical, with logical analysis (in a broad sense) at its heart; the other (represented by Heidegger, for example) is openly hostile to rational analysis and professes to arrive at general conclusions by a phenomenological intuition or hermeneutical interpretations.

Aristotle considers philosophy as "the first and last science"—the first science because it is logically presupposed by every other science, the last because deals with reality in its ultimate principles and causes. He defines it as follows: "There is a science which investigates being as being, and the attributes which belong to this in virtue of its own nature. Now, this is not the same as any of the so-called special sciences, for none of these treats universally of being as being. They cut off a part of being and investigate the attribute of this part" *(Metaphysics* 1003a18-25).Descartes' distinction between mind and matter made it appear that there could be an inquiry into "the inner world" which would be wholly distinct from inquiries into "the outer world." Philosophy came to be thought of as running parallel to physics—the science of man as contrasted with the science of nature. Some of the later philosophers consider that the task of philosophy is to 'coordinate the most important general notions and fundamental principles of the various sciences.'

This unifying and coordinating activity is differently considered in different periods and places: in the medieval period it was done making philosophy theo-centric (God becomes the principle of coordination), in the modern period it was carried out by an anthropo-centric philosophy (a human activity by which the human spirit comes to an awareness of its own potentialities), in the English-speaking countries the coordination is done through the analysis of language. Thus there is no unanimity in the way task of philosophy is considered. But it is generally accepted that the task of philosophy cannot be reduced to that of a mere science, and that

philosophy has a priority and primordially in comparison to the sciences.

The Concept "Artificial Intelligence"

Artificial Intelligence (AI) is a branch of *Science* which deals with helping machines finding solutions to complex problems in a more human-like fashion. This generally involves borrowing characteristics from human intelligence, and applying them as algorithms in a computer friendly way. A more or less flexible or efficient approach can be taken depending on the requirements established, which influences how artificial the intelligent behaviour appears. AI is generally associated with *Computer Science*, but it has many important links with other fields such as *Maths*, *Psychology*, *Cognition*, *Biology* and *Philosophy*, among many others. Our ability to combine knowledge from all these fields will ultimately benefit our progress in the quest of creating an intelligent artificial being. AI is one of the newest disciplines. It was formally initiated in 1956, when the name was coined, although at that point work had been under way for about five years. However, the study of intelligence is one of the oldest disciplines. For over 2000 years, philosophers have tried to understand how seeing, learning, remembering, and reasoning could, or should, be done. The advent of usable computers in the early 1950s turned the learned but armchair speculation concerning these mental faculties into a real experimental and theoretical discipline. Many felt that the new ``Electronic Super-Brains" had unlimited potential for intelligence. ``Faster Than Einstein" was a typical headline, but as well as providing a vehicle for creating artificially intelligent entities, the computer provides a tool for testing theories of intelligence, and many theories failed to withstand the test. AI has turned out to be more difficult than many at first imagined and modern ideas are much richer, more subtle, and more interesting as a result. AI currently encompasses a huge variety of subfields, from general-purpose areas such as perception and logical reasoning, to specific tasks such as playing chess, proving mathematical theorems, writing poetry, and diagnosing diseases. Often, scientists in other fields move gradually into artificial intelligence, where they find the tools and vocabulary to systematize and automate the intellectual tasks on which they have been working all their lives. Similarly, workers in AI can choose to apply their methods to any area of human intellectual endeavour. In this sense, it is truly a universal field.

Philosophy and Artificial Intelligence: The Nexus between new and old artificial crossroads.

World-wide network of connected objects with the possibility of exchanging information with each other, the IoT still raises many questions relating to interfaces, design and the user. Studies on the subject, for example by Magrani[3], Atzori[4]et al, show that the number of IoT events and publications is growing. It is possible to verify public interest in the topic by consulting its evolution through Google Insights with the keyword internet of things. However, there is still a need for studies involving issues of ethics, privacy, legislation, and interoperability, which are important indicators of the breadth of the phenomenon.

Numerous international organizations are carrying out studies and research into the creation of general architectural standards and functionalities for the development of application solutions for the new, large-scale markets that characterize what we can understand as the "thing". The frontiers of discovery are still open. According to Faccioni Filho[5], this new "network of objects" offers a multitude of new solutions in preparation for a near future that will integrate different technologies and social fields in a diversity of elements and areas of coverage. There is still a need for greater clarity as to the consequences of its social impacts.

The prospect is that of connecting an extensive network, with Smartphone's and industrial robots and other connections that require high performance and structured connectivity. It has established itself as a fluid and necessary ecosystem for tackling major problems around the world. Its use has made headlines in the media all over the planet. One example is the use by academia and science of ChatGPT, an artificial intelligence created by a US research laboratory called OpenAI, whose architecture is based on a neural network called Transformer specially designed to deal with text. The name ChatGPT is the acronym for an algorithm developed using neural networks and machine learning, which focuses on improving virtual dialogues by offering users simple ways to chat and get answers by cross-referencing the data available on the internet. According to Fleck et al[6], "a neural network is a system designed to model how the brain performs a particular task, implemented using electronic components or by propagation simulation in a digital computer". A neural network is capable of developing learning processes using parameters set by the stimulation of the environment in which it is inserted. In other words, the learning process takes place because, according to Haykin[7], its functioning resembles the human brain in two basic aspects: a) knowledge is acquired by the network from its environment, through the learning process; b) connection forces between neurons (synaptic weights) are used to store the acquired knowledge. Thus, it undergoes modifications when it responds to stimuli from the environment, as is the case with the human learning process. If the parameters are free, learning will be creative, and AI is capable of contextualizing facts, writing texts, lyrics, poetry, short stories, programming codes, recipes, etc.

This process of authoring texts has mobilized the international scientific community, which is concerned about the ethical issues surrounding the use of AI. In this context, this article aims to problematize the concept of artificial intelligence and the heated debate about whether it is possible to duplicate human intelligence at the current stage of technological development. According to Childs[8], we are witnessing the return of the tension between technological advancement and human experience, which is not new. It is a problem that has been discussed with greater emphasis since the 19th century, when the first major reactions against technology were recorded. In the 20th century, authors such as Heidegger[9], Jacques Ellul[10], pointed out criteria for assessing the rupture between humanism and technology, offering indispensable and

fundamental philosophical tools for understanding the meaning of technology in modernity, since the transformations have been profound and at an accelerated pace. The debate is raging about the implications for various fields of knowledge of artificial intelligence "replacing thought" and impacting on scientific and technological development, as well as the educational bases for building the future. We understand that the problems of the impact of the technological world on social construction require a multidimensional and interdisciplinary understanding involving analysis by generalists from the social sciences, Universalists, from philosophy, computer science and the emerging cognitive sciences. To adequately problematize the issue of techniques versus humanism requires approaches based on multiple fields of knowledge. Thus, in this article, the discussion is based on the double corpus of discourses: the conceptual and the ethical. The methodology used for the study was based on the following descriptors: Artificial Intelligence and Philosophy; Artificial Intelligence; Humanism and Technology; Touring Test; Chinese Room; Internet of Things. Harari[11] brings us elements of the Cognitive Revolution of Homo Sapiens. He provides us with a brief history of humanity with important elements for understanding the current Scientific Revolution, such as the most significant inventions of the last century: combustion engine, electricity, nuclear energy, the three-dimensional structure of the DNA molecule, nanotechnology (control of matter on a molecular and atomic scale), pharmaceuticals (molecular modification and principles of optimization of (prototype compounds), general advances in research involving molecular biology and genetic engineering, among others. According to Harari[12], human beings have always been architects of innovation since the cognitive revolution. Humans have created imagined social orders with successive inventions including the agricultural revolution, writing, the invention of currency, the creation of empires, the systematization of science, capitalism, the engine of great inventions. Since the 19th century, new epistemological paths have been under construction in a historiography based on research that has generated anchor points and clarifications on the various historical contexts that have generated the relationship between technology and humanism. Finally, we have arrived at the current 5.0 society, which advocates repositioning technologies for the benefit of human beings. In other words, human beings are at the center of technological transformations. Usage of AI to enhance a social ecosystem where human skills are geared towards promoting inclusion, ensuring sustainability, and improving the quality of human life. Is this a recovery of the Aristotelian vision of the good life (ethics) and the common good (politics) or a recovery of the currents of contemporary practical philosophy (utilitarianism and liberalism)?

Since Touring[13], the concept of AI has been popularized in many fields of knowledge, with philosophy having an impact of diverse reflections. For methodological purposes, the researchers will present some concepts that are essential for the purposes of this work. The researchers understand what a concept is

when it is possible to conceive a history through it. This means that not every word can be transformed into a concept. The author argues that, in a simplified way, we can admit that each word refers to a meaning and, consequently, to a content. However, concepts have limitations and boundaries. The authors take the understanding that the concept of AI goes through historical metamorphoses.

That being said, the researchers look for the concept of Artificial Intelligence on the border between the exact sciences and philosophy, to explain the methodological route that underpins our approach to clarify controversies at the heart of a possible theoretical debate. As a linguistic phenomenon (Habermas [14], the concept can transcend the theoretical dimension to act in the understanding of facts in concrete reality.

There is no concept without a relationship to a given context. Thus, the authors searched for the concept of AI in the double corpus: computer science and philosophy. In his text Computing machinery and intelligence, Alain Touring[15] presents his "game of imitation" in which he shows the roots of the concept of AI, raising a question that still fuels debates among philosophers of mind: "Can machines think?" This basic question still intrigues philosophers today, since there are many understandings about what kind of intelligence can be produced by algorithms run by electronic artifacts. The so-called Analytical Machine designed by the mathematician, Charles Baddage and Ada Lovelace, brought about the first theological and philosophical inquiries into the fact that thinking is an exclusive act of the human soul.

Touring[21]presents seven arguments and nine objections to the so-called Touring Test or The Imitation Game, in which he tests the ability of a machine to produce a narrative in the same way as a human being. In a nutshell, the game consisted of placing three participants in isolated rooms: a man, a woman, and an interrogator, to discern by written answers, without visual or verbal contact, who is a man and who is a woman, based on questions that the interrogator can ask. In the process, a communication channel with a keyboard and a screen is used to generate the result, which today would be a type of instant messaging tool. One of the interrogated was swapped for the machine and the interrogator had to distinguish between the person and the machine, which happened. For Touring, it would be difficult for a person to imitate a machine, above all because of its inability to perform complex calculations.

A machine can perform, for example, billions of calculations and carry out tasks with small margins of error. This ability is very difficult to prove with human beings. Given the various possible entrances to the discussion on this subject in the field of philosophy, we consider the studies carried out by Jonh Searle[16]as an entry rhizome, when he substantiated the criticisms of Touring's concept of AI, through his Chinese room argument. Searle[17] published the article *Minds, brains, and programs*, which had great academic repercussions. In his critique, he analyzes what he calls formal processes in which the machine performs tasks based on information sent to it. He claims that even if the answers are positive, the machine is still incapable of understanding the information it

is dealing with. The basis of Searle's critique is to challenge Touring's arguments that machines can think, believing that machines are only capable of simulation, have no intentionality and that it is impossible to duplicate the human mind. One of the famous examples of the Chinese room provides evidence that machines can be programmed to perform formal manipulation of symbols, simulate conversations and responses, but they do not have cognitive states like humans. Humans are different. The author raises the assumptions that programs are totally syntactic, and minds have a semantic capacity: "Minds are semantically in the sense that they have more than a formal structure, they have content". The fact of projecting simulations does not guarantee that machines have a mind. Searle's syntactic abilities involve the concept of intentionality, i.e., the ability of the mind to represent objects, understand meanings, present beliefs, and desires in relation to meaning, which is not purely formal. Intentionality can be both original (only humans and animals can desire) and derived (written descriptions and sketches). Dennet[18] criticizes Searle's concept of intentionality by arguing that if an AI is equipped with mechanisms that allow it to interact with the environment, it can learn through interaction. In broader terms, Searle's critique targets the concept of Artificial Intelligence, which he calls Strong Artificial Intelligence – (SAI). Roughly speaking, SAI was the concept coined by Searle to describe a field of understanding in which it is possible to simulate the human mind using a computer model. For the author, consciousness and thinking are biological and uniquely human phenomena that cannot be duplicated, but only simulated. What Searle sought was an attempt to abstractly have an experience, based on his own theory, to elucidate the question posed by Turing: "Can machines think?"

Many objections to Searle's Chinese room test have been raised by various authors of analytic philosophy such as Roger Penrose[19]et al, among others. These criticisms were both of a logical-formal nature and of the author's understanding of fundamental concepts and precepts in the fields of computing and cognitive sciences, pointing to Searle's misunderstanding of Touring's claims, above all because he confused simulation with duplication. Since Touring and Searle, encouraging advances in technology have brought significant elements into philosophical discussion about the nature of the mind, offering various possible ways of enriching this question. Putnan[20] already raised the possible association between Touring's machine and the human being. Authors such as Fodor[21] offer critic's possible alternatives to clarify the problem of what types of evidence are still needed to justify an analytical interpretative choice about human cognition and the possibility of its duplication by machines.

It is also important to note that these questions are being analyzed in the field of culture. Literary and fictional works such as Mary Shelley's novel *Frankenstein* and Aldous Huxley's *Brave New World* are well-known narratives that express the fear of scientific progress in which human relationships are overtaken by technological

development. The authors created dystopian representations of the world with fears about the visible consequences of the relationship between technology and humanism.

Human Intelligence and Artificial Intelligence

Multiple questions raised by philosophers, linguists, and sociologists, for example, take the concept of intelligence as an important parameter. Considered complex, it is still in full development. It varies according to different fields of knowledge, for example: psychology (ability to learn and relate), biology (ability to adapt to new habitats or situations). The debate is open and expanding. Materialist conceptions, for example, deny the possibility of reducing the mind to matter. Thomas Nagel[30] in his article "What is it like to be a bat", in the contemporary philosophy of mind debate, raises the thesis that every mental event or phenomenon can be reduced to a physical event or phenomenon. Nagel denies the irreducibility of the subjective to the objective, in other words, he denies the possibility of explaining our consciousness based on physical phenomena. He maintains that it is impossible to explain the mind from the body. At stake is the definition of what thought is. The project to create an AI that thinks is a purely behaviorist experiment. Many objections to this question are important so that we can follow the research into understanding that a machine can behave similarly to a human being. It can be admitted that this machine can think or have an interior life. The concept of artificial intelligence is complex. Among many approaches, the European Union document, *Ethical Guidelines for Trustworthy AI*, considers that an AI has the capacity to make decisions and adapt its behaviour European Commission[22]

According to Nicolelis[23], AI is not intelligent, since intelligence has to do with solving a specific problem, namely the survival of biological species. Therefore, intelligence is characterized by a type of learning specific to biological organisms to survive. Apparently, this will not be a problem for a machine, or at least it would be something they will not yet have in mind, hence the conclusion that it couldn't be intelligent. Nicolelis[23]presupposes a concept of intelligence and learning derived from a biological view of the phenomenon, but the subject can be dealt with from a more general point of view, such as that according to which intelligence is a human capacity to process and produce information to solve problems and learning is a modification resulting from a stimulus from the environment, in a more general way.

According to this view, learning can mean the ability to dynamically adapt behavior, in a non-deterministic way and susceptible to unexpected behavior. On the other hand, Gardner[24] defines intelligence as the human capacity to process and produce information to solve problems or produce products that are important to a given culture and cannot be measured. It is the human potential to process information and can be activated or not depending on the cultural configuration.

It is made up of the set of skills an individual has for solving problems. If human beings accept that machines can solve complex problems not solved by human beings, this debate has just begun. Many experiments

still need to be carried out to make other interpretations possible.

It is through Machine Learning that computers are acquiring new skills. Machine learning techniques allow the computer to learn by example, in other words, to learn from data. Machine Learning has become key to putting knowledge into computers. Humans have a lot of intuitive knowledge, which can't easily express verbally. We don't have conscious access to this intuitive knowledge. Without a formal understanding of this intuitive knowledge, it is not possible to write programs to represent it.

So, what's the solution? The solution is for the machine to learn this knowledge for itself, in a similar way to how human beings learn. In the last two years research have shown a return to the debate on the question posed by Touring[36] can machines think? To broaden this debate, we ask: can machines learn?

Below are some ethical and legal issues in the debates involving Artificial Intelligence (AI)

Legal Normativity applied to AI

So far, the researchers have tried to address philosophical debates and conceptual clarifications about whether artificial intelligence and human intelligence can be considered correlated, as well as whether machines can think or only simulate human learning. The authors are convinced that there is still a lot of research to be done into the consequences of machine learning in the context of current scientific and technological advances. The debate is not new, but in the last decade it has become more urgent. Therefore, let us now look at the other essential aspect to be considered in this discussion, which is a normative issue. At least two paths open in this regard.

1. The first refers to how robots or AIs should be treated by humans, despite the certainty of being able to predicate them something like intelligence, learning, consciousness, will, intentionality, freedom. This could be done based on the concept of person which dispenses these qualities as already practiced by our normative systems in relation to babies or human beings with dementia.

This perspective can also be extended to artificial agents, as has already been done for corporate legal entities in the legal field. In this sense, it is worth analyzing whether objectives such as those envisaged by Noorman[25]are ethical: "Robots are not going to replace humans, they are going to make their jobs much more human. Difficult, humiliating, demanding, dangerous and monotonous - these are the jobs that robots will do".

2. The second path is one that seeks to regulate the use of AI ethically and legally, with a human perspective in mind. This perspective is based on the massive use that is already being made of AI, and which tends to grow, a use with impacts for humanity. For this perspective there is no need to have a deep understanding of the phenomenon of artificial intelligence, either in the sense of knowing whether it is intelligent, or whether it can learn, since a more pragmatic perspective can be taken, given that there is already sufficient evidence of important impacts

to demand normativity. The European Union document, *Ethical Guidelines for Trustworthy AI*[39], lists concerns of high impact and uncertainty, also lists concerns regarding risks, the possibility of involuntary, non-consensual and mass identification, as well as the possibility of localization; the development of human like robots, which can generate emotional bonds; the possibility of classifying people in all aspects, creating registers; the creation of autonomous lethal weapons.

As has been said, this path takes shapes from the human perspective, something that can be seen in the research carried out by MIT's Moral Machine website [https://www.moralmachine.net/hl/pt], whose aim is "gathering a human perspective on moral decisions made by machine intelligence". This led Savulescu, Gyngell & Kahane[26] to propose a procedure for how to use such preferences in political deliberations, including the consideration of what he calls robust philosophical agreements on certain ethical issues. Let's look at examples of concerns in this direction: a) the use of human brain tissue in AI, i.e., kinds of semi-biological devices: "In practice, this means that semi-biological semiconductors will be able to have what Razi calls 'continuous lifelong learning' - which allows the chips to acquire new capabilities without abandoning old learned skills", the incorporation of language models that allow for something like an artificial brain, which boosts learning possibilities.

There is also concern about the use of AI in various sectors, with significant impacts: the use of robots in care services, the use of robots in the public service, such as granting social, the use of AI in tax jurisdiction, the prediction that artificial neural networks will have a significant impact on cultural production, the unemployment that AI can cause (BBC[27] the use of AI in credit and insurance sectors, which can lead to predictions being a kind of self-fulfilling prophecy Amaral[28] thus giving this tool an unprecedented role. The concept of 'machine' helps to understand the phenomenon, since machines can perform complex activities without any human intervention, such as driving cars, flying airplanes, diagnosing diseases, responding to threats. Leben[29] understands that, in this sense, such machines are no more automatic, but autonomous.

These new developments in technology and new possibilities for the application of AI have generated studies in various areas, as well in the field of normativity, whether legal or ethical: "the use of robots by the judiciary requires active human vigilance", because "there is no way to attribute malice or guilt to AI" Copeland[30], AI needs to be controlled, the use of AI can cause harm, Facebook was ordered to pay twenty million reads in moral damages.

Based on the concepts of unacceptable risk, high risk and minimal risk, as well as potential danger, on June 14, 2013, the European Union approved regulations concerning AI. According to this legal regulation,

"The development and use of systems that present an unacceptable risk are prohibited, while high-risk systems are subject to severe restrictions on development, implementation

and use. Regarding low-risk systems, the requirements relate to transparency".

Specifically, concerning the Brazilian case, from a legal point of view, the Constitution can play an important role in regulating the matter, due to its deeply principled nature, which gives it a more extended applicability. In the words of the commentator,

> *"It is recognized that, although the 1988 Federal Constitution is not exactly digital, Brazil already has a legal system anchored in principles and provisions that guarantee the enjoyment of rights such as intimacy, equality, freedom, security, non-discrimination and the free development of the personality which, in themselves, already consolidate a significant support for opposing resistance to some of the deviations mentioned above" (Sarlet[31])*

In addition, it should not be forgotten that in 2018 the *General Law on the Protection of Personal Data* came into force, which, although it does not deal directly with AI, has applicability to aspects involved in the issue. This law, called the Constitution of the Internet in Brazil (2018), lists various human rights that must be respected, and created the National Council for the Protection of Personal Data and Privacy. In any case, the author concludes that "based on constitutional principles, the catalog of fundamental rights and guarantees, as well as the rights extracted from the Brazilian legal system, aligned in a normative constellation by human rights, it becomes possible to act positively in this area".

Ethical Normativity applied to AI

Ethics, like law, must operate based on the possibility of harm or risk of harm, which puts humans in the position of victims, being this perspective defended, for example, by Leben[32], who mentions Asimov's three laws of robotics, the first of which is precisely not to cause harm to a human being. It is in this sense that studies on prejudice in the uses of AI can be understood, such as in the analysis of credit and insurance profiles as well in relation to gender, race and diversity. With this theme in mind, in2023 the journal *Res publica* launched a special issue (V. 29) on algorithmic discrimination and fairness, the purpose of which was to address the normativity arising from the moral notion of equality and its relationship with the use of AI. The European Union document *Ethical Guidelines for Trustworthy* AI is a good example of how a precautionary principle can be used to deal with AI from an ethical point of view. The document aims to "Develop, deploy and use AI systems in a way that adheres to the ethical principles of: respect for human autonomy, prevention of harm, fairness and explicability". This document has characteristics that allow to take it as a model, since it is positioned at the interface between academic-philosophical studies in the area of applied ethics and deliberative processes for establishing specific ethical standards concerning AI, for the bloc of European Union countries: "these Guidelines seek to go beyond a list of ethical principles, by providing guidance on how such principles can be operationalise in socio technical systems". The document is an explanation of the precautionary principle, since it makes extensive use of notions as risk and harm, including the concepts of unacceptable or high risk, which have already been reflected in European Union

legislation in 2023, as pointed out above, with the aim of an "ethical, secure and cutting-edge AI".

The precautionary principle is widely used in environmental law. Arguably, its ancestor is the principle of responsibility proposed by Jonas. According to him, an ethics for a technological civilization require a principle of responsibility, the basis of which, he claims, is fear: "We know much sooner what we do not want than what we want. Therefore, moral philosophy must consult our fears prior to our wishes to learn what we really cherish".

Conclusion

By way of conclusion, it is worth asking whether the modern world of big data, remote simulators, drones, and the internet of things, 5G technologies, and other inventions using AI justifies and ensures a richer and a more humane future. Will a new philosophy trying to reconstruct the world based solely on technological standards emerge, or will we look to the humanist tradition to develop an interpretation of the world where the relationship between humans and machines remains unclear? Case (2023) in an interview to the newspaper *El Pais* argued that technology humanizes us, at the same time as it gives us the status of cyborgs, since it alters our extensions in terms of our ability to speak, hear and move around. She argues that the synergy between information circuits, digital devices, artificial intelligence, communication networks, social movements and individuals, rewrites and signifies ways of living and ways of building the social, because of the symbiosis between science, technology, and society.

The interaction between humans and non-humans brings to the forefront of discussion the concept of the cyborg and the consequent plasticity in human learning through human-machine interaction. By announcing that the person-machine relationship makes us all hybrids, that is cyborgs, she draws attention to the fact that human beings are not Robocop or the Terminator. Human beings don't wear robotic legs or chips, but we are mental cyborgs every time we use digital devices to live in a world that connects the extensions of our ability to see, hear and move, with AI techniques that are increasingly invading the social world to help us solve complex problems.

The desire to make AI learn is not new. Ever since Alan Turing's test, we've known that systems can be intelligent. Even today, AI systems undergo these tests to verify their ability to learn. So far, it has not been established that humans' ability to learn has not yet been achieved. Research into the learning capacity of machines continues to adjust machines to learn like humans. Although we don't yet know for sure how humans learn, there are already efficient algorithms for "teaching" specific tasks to computers. AI is already learning to drive cars (Google and Tesla cars). But there is still a lack of ethical guidelines and legislation to produce tests in real environments and risk situations. Studies are needed on the social and ethical impacts of AI, as well as studies on its risks and benefits in the development of human actions. Meanwhile, the debate is permeated by both unfounded fears and real problems. AI can have both good and bad impacts. AI can prevent human beings from being exposed to

dangerous tasks that can be carried out by machines without having to stop dealing with exciting and enjoyable creations. The benefits are already numerous: improvements in the field of health; Natural Language Processing; clean energy production; fraud monitoring; safer means of rapid transportation; more productive governance. It must be recognized that AI also has negative impacts: unemployment and increased social inequalities. As pointed out previously, its use involves numerous ethical and moral issues, including: the use of powerful and automatic weapons; invasion of privacy; lack of transparency in the use of data, among others. It is known that the concept of AI is disputed, both in terms of the predicates that make it up and the scope and limits of its use. This text has attempted to make normative considerations, both from the point of view of how humans should treat AI, and in relation to how humans should protect themselves from possible harm arising from the use of AI. In both perspectives, legal or ethical standards can be established.

Recommendations

1. **Interdisciplinary Approach:**

Future research should adopt an interdisciplinary approach, integrating insights from philosophy, computer science, and cognitive science and other relevant fields to better understand the complex relationships between AI and human intelligent.

1. **Ethics and Responsibility:** As an AI system becomes increasingly autonomous, it's essential to develop robust ethical frameworks that address issues of accountability. Philosophers and AI researchers should collaborate to establish guidelines for AI development and deployment.
2. **Human – AI collaboration**: Investigate the potential benefits and challenges of human –AI collaboration, exploring how AI can augment human capabilities while preserving human agency and dignity.
3. **Future of work and education:** To explore the implications of AI on the future of work and education, investigate how philosophical perspective can inform strategies for workforce development, education and lifelong learning.
4. **Epistemological Implications:** Investigate the epistemological implications of AI, exploring how AI challenges traditional notions of knowledge and understanding.

These recommendations can serve as a starting point for further research and exploration at the intersection of philosophy and artificial Intelligence.

References

Amaral, A.J. Governmentalidade algoritmica e novas praticas. 2021

Atzori, L. The internet of things: a survey in computer Networks, 54(15) 2010, P.287-289

BBC News report in Brazil concerning Artificial intelligence. 13/07/23. 2023

Childs, P. Modernism. New yokr oxon. 2017

Copeland, K. Artificial intelligence: a philosophical introduction. Blackwell Publishing. 1993

Dennett, D.C. Chinese room from a logical point of view. New tork: Oxford university press. 2016

Ellul, J. Le systeme technician. Paris. 2012

Faccioni F.M. Internet of Things: Artificial intelligence and Philosophy. 2016

Fleck, L, et al. Principles of artificial intelligence, Vol. no 1(13) 2016 P.47-57

Fodor, J. The Mind-Body Problem. In: HEIL, J. Philosophy of Mind: A guide and Anthology. Oxford university press, 2004 P.168-182

Gardner, H. Human Intelligence and machines. 2000

Habermas, J. La technique et la science comme ideologie. Paris: Gallimard. 1968

Harari, Y. Minds and machines. What is right and wrong about the Chinese room argument. 2020

Haykin, S. Kalman filters in: Kalman filtering and neural networks, john witley and sone. 2011 P.1-21

Heidegger, M. The Question concerning technology. Nova York: Harper. 1977

Leben, D. Ethics for Robots: How to design a moral Algorithm. Routledge. 2018

Magrani, E. Artificial Intelligence and Human Knowledge. 2019 P.23-27. Magrani, E. Opcit 2019 P.29-31

Nicolis, M. Artificial Intelligence and Philosophical inquiry. 2023

Noorman, M. Computing and Moral Responsibility. The standard encyclopedia of philosophy (spring 2023 edition) 2023

Oxford English Dictionary 2023

Penrose, R. Consciousness, computation and the Chinese room. New York: Oxford university press. 2007

Putnam, H. The mental life of some Machines. In: Mind, Language and Reality. Cambridge. Cambridge university press, 1975 P. 408-428

Sarlet, G.B.S. Intelligence artificial in a context. 2021 P.273-305

Savulescu, J, et al. Collective Reflective Equilibrium in practice (CREP) and controversial Novel Technologies. Bioethics. 35. 2021. P.652-663

Searle, J.R. Twenty –one years in the Chinese room. New York: Oxford university press. 2007

Turing, A.M. Computing Machinery and intelligence. Mind, Lix(236) 1950 P.433-460

Turing, A.M. The chemical basis of Morphogenesis. In: The Royal society. 237(641) 1952. P.37-72