Cognitive Priority over Ethical Priority in Artificial Intelligence: The Primordial Philosophical Analysis in Artificial Intelligence

Zapata Flórez A*

Faculty of Engineering, University of Antioquia, Colombia

*Corresponding author: Antony Zapata Flórez, Faculty of Engineering, University of Antioquia, Cra 66 #61-56, Itagüí, Antioquia, Colombia, Tel: +573154048606; Email: antony. zapata@udea.edu.co

Abstract

The general idea that we have of artificial intelligence (AI) consists of the belief that machines will be able to develop conscious thoughts such as those possessed by human beings, and, as computing advances, such thinking will also advance until intelligence to surpass the human being, with which the advancement of AI represents ethical risks in the future. In reality, such a belief hides a cognitive assumption, which assumes that computational engineering explains human intelligence through the mind-computer metaphor. According to this assumption, technology explains cognition, and philosophy, through ethics, reflects on the impact of said technology. However, in this article, I contradict such an assumption and defend that the philosophy in AI is not reduced to the ethics that is present after the use and impact of AI in the world. I intend to expose that a good ethics of AI is the one that reflects on the appropriate risks facing AI, and for this, philosophy, beforehand, must make a cognitive analysis about the possibilities that computing has to create intelligent machines, namely, whether or not the mind-computer metaphor makes sense. My thesis consists in defending that the philosophical analysis about AI must be carried out both on a cognitive level and on an ethical level, but that the philosophical priority in the cognitive analysis over the ethical priority, since the ethical risks of AI depend of the possibilities of technology, and only the cognitive approach can account for this.

Keywords: Artificial intelligence; Cognitive priority; Ethical priority; Philosophy; Intelligence

Introduction

It is commonly assumed that computers are advancing exponentially and that in the future their intelligence will surpass human intelligence. Computers are supposed to have a mind or have to do with the mind at a level equal to the human being, that is, they think like the human being. The difference between human beings and computers is that humans have disadvantages compared to computers, as human beings become physically and organically exhausted, get sick and die. A computational system can function without those factors,[1] thus it is destined for intellectual superiority compared to humans, according to the general belief promoted by computational ideology. So, according to this perspective, the risks are inevitable, that is, the human being

---

1   Such factors that condition human existence are considered limitations of the human being that limit him to a higher stage evolution according to the transhumanist movement, whose philosophy is permeated by computational ideology, so that the human being can access improvements through modifications. computational.

Cognitive Priority over Ethical Priority in Artificial Intelligence: The Primordial Philosophical Analysis in Artificial Intelligence

Philos Int J

in the future will be at the mercy of a superior intelligence, for which the call for the development of an ethics on a par with technological development is necessary.

Naively, the assumption of the *superintelligent* machine is taken as true, which implies an ethical concern vitiated by this cognitive ideology illustrated by the mind-computer metaphor, which posits that the human mind is nothing more than computational information processes, as pointed out by Anibal Puentes [1] "Research on AI has started from an assumption, not necessarily true, that there is an analogy between the intellectual functioning of man and machine. [...] It should be clarified that the mindcomputer analogy is not physical, but functional." (P, 357). So it is assumed that a computer can do the same as a human being, and even better. Ethics in this sense would be aimed at a reflection directed to the future risks of an artificial Superintelligence. You never need to question, for example: what if such an ethic is inadequate because the cognitive assumption turns out to be false? What if ethical reflection is wasted on empty risks motivated by imaginary technological promises? How can we guide a realistic and sensible ethics on the specific problems generated by AI? To guide a good ethics of AI, a theoretical analysis on the cognitive principles of AI is first needed, in order to justify or dismantle the most radical promises of computer technology. Can computers be as intelligent as or more intelligent than human beings? Does the mind-computer metaphor make sense? Hence, the ethical risks of AI depend on the cognitive possibilities of AI. I then argue a cognitive priority over an ethical priority of AI.

This article is developed in three parts:
1. Start by exposing how the traditional idea of AI is developed, its cognitive sense and the exaggerated promises from the computational possibilities.
2. I will expose the ethical risks that arise from the belief of the radical promises of AI, such as the concerns of an artificial superintelligence and the digital consciousness that transhumanism promises.
3. I will expose the philosophical criticism that analyzes the cognitive thesis of AI and the shortcomings of the mind-computer metaphor on which it is based, revealing that the cognitive impossibilities of computers lead to rethinking a more grounded ethics according to the true impacts that they have computers on humanity to this day.

The Cognitive Proposal and the Technological Promise of Artificial Intelligence

AI is a technological project which promises to build machines that can perform any type of task. In other words, based on different techniques, AI machines can solve any type of problem that human beings pose. The promise of such

artificial systems is based on a philosophical assumption about intelligence, which gives theoretical meaning to what engineers do in laboratories, since AI researchers have appropriated the mission of explaining intelligence[2] in general.

The idea that machines can be intelligent, translated later to whether computers can think, is not just a free idea that fell from the sky, but it is the logical conclusion of a particular philosophical approach about cognitive processes. Before the very existence of the AI project, there was a philosophical position on intelligence that gave birth to and nurtured this technological possibility. In the following, we will analyze how the functionalist philosophical idea about cognition configures the philosophical thesis of AI about intelligence as computation.

The philosophical theories that in the middle of the 20th century discussed whether the mind constituted the causal principle of intelligent behavior were *behaviorism* and *materialism* of the theory of identity (I.T.).[3] Both proposals suffered from major explanatory limitations on human behavior. On the one hand, behaviorism denied the existence of an internal rationality of intelligent organisms that caused their actions, something like internal mental states, since human behavior could be reduced to the instinctive dispositions of an organism that reacts to the stimuli of the environment. ambient. In other words, the cause of human behavior is not mental, based on a rational discernment process, but based on automatic processes of an organism causally affected by the outside world. On the other hand, cerebral materialism, in contrast to behaviorism and, in order to rescue the internal processes of a rational mind, proposed the explanation of mental states based on the parallel correlation they have with neuronal states. So the mind was reduced to the rational possibility that arose from the neuronal microproperties of brain biology [2].

However, a third philosophical position known as *functionalism* is put forward to overcome the automatic externalism of behaviorism and the reductive materialism of IT. For this position, unlike behaviorism, there is an internal rational discernment, which are considered as cognitive processes that cause the actions of the system. But unlike materialism, such mental processes have no substantial relationship to the chemical micro-properties of the

---

2   Dreyfus (2007) tells about the rivalries and contempt he found in the AI laboratories at MIT on the part of the engineers: "You philosophers have been reflecting in your armchair for 2000 years and you still don't understand intelligence. We in the Lab have taken over and are succeeding where you philosophers have failed" (P, 247).

3   Mental processes exist as long as they are identical to brain processes. In short, there is a mind because there is a brain.

Zapata Flórez A. Cognitive Priority over Ethical Priority in Artificial Intelligence: The Primordial Philosophical Analysis in Artificial Intelligence. Philos Int J 2022, 5(4): 000269.

Copyright© Zapata Flórez A.

biological brain. As its name implies, this theory posits that internal mental processes can be understood as systematic relationships between fixed functions that influence each other in a causal way. The thought about this or the memory about that, are nothing more than processes with very specific functions that determine the indicated result that the process must reach, for example, what role a certain process must play to produce a thought or a memory. In this way, two characteristic principles of the functionalist proposal are concluded (Searle, 2000):

1. That the actions of the system depend on the functional processes internal to the system and do not depend passively on an external causation, moving away from behaviorism.
2. That these functional processes are not caused by a certain material that physically supports the function, so any other material that can support the function is suitable for the system, moving away from the cerebral reduction of the I.T.

Thus, functionalism proposes that the mind can be understood as functional processes that follow an internal logic to fulfill an end. As well, functionalism raises the possibility that any material that can be configured in such a way that it can comply with this functional process, will be able to fulfill the purpose that is sought, so that if a functional system can execute the logical regulation of the function in question, for example, mental processes that cause certain actions, then that artificial system will have a mind and will be intelligent. In this way we understand how AI has a philosophy that theoretically justifies the possibility of thinking machines. AI then needs the technology that can realize this philosophy and make that promise a reality.

Computation was presented as the technology that could empirically demonstrate functionalist theory. Computers became the devices that, through computer theory, come to materialize computational processes to process information following a logic that regulates certain functions. From digital and computer processes, computers have been able to solve increasingly sophisticated and accurate calculation problems, which for many demonstrates the ability to imitate, match and surpass human reasoning. In this way we understand much better the origin of AI as we know it, since it is born from the combination of functionalist philosophy and psychology (its theory) and computational technology (its materialization). In other words, AI is the immediate product that emerges from the marriage between functionalism and computation. Therefore, the objective of AI in its essence --philosophically-- is to develop psychological processes in computers. This is how Howard Gardner [3] points out: "And most understand that the computer program is a verification of a particular theory about how a cognitive process could operate" (P, 160).

Computational theory and cognitive theory have always had a close relationship from the beginning [4],[4] leading to reciprocal feedback between the two ever since. Alan Turing was convinced of such a relationship in the mid-twentieth century when he raised the possibility of thinking in machines, years before the official birth of AI by John McCarthy[5] and others, given the ability of digital computers to simulate any type of process, among them, the cognitive processes of the human being. In other words, machines can be intelligent like humans because they can replicate their mental processes, and they can do so because people's cognitive processes are computational processes [5].

Turing (1950) proposes three components for the concrete development of machines that can carry out any type of task, and that each component responds to three principles of computational theory, which will also be taken into account and applied in cognitive psychology in the study of the human mind onwards. The computational machinery must be made up of:

1. A central processor (logical rules of the mental process)
2. An executing unit (symbolic manipulation of information)
3. A database or memory (information storage). According to this approach, intelligence in general, both of human beings and of machines, is reduced to these three principles, that is, the mind is nothing more than:
a. An algorithm in which the task to be performed is programmed
b. The computational medium that processes the information quickly and effectively
c. A storage unit large enough to store enormous amounts of information.

In this way, intelligence ceases to be properly psychological, and becomes an engineering problem. Consequently, for the general opinion, the advance of computational technology represents the advance of AI, and with it, the advance of intelligence in general, or if you want, the intelligence of machines represents in principle human intelligence. For committed AI advocates, the issue is not one of theoretical or philosophical principle, rather the issue of AI is technological. The problem is not conceptual but it is a temporary problem in the sense that the progress of time will translate the progress that engineers have to improve technology and can build more sophisticated information processing computational systems in order to enhance the

---

4   Computational and information theory were nurtured by the proposals of analytical philosophy at the beginning of the 20th century. XX, as was the symbolic logic and the explanations about thought and reasoning that were discussed

5   In 1950 Turing raised the possibility of thinking machines, and, in 1957, AI laboratories were inaugurated at MIT, where thinking machines were baptized by McCarthy as artificial intelligence.

Zapata Flórez A. Cognitive Priority over Ethical Priority in Artificial Intelligence: The Primordial Philosophical Analysis in Artificial Intelligence. Philos Int J 2022, 5(4): 000269.

Copyright© Zapata Flórez A.

three computational principles proposed by Turing. In this way, it is believed that enhancing computational principles is already enhancing the principles of intelligence itself, whose assumption has led to configuring the *Transhumanist* movement, from which two promises are raised from the aforementioned computational enhancement, namely: the superintelligence and human enhancement [6].

The Risks and Ethical Concerns of Artificial Intelligence

Over the years, AI continues to generate sensations of all kinds among people based on the technological innovations in which it is presented and developed. AI performs tasks in a faster and more sophisticated way, facilitating the process of certain human activities, even delegating a certain margin of action to machines by people given the level of automation that certain techniques have acquired. It is inevitable for the most impressionable minds not to use their imagination to fantasize about the future of AI materialized and stimulated by fictional stories [7]. Starting from a traditional notion of intelligence in which intelligence is supposed to be reduced to solving computational problems, unsuspecting minds will see the development of AI as the limitless progress of computers towards higher intelligence. The development of AI through different types of computational technologies (starting from computers with programs that process images, text and voice, which can emulate human tasks such as conversing, composing, writing and generating new images, going through algorithms that "learn" through neural networks with Machine Learning, Deep learning and Big data techniques; even quantum computers there; as well as, then research in cognitive sciences, take the computer as an explanatory model of the mind and thought. with superprocessors that perform calculations faster than any common computer) have led us to think that machines are destined to surpass us in reasoning and do all human tasks in one go better shape thanks to his superintelligence.

Machines will be considered super-intelligent in the future when they carry out such exact rational operations through a fast and effective computation process that any human being would be far from being able to carry out by himself with his intellectual capacities [8]. The AI will be more intelligent than the human being because, over time, the computational principles on which its intelligence depends will be enhanced, that is, more sophisticated algorithms, more efficient information processors and coarser memories with the capacity to store large amounts of information and all kinds of knowledge. For the defenders of a conception of the computational mind, machines are destined for super intelligence since their cognitive processes are purely rational, precise and objective, that is, they will never be wrong, while, on the other hand, the human being is doomed.

To his animal intelligence, because given his biological condition, human rationality is vitiated by his organic needs, imprisoned in his corporality, subjected by fatigue, illness and death: "it can be affirmed that for transhumanism the biological body is something despicable, given its limitations and weaknesses, including susceptibility to disease, and ultimately, death" (P, 380) [9]; without saying that the human being is a slave to his emotions. In short, human intelligence is doomed to error, disagreement, and confrontation. Hence, the machines are destined to overcome us since they do not have those "ballasts" that characterize the human being, leading them to see us as obsolete systems and are forced to replicate, building even more intelligent machines, reaching a point without return of super intelligence, whose threshold has been called Singularity [10].

This panorama thus fantasized by the defenders of the explanation of the mind as a computer, and, hence, defenders of strong AI, raises concern in the general public since the supposed Artificial Super Intelligence (SAI) represents a risk for humanity itself [11]. From this technological futurology, precautions have been proposed to be taken from different fields, both from the engineering field and from the ethical field. From this second, which is the subject that interests us, ethical concerns arise such as: is it correct to create super-intelligent machines that surpass us in all fields, leaving the human being obsolete? Is it right to create more intelligent machines that can enslave or exterminate us? Can super intelligent machines be considered people? Will artificial super intelligence have rights and duties? Is an ethic necessary for machines?

Technological speculation about how far computing can go in developing intelligent a machine has raised concerns, which, in turn, has created an ethical priority for future risks. The creation of an official reflection on an ethics of AI in order to design regulations to regulate the use and development of AI and the impact it can have on humanity, is a necessity both for good and for bad: for good in the sense that importance is given to the essential ethical reflection on the role that technology should play in society; for the worse, in the sense that space is given to inadequate or empty ethical reflections that do not deserve any priority. Coupled with this, some other ethical concerns, which enjoy a certain priority without being clear if they are correct or necessary, also come from the same source as the risks of super intelligence, namely, the computational development of cognition, this time not oriented to machines, but oriented to human beings as prophesied by the trans humanist movement.

As its name implies, *transhumanism* promotes the evolution of the human being to an improved version of himself by transforming his current biological conditions through technological advances. Transhumanism places

Zapata Flórez A. Cognitive Priority over Ethical Priority in Artificial Intelligence: The Primordial Philosophical Analysis in Artificial Intelligence. Philos Int J 2022, 5(4): 000269.

Copyright© Zapata Flórez A.

a blind faith in the promises that are glimpsed from the technological advances applied to the human being himself and that will make him transit to a new species that is no longer human, but post-human.[6] According to the futuristic movement of trans and posthumanism, the progress of technology can be applied in various fields and aspects of the human being. Some applications, improvements or transformations would be: the application of nanotechnology for the organic manipulation of people; the development of drugs that will be able to control any emotional alteration or psychological pathology [6]; genetic manipulation for biological improvement; the computational progress of brain prostheses that allow an interface with the neural network of people and thus make up for some cognitive deficiencies, or even improve and enhance these cognitive faculties, such as having better memory, better reasoning ability and better creative abilities; In this line, the creation of robotic prostheses that complement some bodily deficiency or come to enhance some anatomical or organic part with what has been called the cyborg or cyborgization [11], where the natural and artificial difference is diluted; even more radical still, digitally scanning and encoding people's consciousness so that they can make computer copies to download onto a computer and go on to exist in cyberspace and achieve immortality [12].

Like superintelligence in machines, the promises of human superintelligence, its improvement, even the creation of a new human being, have aroused the concern of people in general, activating the need for an ethical discussion about these possibilities. Some concerns center on the social risks that certain alterations or improvements of some human beings can be caused by such technological advances. Some of the most notorious concerns refer to the social discrimination that the existence of certain human beings with better abilities or higher intelligences would cause, dividing humanity into certain types of castes, that is, the evolved people who managed to pay for the technological service of its improvement, and people who cannot afford such technology and maintain their current capabilities. Not to mention the new human beings who download their consciousness into computers or other bodies, so that immortality, and not death, is a social concern for future humanity. Now, leaving aside the futuristic exaggerations of the computational ideology, today some concerns about the technological application on human biology are already

---

6   Technological intervention in human life has configured new debates at an ethical, social and political level, since transhumanism alleges that it is ethical for the human being to agree to its modification to overcome death and guarantee the evolution to a new species, the post-human. human; At the same time, it is alleged that it is not ethical to guarantee the advent of a new port-human species, going against the interests of human existence itself (Dieguez, 2020).

discussed and standardized in the legal field, such as the cases of the regulations and prohibitions of the alteration genetics and human cloning [9].

It will be thought that, because of the risks involved in fulfilling the promises of AI, superintelligence and trans and posthuman enhancement, analysis and philosophical reflection are activated given the ethical concerns to be taken into account in these fields. It is commonly believed that the superintelligence of machines or digital consciences are certain facts in the future that the advance of computing will achieve, as if technology and engineering already had possession of the final explanation to the mystery of human cognition and intelligence; coming to identify the problem of intelligence to technical computing issues, and that it is only a matter of time and technological progress for machines to think. Along with this belief, in turn, philosophy is identified with ethics, as if the only analysis that philosophy could deploy on AI is limited to the ethical implications of technological impacts. However, this is the assumption that I intend to challenge and discuss in this article, since I propose that the priority of the Philosophy of AI (PHAI) is not ethical, but rather cognitive. This thesis is justified on the basis that the explanation presented by computational engineering on intelligence from which it sustains its most extravagant promises such as SAI and digital consciousness are not true, but rather, it is a philosophical assumption that is based on in a discredited metaphor that has lost influence over time, and I mean the mind-computer fallacy.

The Priority of Cognitive Analysis in Artificial Intelligence

The thesis I defend is that the most fundamental philosophical analysis about AI should focus on cognitive theory that brings to life the promises of computing rather than worrying ethically about exaggerated futures for no reason whether they are possible. So the initial philosophical analysis of AI focuses on the possibilities of computational technology, the type of relationship it has with human intelligence, and consequently, analyzes both the scope and limits that AI has in cognitive matters, before entering to reflect on the impact of AI. As Agustí-Cullell (2022) states [13]:

> "If what was predicted were to come about, humanity as a whole would end up being enslaved by those who possessed these intelligent machines. Behind it all there was, and still is, an ignorance of the depth and unity of human intelligence." (P, 2).

In other words, initially the philosophical analysis focuses more on the possibilities of AI than on the impact of AI, that is, the Philosophy of AI (PHAI) is more identified with cognition than with ethics. Contrary to what the

Zapata Flórez A. Cognitive Priority over Ethical Priority in Artificial Intelligence: The Primordial Philosophical Analysis in Artificial Intelligence. Philos Int J 2022, 5(4): 000269.

Copyright© Zapata Flórez A.

computational ideology believes, the cognitive problem of intelligence is not a technical issue that engineering reveals through computational models, since the relationship between intelligence and computation is not technological, but rather a theoretical relationship of principle, namely, it is a philosophical relationship. Precisely, the proposal presented by AI as a technology in which it is intended to demonstrate an identity between mind and computer, is only possible as a justified philosophical thesis based on a particular cognitive position about the nature of cognitive processes. So that such mind-machine identity is not only dismantled from a technological project, but can be refuted from an initial state of its logic, that is, it can be analyzed and refuted from its theoretical principles. AI is above all a philosophical idea among many that presents an image about the nature of human intelligence, just that, and not a technological truth as it has been sold to us by the media propaganda of computational cognitive sciences based on the influence that has had for over seventy years.

In any case, the field of action of philosophy about AI (PHAI) is not limited to the ethical reflection of the social implications generated by the technological impact at a time after its implementation, but rather, in the special case of AI, philosophy is already involved in the theoretical generation and in the explanatory models of the very idea of AI, so philosophy acts analytically both in a theoretical/cognitive moment and in a practical/ethical moment, but above all, primarily, in an initial cognitive moment. Why should be the primacy of cognition over ethics? The thesis of the AI that identifies mind-machine not only poses a technological possibility, namely that the computational machinery possesses a real mind, but that AI poses first of all a philosophical possibility, namely a supported theory, whose fundamental principles are sound. As John Searle (1985) points out: "Unlike many philosophical theses, they (strong AI) are reasonably clear, and admit of simple and decisive refutation." (P, 36). If such a theory does not have cognitive support, its philosophical possibility will collapse under its own weight, and, therefore, its technological possibility will be destined to fail. Indeed, if the technological possibilities of AI such as artificial superintelligence and digital consciousness fail because the philosophical possibilities of the latter two have no theoretical basis, then the social impact they would have would be zero, and, consequently, there would be no why worry ethically about the exaggerated risks of the AI of the future. In this order of ideas, the cognitive priority lies in the fact that the philosophy of AI analyzes the cognitive support of the mind-machine metaphor, unraveling the scope of computational technology, but also its limits in terms of intelligence, so that such reflection it would avoid another subsequent reflection on ethical matters that is unnecessary, or failing that, inadequate. That is the advantage of cognitive priority, which filters philosophical analyzes about the ethics of AI, managing to discriminate the appropriate ethical concerns from the inadequate ones, or the necessary ones from the unnecessary ones.

The primordial philosophical analysis about AI consists of a critical analysis of the concept of intelligence that predominates in computer engineering. The criticism of the computational explanation of intelligence, as we saw it in the first section of this article, refutes the technological possibility in the future, and thus, refutes the ethical concerns that we exposed in the second section, such as the risks of superintelligence and digital awareness. In fact, this has shown us the philosophical development of AI in recent decades, namely, that philosophical criticism has focused on the cognitive proposal of AI, within whose criticism, various authors have refuted the mind-machine metaphor, among whom it is worth highlighting John Searle and Hubert Dreyfus. For these two authors, the mind-computer thesis is simply false, so AI will never be able to generate real intelligence in any of its technological advances and different computational models, whether they are symbolic machines, neural networks, quantum computers or robotic structures. For these two philosophers, from their theoretical positions respectively, they argue that machines perform tasks that appear to be intelligent, but such intellect is based on a sophisticated follow- up of rules programmed in an algorithm, but they do not produce real cognitive processes based on intentional or state states awareness. So AI simulates human behavior through computer code, but it does not replicate it based on its meaning.

Both Dreyfus and Searle accuse computing of not being able to account for the problem of understanding, which consists of defining intelligence as the ability to understand the meaning or content of the information that computers process or store, that is, the Computers execute knowledge about things, but they do not consciously understand the meaning of things. Searle goes in this direction with his criticism of strong AI through his conceptual differentiation between syntax and semantics[7] (Searle, 1980): a problem can be solved in two ways, one by dictating rules that direct the action of a designed mechanism, and another, based on what the information he is facing means. The first way is how mindless machines operate, the second way is how people

---

7 The differentiation between syntax and semantics is done through Searle's famous thought experiment known as the *Chinese Room*. In this experiment, Searle criticizes the Turing Test (Imitation Game) in which a machine pretends to be a human being, now, in Searle's argument, the human being pretends to be a machine, where to give correct answers you can only do it following rules. This shows that even if the semantic content of a message is not understood, it can be responded to effectively, only by accessing its syntactic structure. Thus, computers perform precise tasks but never understand what they themselves do.

always act. Computers depend on the syntactic design of their own making, for their part, human beings depend on the semantic content that things have for themselves. The fact that human beings have biological brains is what has allowed their minds to make sense of things, which machines obviously do not have, and therefore from their syntactic silicon brains it is impossible. that computers originate semantics and that they can understand the meaning of information.

Marvin Minsky [14] called this ability to understand as intelligence or commonsense knowledge, which AI research should take very seriously, since the intelligence of machines would depend on the computational mechanization of said knowledge so that the machines could solve any type of situation, that is, the ability to recognize the contexts in which things are framed. Minsky (2006) [15] illustrates it as follows:

> "That certainly is a tempting idea, for the World Wide Web must contain more knowledge than any one person ever could learn. However, the texts on the Web do not explicitly display the knowledge that one would need to understand what all those texts mean." (P, 179).

This cognitive challenge became known in AI history as the frame problem. For his part, Dreyfus (2002) calls this problem the problem of relevance, which is none other than the ability of an agent to discriminate which, among the vast amount of information and things, are those that are relevant to me. action in a given context. Said ability to discriminate between what is relevant and what is not, cannot be programmed in the logic of an algorithm that has previously been defined for the machine and that indicates the contexts and meanings of things. The relevance and significance of each thing depends on the projects that each agent has in their daily life and carries out in an interested manner [16]. Intelligence consists of the ability to adapt to unforeseen changes in the world and things. Such changes can only be intuited by the body of the agents, which leads to an interested discrimination that the agent has with his projects, and not through the mental and computer representation of a system that already has the order of things fixed and does not can adapt to changes in the environment. The new AI precisely seeks models that are more concrete and embodied and less abstract or rational, as stated by Andy Clark [17]:

> "The 'new robotics' revolution rejects a fundamental part of the classical image of the mind, [...] The problem with the central scheduler is that it is, at its core, highly impractical. It introduces what Rodney Brooks aptly termed a 'representational gridlock' that blocks any quick response in real time." (P, 61).

These criticisms are bad news for the more outlandish possibilities of AI. Well, as the philosophical analysis indicates, the ability to understand the meaning of things due to common sense knowledge, typical of the human mind, is due to the biological reality of human brains and body flexibility to adapt to any change in context. , managing to intuit the relevant information from what is not. For these reasons, common sense would never originate in machines through computational methods such as providing artificial systems with a large memory capacity in which millions of data can be stored together with the design of sophisticated algorithms that tell the machine what to do. what is relevant and what is not. So the ultra-computational claim of transhumanism would be in serious trouble when it comes to stripping away what it considers a hindrance to the intellectual improvement of the human being, such as their biological, bodily and emotional condition, since what cognitive analysis tells us it is that such conditions determine human consciousness, reasoning and intelligence. Common sense depends on such factors. This has been demonstrated by contemporary research in cognition and AI: a shift in the cognitive conception, no longer computational but biological [18]. Philosophically, it is revealed that the theoretical principles of AI have no cognitive support, that, as Dreyfus [19] demonstrates, the mind-machine metaphor is nothing more than an idea that is based on assumptions (biological, psychological, epistemological and ontological)[8] ideologically. accommodated to computational theory.

The deepest objective of AI is to imitate intelligent behavior, and this is none other than human behavior, since the theoretical and technological goal of AI is to replicate the action of the human being when he performs his tasks, properly human, therefore. that the ultimategoal of AI is to reproduce the human mind as an artificial mind. But human intelligence is not reduced to the combination of functional rules that describe a process, such as what is the algorithm of a thought to open a door or drive a car; human intelligence is defined by the understanding of things, that is, to give meaning to things, to the world around him, basically, to become aware of a world. The problem with how AI understands cognition is to believe that intelligence is a property internal to the system, as if intelligence were a hidden property in the processes of the artificial or biological brain. But the problem of common sense, or the problem of understanding things, is outside, in the world, in the relationship of the system with the world. The challenge of AI has been to make a machine understand the contexts in which it finds itself, to realize a

---

8   According to Dreyfus, the computational ideology conceives the human being as a computing system thanks to 4 assumptions: 1, that the brain is a biological computer; 2, whose psychology consists in following discrete syntactic rules; 3, that such rules can be formalized in computer code; 4, so that human intelligence resides within the set of natural processes of the brain.

situation, to realize a world. Machines cannot be intelligent if they do not have a world, if things do not have a meaning of what they are, only there, the machines will behave with their own intelligence, and not as an intelligence derived from the intentions of the programmers who design them.

That is what human intelligence is all about, acting according to what a particular situation means based on the particular interests of each person in their daily lives. Human beings adopt contexts because they have a world, and this world is constituted by the unforeseen and dynamic situations in which things are presented, so that the human being is constantly adapting to the unforeseen environment. So his behavior is always relational to the outside, and not dependent on an internal code of rules in the brain that tells him how to act. Comprehension is not in the already programmed definition of information or of the significance of data, but rather it is a human ability to, as Dreyfus says using Heidegger's concept, Being-in- the-world.

The human being, unlike the most intelligent computer that exists today and may be created in the future, has a world and has a daily life, and therein lies the difference between the human and the algorithm. The algorithm processes information faster and faster and stores more and more vast information, but it does not understand the content of that information for the simple fact of not having an everyday world. The Heideggerian concept for it is that of Facticity or Worldliness. That is to say, the sense of things, human intelligence, is based on his interested daily life, on behaving based on a need, whether basic or organic, or personal emotional. The human being gets up from his bed to learn, do and talk based on a daily project or even a life project. That is the world of the human being, the personal life of each real and concrete person. Human intelligence is above all interest in living and existing in a particular way with an individual meaning, how each person interprets their existence, their being, as Heidegger would say. Machines only process information faster and faster and exceed, but they don't give meaning, because they don't understand contexts, and they don't understand contexts, because they don't have a world, and why don't they have a world? Because they do not have a daily life where the needs and interests of the world motivate them to act and learn skills. The human being has a world, he has existence, intelligent machines do not. This is how Dreyfus (2002) [20] declares it:

> "Thus, research in AI has revealed the bluff of the Cartesian cognitivist. It is easy to say that to account for the equipment nexus, one need only add more and more rules and function predicates that describe what to do in typical situations. However, the real difficulties in AI—its inability to make progress with the so-called common-sense problem of knowledge and its inability to define the prevailing situation,

sometimes called the frame problem—suggest that Heidegger is right. Apparently, the phenomenon of the world cannot be built from meaningless elements." (P, 97).

Cognitive priority exposes what computationally cannot be done and what can be achieved. The computational advance of AI can only reach the simulation of a mind, but not a genuine mind (Searle, 1985). Computers will not be able to be intelligent in the human way, much less superintelligent. Just as it will not be possible for humanity to transcend to a higher intelligence when it manages to get rid of its needy organic body and its irrational emotions by being able to download its digital consciousness in cyberspace, since it is thanks to its emotional condition embodied in a needy organism that the human being can relate (P, 46). interestedly with a world, that is, see things with meaning. As Dreyfus [21] says, it is thanks to the body that we learn, relate and transmit with others: "If we were to leave our bodies behind and live in cyberspace, nurturing children and passing on one's variation of one's cultural style to them would become impossible" (P, 46). Neither an apocalyptic future dominated by conscious machines nor immortal humans existing in digital space, such risks are false and empty like the theses that compare the mind with a computer program from the imaginary reinforced by science fiction and popular culture [13]. The real ethical risks are the ones that should concern us right now given the social effect that real AI is having today.

AI computing technology cannot create artificial minds, but it can create negative impacts on humanity today and even more so in the future if its use is not regulated immediately. The creation of machines that process information precisely manages to emulate behaviors and improve them in highly specialized tasks such as expert systems; neural networks that collect large amounts of data with which dynamic patterns of response and interaction with humans themselves can be reorganized; robotic systems that automate processes and accomplish tasks more effectively than human labor. In short, algorithms do not think, but they do, and the effect of their actions negatively affects the human being, for example, processing information in a calculated way in predictions, statistics and evaluations whose goal is precision, without being aware of who is affected , how and why; collect data from millions of people for corporate purposes without knowing if the privacy of individuals is violated; automate production chains affecting the activity, profession and employability of many human beings; such as imitating biased conceptions and interpretations of human beings themselves, generating artificial biases with which they become instruments that reinforce social discrimination [22]. Right now, the AI that we have available, already generates enough ethical concerns that must be addressed in a necessary way, both in the direction of the effects caused by the human-computer

Zapata Flórez A. Cognitive Priority over Ethical Priority in Artificial Intelligence: The Primordial Philosophical Analysis in Artificial Intelligence. Philos Int J 2022, 5(4): 000269.

Copyright© Zapata Flórez A.

interaction, as well as in the direction of the types and the very meanings of ethics[9] to be taken into account [23], which shows new ways of deepening the social reality that the same practical problem of AI manages to reveal.

AI constitutes a technology created by the human being whose sole purpose must be the common good, achieve the global well-being of humanity, social development and environmental sustainability [24]. The real ethics of AI must be aimed at finding the guarantees of a humanitarian and humanistic deployment of AI in its possibilities of interaction with human beings.

Conclusion

The cognitive priority over the ethical priority of AI can be translated into the following formula: the philosophical possibility determines the technological possibility, and the latter, in turn, determines the ethical possibility. With this I do not intend to undermine the ethical reflection that philosophy carries out in AI, much less insinuate that philosophy cannot think about AI ethically. In fact, I consider the ethical analysis of the impact that AI has and will have on society to be very necessary, and hence its enormous importance. Ethical reflection and regulation of technology, especially AI, is so fundamental that it is therefore essential to guide said reflection in the right direction with sensible and necessary concerns about the realistic impact that AI has today with its available techniques.

The cognitive foundation allows the ethical foundation: cognitive analysis exposes the limits and possibilities of AI, these possibilities present a tangible reality that we observe and affect us today. The cognitive priority limits the order of the risks of AI, so at the same time, it grounds ethical reflection and allows for realistic regulatory regulations aimed at the true impacts that we must face immediately. AI is not developed as naively believed to cause thinking consciousness in a computer by magic in the best style of science fiction cinema, rather, the purpose of AI must be understood according to the purpose of any other type of technology, namely, to enable the overall development of human well-being. Ethical regulation must be aimed at

seeking and guaranteeing said purpose, so it must be vigilant that the large corporations that produce AI and that other users do not go against social justice, freedom and dignity of all people in this contemporary world in the overflowing use of computing. That the possibilities of AI be an instrument of connection and global equity between humans and not an instrument that accentuates discrimination between them, since it is the main problem to be faced right now.

References

1. Puentes A (1999) El cerebro creador. Alianza Editorial. Madrid.

2. Martinez-Freire P (2004) Psicología y materialismo. Revista internacional de Filosofía 9: 123-142.

3. Gadner H (2000) La nueva ciencia de la mente. Paidos. Barcelona.

4. Martinez-Freire P (1995) La nueva filosofía de la mente. Gedisa. Barcelona. pp: 168.

5. Fodor J (2002) La mente no funciona así: alcance y límites de la psicología computacional. Siglo veintiuno. Madrid.

6. Bostrom N (2021) ¿Qué es el Transhumanismo?. Futuro hoy 2(2): 7-12.

7. Borges P (2021) Hasta la vista, baby: refletions on the risks of algocracy, killer robots, and artificial superintelligence. Revista de la Facultad de Derecho de México 70(279): 45-72.

8. Terrones A (2020) La actualidad del concepto de Sobrenaturaleza en José Ortega y Gasset: una mirada desde la inteligencia artificial. Análisis. Revista colombiana de Humanidades 52(96): 165-175.

9. Dieguéz A (2020) La función ideológica del transhumanismo y algunos de sus presupuestos. Isegoria. Revista de Filosofía moral y Politica 63: 367-386.

10. Gayozzo P (2021) Singularidad tecnologica y Transhumanismo. Teknokultura. Revista de Cultura Digital y Movimientos Sociales 18(2): 195-200.

11. Batin M Turchin, Markov A, Zhila S, Denkenberge AD (2017) Artificial Intelligence in Life Extension: from Deep Learning to Superintelligence. Informática 41(4): 401-417.

12. López M (2019) Del hombre máquina a la máquina hombre: materialismo, mecanicismo y transhumanismo. Naturaleza y Libertad. Revista de estudios

---

9   There may be a phenomenon of ethical colonialism in the type of ethical norms and regulations that are imposed from the centrality that is given a certain particular ethical vision, for example, the main ethical regulations about AI are formulated in Europe, without knowing if they take into account other different moral visions. The type of ethics that I criticized here could also be seen in that direction, that is, the ethics directed at artificial superintelligence. This ethic can be seen as a morality that exports the North American vision about AI given the trust that said society has in technological development, and that we see exported in its cultural products such as science fiction movies, in which machines they dominate man, so there is the correct ethics oriented to the superintelligence of machines.

Interdisciplinares 12: 179-190.

13. Agustí-Cullell J (2022) Beyond the AI Conundrum: the future of intelligence lies in its social flourishing. Philosophy International Journal 5(2): 000251.

14. Minsky M (1986) La sociedad de la mente. La inteligencia humana a la luz de la inteligencia artificial. Galápagos. Buenos Aires.

15. Minsky M (2006) The emotion machine: commonsense thinking, artificial intelligence and the future of human mind. Simon & Schuster Paperbacks. New York.

16. Dreyfus H (2002) Ser-en-el-mundo. Cuatro Vientos Editores. Santiago de Chile.

17. Clark A (1999) Estar ahí: cerebro, cuerpo y mundo en la nueva ciencia cognitiva. Paidos. Barcelona.

18. Muñóz M, Díaz S (2010) Inteligencia artificial ecológica para una nueva generación de robots. Factotum. Revista de Filosofía 7: 53-61.

19. Dreyfus H (1992) What computers still can't do: a critique of artificial reason. The MIT Press. Cambridge, Massachussette.

20. Dreyfus H (2002) Intelligence without representation: Merleau-Ponty's critique of mind representation. Phenomenology & Cognitive Science 1: 367-383.

21. Dreyfus H (2009) On the internet. Routledge. New York.

22. Nurock V (2020) Puede prestar cuidados la inteligencia artificial? Cuadernos de Relaciones Laborales. 38(2): 217-229.

23. Piedra J (2022) Decolonizando la Ética de la inteligencia artificial. Dilemata 38: 247-258.

24. Monasterio A (2022) La inteligencia artificial para el bien común (AI4SG): IA y los objetivos de desarrollo sostenible. Arbor, Ciencia, Pensamiento y Cultura 197(802): 1-19.