Beyond technology: Ethics and responsibility in the age of Artificial Intelligence

https://doi.org/10.56238/sevened2024.007-035

Helen Cristina Minardi Baumgratz[1] and Rodrigo Otávio dos Santos[2]

ABSTRACT

This study explores the relationship between artificial intelligence (AI) and human ethics, focusing on Ulrich Beck's "risk society" theory, highlighting the challenges and risks that AI presents in modern industrial society. The work emphasizes the need for effective governance and regulation of AI, addressing the ethical issues related to its implementation, and the importance of aligning the technology with human ethical principles. In addition, the possibility of different ideologies sharing similar core values is discussed, highlighting the integration of universal human values in the development of AI systems. The research underscores the need for accountability and transparency in automated systems decisions, and the importance of collaboration between human and artificial intelligence for safe and balanced decisions. This bibliographic and exploratory study seeks to offer new perspectives for future research at the intersection of AI and ethics.

Keywords: Technology, Artificial intelligence, Risk Society.

[1] Highest degree: PhD in Education and New Technologies
Academic institution: Colégio Militar de Brasília and UNINTER
[2] Highest degree: Post-doctorate in Technology and Society
Academic institution: UNINTER

INTRODUCTION

The evolution of civilization and scientific and technological developments have gone hand in hand throughout history and are deeply intertwined with social and environmental changes.

Technology, which permeates virtually all spheres of life, advances rapidly and creates new possibilities, but the lack of control over its growth raises debates about risks, ethics, and social justice.

The action of man and his scientific and technological achievements have allowed the solution of several vital problems for society. However, new problems have arisen and the prediction of consequences is necessary to the extent that man threatens his own existence as a result of this activity, as well as promotes the depletion of natural resources and environmental degradation.

Understanding this phenomenon and assessing its impact on society, culture, and people's lives contributes to the projection of future scenarios capable of anticipating possible risks, liabilities, opportunities, and alternatives. By critically reflecting on the processes of technological transformation, we seek to ensure that future innovations will bring benefits and not the ruin of society.

THE TECHNOLOGY

Throughout human history, technology has traveled a long path of progress that has radically changed society. From the evolutionary historical and cultural perspective, *Homo sapiens* is the only one among the species of the genus *Homo* that continues to exist, transforming the set of ecosystems and preserving the unity of the species.

For Harari (2015, p.361), *Homo sapiens* has established itself as the dominant species in the world and highlights the process of the three revolutions that have shaped human history: cognitive, agricultural, and scientific. He points out that, in the cognitive revolution, *Homo sapiens* was able to prevail over other animals for several reasons, among them the difference of the human brain and abilities such as: unique language, capacity for abstraction and memory, cooperation and the creation of subjective realities and shared myths. With the cognitive revolution, they transformed their social relations and developed the ability to cooperate; They acquired technical skills and created tools. In the agricultural revolution, they abandoned the life of hunter-gatherer and began to manipulate some plants and animals, using tools that allowed them to progressively increase the periods they spent in the same place, as well as the probability of survival and the development of the species. After the Agricultural Revolution, the emerging societies sought to improve their way of life, acquiring new knowledge at a frenetic speed, initiating the third revolution, the scientific Revolution (Harari, 2015, p 11)

Scientific revolutions are stages in the development of science in which there is a change in the research strategies established by their foundations.

The American philosopher T. Kuhn (2012, p.15-17) divided the evolution of scientific knowledge into three distinct stages. The first was called pre-paradigmatic, where the discovery of a theory to solve the problem had not yet been identified. The second stage, called normal science, comes along when a theory comes along that is capable of solving most of the problems, and researchers do not object to it because they believe it to be true; on the contrary, they seek to perfect it. However, with the passage of time, anomalies appear that are difficult to be remedied.

Para Kuhn (2011, p. 29):

> "[...] "Normal Science" means research firmly based on one or more past scientific achievements. These achievements are recognized for some time by some specific scientific community as providing the foundations for their further practice.

The third stage, entitled revolutionary science (Kuhn, 2012, p. 21), is constituted when numerous anomalies are present, leading scientists to deny a theory, establishing a clash between distinct groups of scientists who start to advocate in favor of different theories.

Kuhn (2012, p.17) also observed that in scientific practice, people invested so that nature would adapt to pre-established paradigms. According to his philosophy, normal science is anchored in the metaphor of "*puzzles*", where the result to be achieved is already known, and it is enough to unveil the means to achieve it, just like a Chinese cube. Kuhn contradicted the scientists by declaring that they did not have the truth as their purpose, but rather to decipher the riddles and solve them. He asserted that during this period, dogmas are fundamental, since they determine the "*puzzles*" and institute the criteria for their solution. Conflicting paradigms are discouraged and the scientific community converges in the search for the solution to the puzzle. Solving puzzle secrets brings fame and recognition, but for the scientist's work to be legitimized, he needs to be part of a scientific community.

In this process, the adoption of a paradigm to support research is fundamental and assures scientists of control of research, as well as stimulates its advancement. In this sense, it is necessary to emphasize that the production of scientific knowledge is carried out by scientists who make experiments, inferences and observations to prove their findings and, in order for the result of a research to be considered valid by the scientific community, it is imperative that some criteria and standards are met.

According to Kuhn (2011a, p. 220),

> [...] The term "paradigm" is used in two different senses. On the one hand, it indicates the whole constellation of beliefs, values, techniques, etc., shared by the members of a given community. On the other hand, it denotes a type of element of this constellation: the

> concrete solutions of puzzles that, used as models or examples, can replace explicit rules as the basis for the solution of the other puzzles of normal science.

In the course of the period of revolutions, scientists make new discoveries and achieve new results, bringing significant contributions such as the emergence of new techniques and technologies. The origins of the terms "technique" and "technology" can be traced back to the ancient Greek notion of "*techne*" (art or craftsmanship), meaning the set of knowledge pertinent to a particular productive practice.

Heidegger (2007, p. 376-377) states that in the current conception technique is a human activity and that its essence is a means to an end. Thus, it is perceived as the fulfillment of the human will, free to act and to benefit from its achievements, while contradicting historical facts. From these answers, the instrumental technique, that is, designated to "doing things", and the anthropological technique, that is, a human activity. He considers this definition to be correct, but invites a more complete understanding of the term, considering what lies behind the meaning of instrumentality in the light of causality.

Using the example of a silver chalice, he observes the existence of four causes: the first called causa *materialis*, the matter which constitutes the chalice, the causa *formalis*, the form which it acquires, the causa *finalis*, the end and the causa *efficiens*, the finished royal cup. However, he concludes that there is no thought that clarifies the connection between these four causes (Heidegger, 2007, p. 377).

He compares the traditional model of causality, starting from the Greek conception, revealing that Aristotle did not perceive the goldsmith as the agent who "effectuates" the production of the chalice. For him, the goldsmith would be the starting point for the chalice to exist. In this sense, the goldsmith is responsible for the chalice; The chalice would be "indebted" to the goldsmith, in a relationship of responsibility and indebtedness.

For the Greeks, *poeisis* means "to generate" and is closely related to "to be responsible." The author distinguishes between the two forms of generating, the first being associated with *poeisis*, referring to what the craftsman and the poet practice, that is, the poet makes the poem, the craftsman makes the sculpture. The second is *physis*, the production that takes place in nature. However, these two forms are considered *poeisis* in the sense that something that was not present becomes present.

This concept of "*poeisis*" as a quality of revelation led him to yet another Greek word: "*aletheia*," which means "to reveal," as well as being used as "truth." In this sense, Heidegger argues, that technique is a kind of *poeisis*, i.e., a way of "revealing" as well as "the realm of truth." In this way, it can be seen that Heidegger leads to an alternative way of reflecting on technique, detached from instrumentality, and its similarity to poetry in the way of opposing the world.

The word technique originating from the Greek *technikon* is correlated to the word *techne*. He points out that "*techne*" can be considered a form of knowledge since it refers to both manufacturing and the arts and that when used together with *episteme*, from which the word epistemology derives, it is understood as the philosophy of knowledge.

Techne and *episteme* are linked. The former is related to that which arises only from man's intervention in nature, and the latter is related to that which arises only from his own nature. *Techne* reveals everything that doesn't present itself. Thus, what is decisive in *techne* is not in doing, manipulating or in the use of means, but rather in the aforementioned revelation.

> Both *episteme*, *poiesis*, tekne – the three ways in which the Greeks conceived knowledge – are configured as ways of manifesting something, apprehensible by thought, but never created by it. Tekne, which is inappropriately taken as equivalent to the technical, therefore has an original meaning quite different from the modern one: it is also a form of unveiling, it uncovers what is not produced by itself, it is an auxiliary way for something to become. (...) "the decisive factor of *tekne* does not reside, therefore, in the making and handling, nor in the application of means, but in the aforementioned uncovering" (HEIDEGGER, 2002, p. 18, apud WEBER, 2011, p. 5).

Heidegger then wonders if it is appropriate to apply the concepts of Greek philosophy to modern technology, which has followed a path closer to the physical sciences than to the arts and crafts itself, concluding that science cannot be seen as a "cause" or "origin" of technique, but rather in its orientation to the world.

Heidegger (2007, p. 381) distinguishes the ancient forms of technique, which produces energy from the force of the winds, without being able to store it; capable of depleting the planet's resources. He also introduces the concept of "permanent reserve", strongly related to the concept of "instrumentality", stating that the instrumental orientation of technique transforms the world into a permanent reserve. He also points out that the use of the expression "human resources" equates human beings with raw materials such as coal or oil, but points out that just as man has no control over the formation of these resources, he can only control or guide thoughts and actions.

In this sense, Heidegger clarifies that in the bond between man and the world, an attitude called framing arises, using the German word *Gestell*. This framing alludes to the human drive to control and shape thought. Thus, understanding modern technique as the use of framing also encompasses the problem of understanding the essence of technique and its historical relationship with man.

However, it suggests the existence of a "danger" or threat associated with man's alienation and inability to depart from modern technology, just as it suggests a growing saving power where danger dwells. He also points out a path of salvation in a poetic look from the perspective of the world, seeking a free and inquiring relationship about the role of technology in the life of each one.

(Feenberg, 2003), presents the conception of technology from the perspective of axes that reflect its relations with human values and capacities, in a critical approach to essentialist philosophy. It rests on the distinction of the Greek words *physis*, translated as nature, denoting the ability to create oneself; and *poiêsis*, which is the practical activity of doing something. *Poiêsis* means to reveal or bring into existence something that did not exist before and its product is called an artifact.

He expounds on the concept of *techne*, which means the knowledge or discipline associated with some form of *poiêsis* (Feenberg, 2003). It also states that in the Greek view of things, each *techne* includes a purpose and a meaning for the artifacts whose production it guides. To the Greeks, *thecnai* show the "right way" of doing things in a very strong, even objective, sense. Although artifacts depend on human activity, the knowledge contained in *thecnai* is not subject to subjective opinion or intent. Even the purposes of the things produced share this objectivity, insofar as they are defined by *thecnai* (Feenberg, 2003, p.2).

In the second fundamental distinction, existence clarifies whether something is or is not. Essence, on the other hand, must answer what the thing is. For Plato, nature is divided into existence and essence, just like artifacts, but there is no separation between them. The idea of the artifact for the Greeks does not emerge in a subjective way, rather, it is necessary to possess dominion over a *techne*.

In Greece, the processes of acquiring knowledge (*episteme*) and producing artifacts (technê) were distinct. In this way, for the Greeks the concepts of t*echnê* and*and pisteme* differed. *Episteme* expressed knowledge in its pure state, but today it is understood as science. In this sense, the construction of artifacts is achieved through the combination of knowledge, practice and experimentation. In modern times, despite sharing these fundamental concepts, common sense points to a dichotomy between the existence and essence of objects. Technology does not respond to the essential purposes that are contained in nature, but rather to the subjective intention of prioritizing the will, the will. In this way, a collapse of civilization was installed from then on due to the dispute for technological power, as well as for the strategic dominance portrayed in the two great wars.

Feenberg argues that modern societies, especially from the eighteenth century onwards, emerge from the questioning of traditional forms of thought and force institutions to justify themselves as useful to humanity. And, under the impact of this demand, "science and technology became the basis for the new beliefs", decisively influencing culture and positioning itself as the 'rational' choice. From this perspective, "technology becomes ubiquitous in everyday life and technical modes of thought come to predominate above all others" (Feenberg, 2003, p. 1).

For Feenberg, contemporary technology emerges in four strands: instrumentalism, determinism, substantivism, and critical theory. (Feenberg, 2003, p. 6) Instrumentalism is the

conjunction between human control and the neutrality of technology, so that it presents itself only as an instrument to satisfy human needs.

Determinism holds that technology is not under human control, but drives development through the progress of knowledge and mastery of the forces of the natural world.

In substantivism, it is admitted that technology is autonomous and loaded with values. According to the author, if technology embodies a substantive value, it is not merely instrumental and cannot be used for the different purposes of individuals or societies with different ideas of good. (Feenberg, 2003, p.7).

It proposes a new critical theory in relation to technology, admitting the possible threats expressed in the substantivism, but provides an opportunity for a philosophical reflection on the control of technological development, through institutions more appropriate for this process. In this approach, it is possible to verify the sharing of common characteristics of both instrumentalism and substantivism.

In this reflection, Feenberg departs from the position of the philosopher Martin Heidegger, who believes that technology is dominant, uncontrollable and does not offer a way out, so that only a god could save humanity. For the author, there are pitfalls that often become prejudices about the way technology is used and warns of the adoption of a critical stance in the face of this reality, opening space for questioning its importance and usefulness.

This conception of man's relationship with technology finds its source in the ability to conceive them as revealing exercises under a useful eye and critical thinking, which are often neglected in favor of a linear economy that meets the objectives of the modern society dominated and controlled by technological innovation, its power to influence people and completely modify the way of life of a population.

From this perspective, the preponderance of technology and its implications in modern life cannot be portrayed as a dilemma situated between utopian aspirations and dystopian fears, but rather reflected in the light of the privilege of action and responsibilities inherent to every human being who are capable of transforming a society and indicating possible paths and solutions capable of satisfying democratic demands embraced by technological development in the search for a possible reconciliation between human beings and technologies.

For Vieira Pinto (2005, p.2019-220), the term technology can be conceptualized from four different meanings. The first one presented is the epistemological or as logos of technique, where technology is seen as the theory of technique, being related to the skills of knowing how to do and produce. As an appropriation of the technique of the means of production, technology is confused with technique and the worker who plays this role is called technician, who in general does not have a critical training to think critically in relation to the product he is delivering. He clarifies that

technique, as a productive activity of man, represents an object of study of technology, "there is undoubtedly a science of technique, as a concrete fact and therefore an object of epistemological indication" (VIEIRA PINTO, 2005, p. 220).

In the second sense, technology is related to technique, but excluding the social and cultural aspects involved. In the third, the author presents technology as "a set of all the techniques available to a given society, at any historical stage of its development". And in his latest conception of the term, he perceives technology under the aspect of the ideologization of technique, where political and ideological issues influence certain groups or populations. From these definitions, it can be seen that from his point of view, technology could be understood as a science of technique, but duly elevated to the condition of political and social ideology. In this sense, technique is described as the result of man's perception, which takes shape through instruments and machines capable of influencing the thinking and culture of a given group.

Vieira Pinto (2005, p. 233) observes that current technological development is the result of the historical accumulation of knowledge and social practices, in addition to highlighting that when a technology becomes current, the next technological explosion is launched, in view of the limitations of the potential improvement of the one in use. From this angle, it can be seen that through globalization, the idea that the whole world enjoys these technologies motivates the incessant search for increasingly advanced artifacts, without any reflection on the real need for their use. Thus, conditioning is introduced capable of provoking changes in people's way of life that will be reflected in social, cultural and productive relations, always being linked to the political and market aspect that flows into the question of ideology.

With the increasing development and convergence of technology and science and their expressive dominant action in all areas of human life, the benefits and adversities have become the object of philosophical reflection due to the risks and harms caused to man and nature. Man, by means of new techniques, is confronted more and more with the danger that he himself creates. Thus, new questions arise about the use of technology and the transformation of human action, demanding a moral and ethical evaluation from the perspective of the concept of responsibility.

THE MAN

Hans Jonas, noted in his work *The Principle of Responsibility*, that the nature of human action has changed and technology, which is an important tool in solving the problems of modern life, has become a threat.

Man was not content with dominating nature, he went further, he began to dominate himself, influencing its fundamental characteristics and proving that he is not an immutable being. In light of the rise in technological power, *homo faber* has become the object of technology. Jonas says that

homo *faber* applies his art to himself and prepares to invent a new fabrication of the inventor and creator of everything else (OLIVEIRA, 2013, p.14). For him, man is a being condemned to death, however, with the development of new technologies for biomedicine, man has been seeking to age or live eternally through the prolongation of life, control of behavior and genetic manipulation.

However, the prolongation of life has ethical consequences, such as the increase and predominance of the elderly population and, consequently, the absence of innovation with a view to the end of procreation. These issues have taken on a significant role since they deal with the rejection of death and bring to the fore issues such as euthanasia, cloning, *in vitro generation,* among many others that are important from an ethical point of view. Faced with these questions, traditional ethics became ineffective and Hans Jonas resorted to the ethics of responsibility that finds its sources in metaphysics. He placed responsibility at the center of ethical theory.

From Jonas' perspective, in the face of the threats of danger associated with current techno-scientific progress, as well as in the future, the effects of human actions conducted by technology must preserve humanity and nature. In this sense, man's freedom, evident in the power of technology, can also be conceived for the destruction of man himself and nature, leading to the subjugation of his own image. This power provided by technology obliges man to assume his responsibility and to control the enterprise of technological advancement, since the exercise of power unaccompanied by duty results in morally reprehensible irresponsibility. It is noteworthy, then, that for Jonas, the ethics of responsibility will reorient man as an ethical subject to overcome nihilism, which points to a scenario of uncertainties in relation to the future.

Faced with this panorama, Jonas points out the need for a reflection on what is humanly desirable and what should determine the choice. With the advances of nano and biotechnology, problems related to life, death, today and tomorrow arise, which, in the light of Jonas' concept of responsibility, reveal opportunities for man to take control of the development of technology for the benefit of an ethical action compatible with the preservation and integrity of human life and nature. Since the inclinations of *Homo Faber* are to act without considering the consequences of their attitudes.

Globalization and technological advancement have raised the need for responsible behavior in the face of the anthropogenic impact and the possible consequences of decisions made, based on the principle of the continuation of human existence. Freedom to act and responsibility are intrinsically embedded in each other. Man becomes responsible as he engages in any activity and knows its consequences, raising the issue of anticipation and prediction of risks pointed out by Beck. In this sense, the ethics of individual responsibility spreads to social responsibility, making it essential to assess the side effects and risks, even with a very small probability, of any action that threatens the existence of humanity. This action must be inserted in an expanded spatio-temporal

dimension, so that its possible future consequences can be clearly prognosticated. The implications of the unchecked power of technological advances transcend the boundaries of time and space, and crises, once local, have become global to the extent that everyone is affected.

The technology that integrates modern daily life and the search for its constant updating, reveals critical aspects about risk possibilities where science and the state exert power of persuasion for practice, which must be in accordance with the orientation of moral principles. However, divergence in ethical and moral values from one society to another hinders social order and endangers world peace.

Technology calls all societies to reflect on the dimensions of power and its political, social and ethical consequences, as well as its impact on human relations and governance structures. In this sense, the control of territories also unfolds into the control of citizens and society. Through new technologies, surveillance and control systems bring to light information about the activities of the individual in the private and public spheres, sometimes leading to violations of fundamental rights and perverse consequences for the security of the citizen's life.

The technological advancement that empowers man has ambivalent effects that benefit and destroy, as well as generates the need for responsible conduct, complemented by freedom and based on ethical principles that should provide new normative principles for human action, based on the image of the object that can only be determined from a philosophy that encompasses life and human dignity.

For Beck (1999a, p. 178), the current threats to humanity come from man's own action, which generates unpredictable implications, opposing the conception that the advancement of science and technology would guarantee protection against the scourges of nature. Man, who at the beginning of the modern era sought to dominate nature through technological advances and the progress of science, begins to provoke a reaction on the planet to his disorderly intervention, dissolving the common sense of progress.

It states that:

> Risk society means: the past has lost its power of determination over the present. In its place comes the future – that is, something that does not exist, something fictitious and constructed – as the cause of life and action in the present. When we talk about risks, we are discussing something that does not occur but that can arise if the direction of the boat is not immediately changed. Imaginary scratches are the whip that makes the present time move. (BECK, 1999a, p. 178):

Beck argues that a rupture in modernity emerges and a new society is revealed, dissolving the structures of industrial society and replacing them with the risk society and it is no longer possible to consider traditional certainties as guarantees and that humanity is subject to risks generated by man himself and quite different from the past.

For him, one of the main characteristics of the risk society is that modern society becomes reflexive by manifesting itself as an issue and a problem for itself. Quoting François Ewald, he states that risk is a form of control and consists of a real kind of virtual reality, which increases the obligation and power to modify events as a threat increases. The author points out the existence of a paradox in the probability that surprises will erupt, as attempts to colonize the future are undertaken; justifying the two stages through which the notion of risk passes.

In the first stage, risk is presented as an elementary calculation, which attempts or converts the unforeseeable into the predictable. In the second stage of risk, it presents manufactured uncertainty, where the generation of risks stems from scientific and political efforts to control or reduce them. From this perspective, risk becomes random and inevitable, but manufactured uncertainty would be the origin of new risks.

Another aspect that Beck develops is to understand society as a laboratory, but without those responsible for decisions and results. In view of this, the question arises about the role of politics in the face of technological development. In this case, Beck identifies a mismatch between the industry and the politicians, since the former has autonomy in decisions and a monopoly on the application of technology while the politicians remain oblivious to this development, but favorable to the economic future and employment in the country. Thus, the industry confirms itself as a decision-maker without assuming risks and responsibilities, while politics loses its contour, leaving it only with the role of endorsing the decisions not conceived by it. In view of this, industrial society constitutes rules and institutions in response to the unexpected implications of the risks it causes in the risk society, previously calculated in industrial society.

It is in this space that he argues for the institution of a broad debate on decisions, scientific agendas, implementation of new technologies, as well as the construction of a legal and institutional framework in order to achieve democratic legitimacy.

Finally, Beck proposes the creation of new parameters to measure the social responsibilities of risk production, in addition to demanding public justification from agents of industrial society. However, this issue becomes complex to the extent that multiple subjects are involved and the chance is often not linear.

Risks have become global, cross generations, cross borders, and affect everyone, regardless of social class. In this way, some risks, such as those of social, political, scientific or economic origin, interfere with the course of societies, sometimes generating unintended consequences.

However, although global risks are shared, knowledge capital also involves learning about multiple risk situations in a dynamic of cumulative knowledge construction that can be strongly linked to the development of strategies to combat risks.

In this scenario, the risks may become imperceptible to the layperson and only specialized scientific knowledge could be able to perceive them, thus becoming a threat to the layman's autonomy to assess these risks. In this sense, the role of science in public life and politics changes significantly, influencing society through its voices and giving rise to new political forces.

In this escalation of risk production, the very scientists who should determine the level of risk of new technologies and technical systems, assume a privileged position as they are called upon to demystify knowledge into the language of political dialogue and decisions before public opinion.

However, experts may be unaware, omitted, or misrepresented information if they are politically compromised. From this perspective, they can deprive the population of the real purposes, means, threats and implications of a given risk, making it impossible to find an adequate solution, opening the way for the construction of authoritarian or even totalitarian political regimes that can project a scenario of catastrophic future situations and impose security laws, as well as instruments of surveillance and control.

On the other hand, the progress of science, new technologies and industry combined with political and economic interests, in a subtle and imperceptible way, transforms society and the planet into an immense laboratory where experiments are carried out with unknown, unpredictable, incalculable and unaccountable effects.

According to Krohling (2011, p.127):

> The priority is the cultivation of the moral conscience of the human being, who must place the ethics of responsibility in all his acts. As a rational and social being, he is an alteral being and depends on others. This calls for a conduct of precaution, prudence, and factual responsibility, since man is the creator and main actor of technological innovations.

This emphasis on the ethics of responsibility suggests the need for a continuous critical evaluation of technological development. It is not enough to move forward in the creation of new technologies; It is crucial to consider the ethical consequences of these advances. Such an approach requires deep reflection on the role of technology in society and the establishment of governance and accountability mechanisms that ensure that technological advances are utilized in ways that benefit humanity and preserve balance with fundamental human values and collective well-being.

THE RISK SOCIETY

In the theory of the "risk society" developed by Ulrich Beck, the notion that the advent of new technologies and tools for the development of industrial society entails the emergence of more serious and complex risks is emphasized. This theoretical approach is based on the importance of understanding and characterizing the phenomenon of risk.

He introduced the concept of the risk society in the context of the significant transformations brought about by advanced modernization. At the heart of this theory is the idea that we live in an era defined not only by the opportunities and benefits brought about by technological progress, but also by the risks and dangers that those same technologies generate. This concept becomes particularly relevant when considering the disruptive impact and uncertainties associated with artificial intelligence (AI).

The rise of AI represents a technological revolution, redefining the contours of society, economics, and human interactions. AI, with its ability to learn, adapt, and perform certain tasks efficiently, brings with it a complex set of challenges and opportunities. In the light of Beck's theory, the debate about artificial intelligence reveals itself as an intrinsic need to carefully assess the associated risks. This underscores the importance of considering several interconnected aspects when discussing the reliability and security of the future of artificial intelligence in society.

The balance between Artificial Intelligence, and the human ethical essence, which focuses on distinguishing between right and wrong, poses a significant challenge in the modern age. This challenge becomes even more complex when it comes to making difficult moral decisions that transcend ideologies and are rooted in fundamental ethical principles.

This intersection between advanced technology and moral ethics raises crucial questions: How can AI be aligned with core human values? How can automated decisions reflect deep ethical considerations? The answer to these questions is not trivial and requires continuous reflection, highlighting the importance of a multidisciplinary approach that integrates technology, philosophy, ethics, and politics.

In this regard, it is critical to recognize the inherent complexity of AI and its ability to significantly influence human lives, as it operates in a domain that has traditionally been exclusive to human decision-making. This raises questions about responsibility and *accountability* in decisions made by automated systems. To address these issues, it is necessary to establish a clear ethical framework that can guide their development and implementation. This framework should be built on the basis of a broad consensus that takes into account universal human values, rather than being shaped by specific ideologies.

The integration of ethical principles into the design and implementation of AI systems is necessary to ensure that the technology operates within a morally acceptable context. This entails adopting ethical guidelines in the development of algorithms, ensuring that they are transparent, fair, and impartial. Transparency is particularly important as it allows users and regulators to understand how decisions are made, which is essential for establishing trust and holding developers and users of the technology accountable.

Additionally, AI development should be accompanied by an ongoing, interdisciplinary dialogue between technology developers, ethicists, policymakers, and the general public in order to ensure that ethical considerations are integrated into AI in a way that reflects a broad spectrum of human perspectives and values. Such an approach also helps mitigate the risk that AI will be used in ways that perpetuate existing biases or introduce new forms of discrimination.

In this context, governance and regulation emerge as fundamental aspects. The need for these mechanisms stems from the potential for abuse associated with AI and the significant impact this technology can have on individuals and society. Effective regulation involves the creation and implementation of policies, laws, guidelines, and standards that aim to guide and control how AI is developed, distributed, and used, against abusive uses of the technology, which can include privacy violations, algorithmic discrimination, and automated decision-making with negative consequences for people.

In addition, the regulation ensures accountability and transparency in the use of AI, so that there is the possibility of taking corrective action in case of failures or abuses, as well as to establish clarity on legal and ethical responsibility. By establishing clear guidelines, AI developers can understand the expectations and limits imposed by society, thereby encouraging the development of technologies that respect core human values and promote social well-being.

FINAL THOUGHTS

In the modern era, the rise of artificial intelligence (AI) brings with it not only technological advancements but also a host of complex risks and ethical challenges. The development and application of AI has the potential to profoundly impact diverse aspects of society and human life. Given this scenario, the assessment of the risks associated with AI becomes a relevant aspect, where ethics emerges as a fundamental component in the governance and implementation of these technologies.

The interplay between AI and human ethics requires deep reflection on the importance of universal human values, even amid a spectrum of divergent ideologies. An emphasis on common human values can help avoid the pitfalls of indeterminate AI or systems in which values are unilaterally imposed by developers or the technology itself. By recognizing that different ideologies may share similar values, it is possible to develop AI systems that are more inclusive and representative of diverse human perspectives.

The issue of responsibility is central to the AI debate, since when a machine makes a wrong ethical decision, the responsibility for that decision remains ambiguous. The lack of clarity about who is responsible for these decisions – whether it's the developer, the user, or the system itself – raises complex questions about *accountability* and fairness. Therefore, it is imperative that there are

clear accountability and governance mechanisms in place to ensure that decisions made by AI systems are fair and transparent.

Additionally, over-reliance on AI can lead to the loss of critical decision-making skills and moral judgment among humans. However, the harmonious and collaborative coexistence between human and artificial intelligence is a path with the potential to provide balance between the efficiency and analytical capacity of machines and human ethical and emotional discernment, to ensure that the decisions made are not only technically competent, but also ethically sound and socially responsible, where technology amplifies and enriches the human experience, rather than subvert it.

This will not only ensure the safety and reliability of the decisions made by AI, but it will also ensure that those decisions are aligned with the core ethical values of human dignity and rights.

REFERENCES

1. Beck, U., & Franklin, J. (Eds.). (1998). *The Politics of Risk Society.* Cambridge: Polity.

2. Beck, U. (1999). *O que é globalização? Equívocos do globalismo: respostas à globalização.* São Paulo: Paz e Terra.

3. Feenberg, A. (2003). *O que é Filosofia da Tecnologia?* (17 p.)

4. Hans, J. (2013). O homem como objeto da técnica segundo Hans Jonas: o desafio da biotécnica. *Problemata: R. Intern. Fil., 4*(2), 1338. DOI: http://dx.doi.org/10.7443/problemata.v4i2.16966

5. Harari, Y. N. (2015). *Sapiens: Uma breve História da Humanidade.* São Paulo: L&PM Editores.

6. Heidegger, M. (2007). A questão da técnica. *scientiæ zudia, 5*(3), 375-398.

7. Krohling, A. (2011). *A ética da alteridade e da responsabilidade.* Curitiba: Juruá.

8. Kuhn, T. (2012). *A função do dogma na investigação científica; organizador: Eduardo Salles O. Barra; tradução: Jorge Dias de Deus.* Curitiba: UFPR.

9. Oliveira, J. (2013). O homem como objeto da técnica segundo Hans Jonas: o desafio da biotécnica. *Problemata, 4*, 13-38.

10. Sirvinskas, L. P. (2006). *Manual de direito ambiental.* (4ª ed.). São Paulo: Saraiva.

11. Weber, J. F. (2011). Técnica, tecnologia e educação em Heidegger e Simondon: Destruição do pensamento ou ampliação da experiência? *X congresso Nacional de Educação.* Paraná: PUC.