ISSN 2227-7242 (Print), ISSN 2304-9685 (Online)
Антропологічні виміри філософських досліджень, 2026, Вип. 29
Anthropological Measurements of Philosophical Research, 2026, NO. 29

UDC 004.8:17.02

A. I. KUZMENKO[1*], N. V. KHALIPOVA[2*], I. Y. LESNIKOVA[3*]

[1*]University of Customs and Finance (Dnipro, Ukraine), e-mail albinakuzmenko03@gmail.com, ORCID 0000-0001-7278-3647
[2*]University of Customs and Finance (Dnipro, Ukraine), e-mail khalipova@ukr.net, ORCID 0000-0001-5605-6781
[3*]University of Customs and Finance (Dnipro, Ukraine), e-mail lesnikova@i.ua, ORCID 0000-0002-2750-6031

AI and Human Responsibility

Purpose. The purpose is to analyze the contemporary request for studying forms of responsible attitudes toward the development of artificial intelligence (AI). Achieving this involves addressing the following tasks: a) clarifying the core components of the modern interpretation of AI; b) studying the nature of the threat posed by AI to the further development of humanity. **Theoretical basis.** The principles of humanism, worldview pluralism, and objectivity form the theoretical and methodological basis of the article. **Originality.** The nature of the AI threat to humans has been examined through the lens of naive acceptance of instrumental rationality as universal, accompanied by a disregard for the spiritual rationality inherent in human beings. The peculiarity of the modern interpretation of AI stems from an abstract vision of reality that focuses primarily on its quantitative parameters. This is a departure from humanistic values and meanings, which are prerequisites for reducing humans to a means of servicing AI. The existence of a heuristic potential for understanding the problem of AI responsibility is substantiated in those texts of 20th-century humanists that are devoted to the phenomenon of technology. **Conclusions.** It has been revealed that the basis of the irresponsible interpretation of AI is the naive assumption about the universality of instrumental rationality and nihilism regarding spiritual rationality. It has been found that an abstract vision of reality is one of the substantive prerequisites for the prevalence of naively optimistic interpretations of AI, which consists in focusing attention exclusively on the quantitative parameters of the world, which is complemented by the promotion of man to the forefront as the center of the universe. Understanding the essential difference between the reduced type of rationality underlying AI and the spiritual rationality inherent in humans allows us to concretize the red lines that define human responsibility toward the spontaneous tendencies of AI development in modern civilization. The authors see prospects for further research in understanding the specifics of the existence of human freedom and responsibility at this stage of history.

Keywords: human; artificial intelligence (AI); instrumental rationality; spiritual rationality; ethics; sacred

Introduction

Our era is marked by the rapid development and widespread use of artificial intelligence (AI) technologies. It is radically renewing most spheres of social life, including the economy in a broad sense and the spheres of spiritual life and education. It is now clear that the widespread integration of artificial intelligence systems largely drives the dominant trends in technology and culture. When analyzing AI as a technical tool in the modern world, it is worth noting the wide range of emotional attitudes toward it, from admiration and awe to skepticism and depression. An additional factor in rising emotional tensions is the growing scale of human influence, which increases the likelihood of omnicide. Analyzing the consequences of AI's expansion into all spheres of human life, there is the displacement of humans and their values to the margins of culture and history. The aggressive interference of AI into domains tied to the specifics of human nature is a cause for grave and well-founded concern. This points to a threat of significantly distorting the quintessential, centuries-old understanding of the human being as a rational and spiritual entity. The particular danger of this trend lies primarily in the subtle encroachments aimed at dismantling those forms of human rationality that represent the realm of higher spiritual meanings. Therefore, the task of understanding and protecting the right to autonomous existence of the sphere of higher (spiritual) meanings in the conditions of an undeclared war by AI is now becoming more urgent.

Creative Commons Attribution 4.0 International
doi: https://doi.org/10.15802/ampr.v0i29.365948
© A. I. Kuzmenko, N. V. Khalipova, I. Y. Lesnikova, 2026

When analyzing the role and potential of AI today, it is impossible to overlook the profound research into information processes – a key to understanding the present – conducted by the renowned Israeli philosopher and historian, Yuval Noah Harari. Recently, a Ukrainian translation of his book "Nexus: A Brief History of Information Networks from the Stone Age to AI" was published by Bookchef. Throughout these pages, the author explores the phenomenon of AI, making a compelling case for its nuanced impact on humanity's future while examining the potential threats AI poses to mankind. Schematically reviewing the author's main theses, we consider it appropriate to emphasize the validity of the author's thesis regarding the paradoxical result of centuries of human effort to significantly alter its ontological status. And although, over the past 100,000 years, the author writes, we, as intelligent beings, have accumulated enormous power, despite these undeniable achievements of ours, humanity has found itself in an existential crisis. And therefore, skeptical voices are increasingly heard about those dark prospects that await us in the era of the new information network (the era of AI), which contain the threat of suicide.

Analyzing the factors of this critical situation, Harari rightly notes that today, many tend to equate information, as knowledge, with wisdom. Emphasizing the urgency of the problem, the philosopher poses a rhetorical question: "Why", Harari (2025) writes in the Prologue, "are we so good at gaining more information and power, but so poor at gaining wisdom?" (p. 10).

The research literature convincingly demonstrates that concerns about the current state of affairs are increasingly gaining attention. This concern is about the importance of the task of critically rethinking the basic ethical principles that underlie the spontaneous development of AI. Here, it is difficult to disagree with authoritative American researchers who emphasize the importance of targeted efforts both within individual states and on an international scale:

> Moreover, the development and deployment of conscious AI demand robust ethical governance to tackle the complex moral issues that arise and to prevent potential harm. The establishment of independent oversight bodies and ethical review boards is crucial for monitoring AI research and development, ensuring that AI advancements comply with ethical standards. Given the global nature of AI development, international cooperation is essential to address the ethical and regulatory challenges that transcend borders. Establishing international frameworks and agreements is necessary to harmonize regulations and promote global ethical standards, ensuring that development benefits humanity universally and responsibly. (Barnes & Hutson, 2024, p. 120)

Creative Commons Attribution 4.0 International
doi: https://doi.org/10.15802/ampr.v0i29.365948
© A. I. Kuzmenko, N. V. Khalipova, I. Y. Lesnikova, 2026

ISSN 2227-7242 (Print), ISSN 2304-9685 (Online)
Антропологічні виміри філософських досліджень, 2026, Вип. 29
Anthropological Measurements of Philosophical Research, 2026, NO. 29

Kakembo Aisha Annet holds a similar position regarding the expediency of critically reimagining the ethical concepts that have determined the human way of life for centuries. Today, the researcher emphasizes, we are in a radically different situation, and therefore, the relevance of researching the artificial intelligence (AI) system is axiomatic in conditions when it is increasingly aggressively interfering in human life. It is difficult to disagree with the thesis that today we need to look with fresh eyes at the relationship between AI and philosophy, which implies a critical attitude toward superficial ethical theories. The thesis about the importance of an impartial search for those values that will contribute to a person's responsible attitude in conditions of uncontrolled spread of AI is relevant (Annet, 2025). The above-mentioned factors determine our vision of the goal.

Purpose

Analysis of the contemporary demand for studying forms of responsible attitudes toward AI development. Achieving this involves: a) clarifying the main components of the modern interpretation of AI, and b) studying the nature of the threat AI poses to the further development of humanity.

Statement of basic materials

As the analysis of the research literature shows, the term "artificial intelligence" primarily refers to the development of computer systems capable of performing tasks that, as a rule, have so far been solved with the help of human intelligence. These include visual perception, decision-making, language translation, and the optimal use of autonomous vehicles, among others. In other words, artificial intelligence is understood as the ability of automatic systems to perform certain functions of human intelligence, which are associated with the search for optimal solutions to tasks that are based on previous experience and involve considering the nature and likely consequences of external influences. We do not consider it appropriate to dwell in detail on the main approaches to defining AI and various options for its application. Domestic research literature already has a thorough and sufficiently complete analysis of existing approaches (Kozlovets, 2024).

The prerequisite for a holistic understanding of the AI phenomenon is attention to a) its history and b) the opposing emotional options for its assessment, namely, the optimistic and the pessimistic. For many centuries, mankind has dreamed of conquering the external world through its own will. One of the important preparatory steps on this path is mathematics. This feature was implicit for a long time, since the individual was perceived as that part of the Universe that must live according to its laws. The situation changed radically in the Modern Era, when humans became the center of the Universe, and their consciousness became the starting point for reasoning about the external world and the measure of the significance of individual things. As is known, the heliocentric hypothesis of Nicolaus Copernicus, as well as the philosophical teachings of Francis Bacon and René Descartes, became milestones on the path to the formation of anthropocentrism in the worldview of Europeans. The latter, as is well known, emphasized the ability of men to master the world and subordinate it to the human will. For us, the important fact here is that, for Descartes, as well as for his older contemporary Galileo Galilei, mathematics was the language of nature. Somewhat later, Gottfried Leibniz emphasized the idea of the key importance of mathematics in the process of comprehending the way nature exists. For him, it was obvious that the human mind could be reduced to mechanical calculation. The axiomatic

Creative Commons Attribution 4.0 International
doi: https://doi.org/10.15802/ampr.v0i29.365948
© A. I. Kuzmenko, N. V. Khalipova, I. Y. Lesnikova, 2026

ISSN 2227-7242 (Print), ISSN 2304-9685 (Online)
Антропологічні виміри філософських досліджень, 2026, Вип. 29
Anthropological Measurements of Philosophical Research, 2026, NO. 29

nature of the ideas about the possibility of reducing human rational thinking to a system of mathematical knowledge by analogy with algebra and geometry is a common point of the positions of such bright thinkers of the 17th century as Leibniz, Hobbes, and Descartes. Thomas Hobbes emphasized in the pages of his famous work "Leviathan" that "reason is nothing but reckoning". Retrospectively evaluating the ideas of the above-mentioned philosophers regarding the importance of a system of physical symbols for the knowledge of the world, it is difficult not to see in them a signpost in the direction of modern concepts of AI. Here, as is easily seen, we are discussing the substantive prerequisites for the emergence of modern technogenic civilization. A detailed analysis of the evolution of the aforementioned ideas at later stages of European civilization can be easily found on the Internet today.

In the context of our research, it is worth noting that extrapolating the image of mathematized nature to the understanding of a human being leads to a despiritualized image of humanity. This point is rightly emphasized in the article by the national author in 2024. The foundation of AI, he writes, is the image of man as a machine:

> Deprived of the naturalness of live, direct communication, the unique thrill of spirituality and emotionality, this computerized space gives birth to and shapes a mechanical human being who functions according to the laws of mechanics and whose activities increasingly revolve around interaction with hardware. (Kozlovets, 2024, p. 57)

In understanding the unique status of AI, it’s important to keep in mind that the reduced rationality underlying artificial intelligence exerts a powerful feedback influence on how humans understand themselves. The existence of this feedback influence is rightly noted by M. Bogachov (2021): "Man imitates artificial intelligence, which is a technogenic, modified, and incidental reflection of human nature…" (p. 196).

In the context of researching the influence of mathematics on the formation of the modern picture of the world, it is difficult not to agree with Skolimowski’s (1979) correct observation about the central role of its quantitative parameters. In this context, it is also right to note that, since technology is a set of means to achieve a goal, the values of technical civilization itself are concentrated around these means. In other words, the main subject of attention and concern here is the specified means. Moreover, we see a deification of these means on the human path to perfection. Of particular interest to us is the question of what a person is and what his/her value is in this coordinate system. Bearing in mind the historical experience of mankind over the past few centuries, associated with the practical implementation of the ideas of anthropocentrism, it is reasonable to focus on the key question of whether humans have their ability to act as the demiurge of the new world, or, on the contrary, whether the role of a man is secondary and reduced to one of the means of the successful functioning of technology as a self-sufficient substance?

In analyzing this issue, it is first worth focusing on the main points of the naively optimistic interpretation of AI. Its genesis is assessed as the emergence of an intelligence of a higher level than that of men on Earth, that is, we are talking about the emergence of an entity that deprives

Creative Commons Attribution 4.0 International
doi: https://doi.org/10.15802/ampr.v0i29.365948
© A. I. Kuzmenko, N. V. Khalipova, I. Y. Lesnikova, 2026

ISSN 2227-7242 (Print), ISSN 2304-9685 (Online)
Антропологічні виміри філософських досліджень, 2026, Вип. 29
Anthropological Measurements of Philosophical Research, 2026, NO. 29

men of the status of the pinnacle of evolution. Schematically outlining their position, some authors consider it possible to speak of a modern man descending to the level of Neanderthals, whose rationality existed at a primitive level. As for the further prospects of AI evolution, they are perceived in a rosy light and do not cause any concerns. These bright prospects are proposed to be assessed as an unconditional approximation to the absolute good, even though new versions of artificial intelligence will be much smarter than humans. This is a qualitative difference, namely, humans will cease to understand and predict their work and their direction. Illustrating the naivety of the aforementioned utopias, their authors compare modern humans and chimpanzees. And since even the most intelligent chimpanzees are unable to understand 99 percent of human actions, a similar sad fate may await modern humans.

Supporters and enthusiasts of the naively optimistic vision of AI have no doubts that further spontaneous evolution of AI can lead to grandiose problems and even to the self-destruction of civilization. This concerns the tendency to completely replace the human dimension of reality, so that people will lose the meaning of life. At the heart of such ideas is the absolutization of AI's creative ability, which is capable of realizing absolutely everything a person can dream of. A powerful superintelligent machine will do this a million times better and faster, without the need for prior training. As is easy to see, this is about the prospect of eliminating humans from the creative process. In such a context, our usual ideas about human life undergo a radical transformation. From now on, human life must take place under the supervision of artificial intelligence, which is more like a computer game with open cheat codes. In the fundamentally new world created by AI, all previous problems of humanity have already been solved, all dangers and threats have already been eliminated, all laws of nature have been fully studied, all mysteries of the universe have been solved, and all possible and impossible desires have been satisfied. Men find themselves in a situation of complete absurdity when they have nothing more to strive for.

Under such conditions, a person who cannot exist in a closed space, a world transparent to the mind, devoid of secrets, has no choice in reality and is doomed to go en masse to virtual worlds. Only there does it [humanity] find everything it has grown accustomed to previous eras: there will be suffering, failures, and defeats, pain, anguish, and injustice. These are meaningfully complemented by incredible joys and discoveries, creative achievements, and personal triumphs. In other words, only there will human existence possess any meaning.

However, as the pages of attentive AI researchers attest, the above naive vision of it is subjective and biased. We imply the illusory nature of our perceptions regarding the possibility of an exhaustive definition of AI. Therefore, in the process of modern understanding of the AI phenomenon, it is right to pay attention to those deep insights, the authors of which rightly emphasize the dangers of a literal understanding of this combination. This point is rightly noted by Kozlovets, referring to Western colleagues:

> J. Lanier, a renowned inventor of our time, even claims that the use of such a term is misleading and dangerous because no 'artificial intelligence' actually exists: the latest programs, like GPT-4, only reproduce and combine the work people had done before; they can, for

Creative Commons Attribution 4.0 International
doi: https://doi.org/10.15802/ampr.v0i29.365948
© A. I. Kuzmenko, N. V. Khalipova, I. Y. Lesnikova, 2026

> example, find hidden coherence in human creations, but not create anything on their own. (Kozlovets, 2024, p. 58)

The validity of a skeptical assessment regarding the perceived total transparency of the concept of AI is axiomatic for his attentive researchers.

> There is no clear definition of artificial intelligence. Margaret T. Boden writes in her recent book *AI: Its Nature and Future* that an artificial general intelligence could have general powers of "reasoning and perception—plus language, creativity, and emotion." However, she does not forget to add that "that's easier said than done". (Morioka, 2023, p. 29)

The above-mentioned circumstance regarding the opacity of AI as a technical tool for the human mind is an additional reason for a more balanced attitude towards it today. Its significance increases substantially as attention focuses on the ways a technical mindset influences our vision of the human essence.

Analyzing the current spiritual landscape and the course of events in global politics, it is difficult to disagree that we are dealing with a practical threat to humanistic values. One of the embodiments of these threats, which is increasingly coming to the fore in modern culture, is AI. One of the key fateful questions of the current historical stage is whether humanity will manage to maintain its central status in the world. In this context, it is difficult to overlook several general points: first, the technical tools that men have so far used and continue to use to strengthen their position in the universe; second, a reminder that humanism is based on concern for the conditions of human self-realization. It is now quite obvious that, in this context, technology is often understood as a reliable means of promoting humanism. Although such optimistic expectations have become widespread, we need to go beyond these naively optimistic ideas. In the first decades of the twentieth century, Lewis Mumford (1934) rightly drew attention to the vital need for individuals to recognize the primary importance of self-care, warning against the trap of replacing it with a sense of possessiveness: "the machine is no longer a model of progress and the full expression of our desires: it is merely a series of tools" (p. 365). The author concludes his research by warning against the persistence of naive views on the progression of technology's linear advancement (Mumford, 1934, p. 434). How true is this warning about AI, that is, about our current situation, marked by the rampant expansion of AI into every domain of technology and culture?

In studying the nature of the technology's influence on the way of human behavior, it is appropriate to look back at the last few centuries, in particular at the creative heritage of representatives of non-classical philosophy of the 19th century – Arthur Schopenhauer and Friedrich Nietzsche. Their worldview pessimism is largely a form of disillusionment with the positive influence of technology on the way of human existence. One of the significant manifestations of the destructive influence of the technical mindset is the god-forsakenness of man. Fromm documented the extraordinary significance of the concept of the "sacred" for

Creative Commons Attribution 4.0 International
doi: https://doi.org/10.15802/ampr.v0i29.365948
© A. I. Kuzmenko, N. V. Khalipova, I. Y. Lesnikova, 2026

humanity in the 20th century. In the process of his critical understanding, it is worth paying attention to the original version of the definition of the uniqueness of our era in the form of a well-known aphorism about the death of man in the 20th century. Here, it is appropriate to draw attention to the fact of its meaningful connection with Nietzsche's aphorism about the death of God.

What makes Fromm's analysis particularly compelling today is his dispassionate look at how 20th-century views on humanity and technology shifted, specifically highlighting how technology – as an artificial, man-made world – became a new object of worship. Here, one should draw attention to the evocative heading of a passage in his fundamental work on human destructiveness – "The Worship of Technique and Necrophilia" (Fromm, 1973). Today, in times of radical ideological shifts, a person's need for an object of worship often takes the form of adoring AI, that is, as an attempt to restore direct access to the sacred.

The thesis about the danger of the adoration of technology is one of the dominant themes in Fromm's later work, namely, his book "To Have or to Be?" 1976. Diagnosing the threatening nature of the reductionist tendencies in the development of technocratic society, he first of all warns about the danger of transferring the technical way of thinking to humans. Realizing this point, he warns: "Dehumanized man will become so insane that he will not be able to maintain a viable society in the long run, nor will he be able, in the short run, to refrain from the suicidal use of nuclear or biological weapons" (Fromm, 2020, p. 288).

In Fromm's opinion, in the face of the increasingly obvious negative and destructive consequences of the cult of technology, humanity must work to develop a meaningful alternative. This is, he wrote 50 years ago, about the creation of a kind of supreme council, consisting of humanitarians and empowered to adopt recommendations in the process of understanding the practical problems of today: *"A Supreme Cultural Council should be created, which would advise the government, politicians, and the public on all questions where knowledge is needed"* (Fromm, 2020, p. 282).

Unfortunately, space does not permit a detailed examination of Fromm's concept here; however, it is fitting to at least provide the summary he formulates at the conclusion of "To Have or to Be?" Humanity today, as this firmly convinced humanist believes, must seek to create a synthesis of previous eras, crowned by the image of spiritual rationality: "the synthesis of the spiritual core of the Late Medieval world and the development of rational thought and science since the Renaissance. This synthesis is the City of Being" (Fromm, 2020, p. 295). It would seem that we live in a qualitatively different era today, which implies other ethical principles. However, although we are separated from Fromm's era by fifty years, the idea of the importance of ethics in determining the way a person relates to technology is axiomatic even today. What is the type of rationality inherent in human nature?

The Ukrainian thinker Serhiy Krymskyi proposed a deep and meaningful concept regarding this problem. In developing it, he proceeded from the widespread, superficial ideas about the universality of reduced rationality dominant in technological civilization. Our philosopher interprets the naive exaggeration of the significance of reduced rationality as a threat to civilization, since the world is not fully rationalized without a residue. Therefore, it is difficult not to recognize the correctness and relevance of the domestic thinker's thesis about the expediency of a broader understanding of rationality. And although positivist approaches to interpreting society and culture are currently widespread, Krymskyi writes that "we must be the kind of thinkers who serve as guardians of spiritual sanctuaries". In other words, the Ukrainian

Creative Commons Attribution 4.0 International
doi: https://doi.org/10.15802/ampr.v0i29.365948
© A. I. Kuzmenko, N. V. Khalipova, I. Y. Lesnikova, 2026

scholar links his search for a meaningful antithesis to the narrow understanding of rationality with eternal spiritual sanctuaries. It is about "rationalism, addressed to the eternal sanctuaries of existence – Freedom, Love, Personality (as the earthly hypostasis of the "infinite face"), which turns out to be involved in the world-historical affirmation of the SPIRIT" (Krymskyi, 2003, p. 109). We have no right to forget that the renewed rationality of the present includes the problem of men. Meaningfully outlining the antithesis of reductionism as the idea of the universality of instrumental rationalism, Krymskyi turns to the specifics of Ukrainian philosophy as the most expressive form of its manifestation. It is about "spiritual reason" (Krymskyi, 2008, p. 351).

Studying the collisions and transformations that the concepts of rationality and a man as a rational being undergo at the beginning of the 21st century, Krymskyi rightly emphasizes the narrowness of their superficial reception. We find it particularly valuable that he turns to AI to illustrate the unique philosophical demands of our century. The latter is closely linked to the destruction of the traditional human monopoly on the status of the sole bearer of intelligence. The danger of the rapprochement and identification of humanity and intelligence lies in the fact that morality remains in the shadow as a manifestation of spiritual rationality as a meaningful "alternative to machine intelligence. Humans are primarily beings possessed of an inner world, spirituality, their own destiny, and a commitment to higher values" (Krymskyi, 2008, p. 348). The above thesis assumes a responsible attitude of modern philosophers to the implementation of the task in protecting human spirituality. How do contemporary scholars interpret the problem of responsibility?

A striking example of human concern regarding the existing ambiguous prospects for the further development of technogenic civilization is Hans Jonas's research in 1979 (Ukrainian translation in 2001). For us today, the fact that he focused his main attention on the significance of the artificial environment is of fundamental importance. He qualifies the situation of *homo faber* dominance over *homo sapiens* as a challenge to the development of a new ethics: "… technology acquires ethical significance… Its cumulative creativity, namely the *creation of an artificial environment*… (emphasis added – A. K., N. K., I. L.) strengthens those special forces that it (men) have caused…" (Jonas, 2001, p. 24). We consider it appropriate to draw attention to the fact that today, the embodiment of the artificial environment is AI. The destructive tendency in the development of technical civilization, as indicated by Jonas, causes him anxiety and concern. Since the author largely identifies modern man with a traveler who has lost his bearings, he tends to define his own duty as an appeal to the compass. In other words, we are talking about a situation of "value relativism", the overcoming of which, for the author, is associated with the substantiation of a new ethic. Here, it is appropriate to draw attention to the subtitle of this research: "In Search of Ethics for Technological Civilization". Jonas means his own mission in the Preface with the word "warning". It is worth noting that this genre gained popularity already in the middle of the twentieth century. This is, in particular, the book by Jacques Ellul "Technique or the Challenge of the Century" of 1954, which is rightly qualified as one of the most important scientific works of the second half of the twentieth century. The purpose of this book is representative – "a call to the sleeper to wake up".

Considering Jonas's proposed constructive options for a meaningful response to the destructive tendencies in the development of technical civilization, we should pay attention to his way of understanding man. For him, it is more than a carrier of reduced rationality; it is a spiritual being. And since, in the conditions of the development of technogenic civilization, the

Creative Commons Attribution 4.0 International
doi: https://doi.org/10.15802/ampr.v0i29.365948
© A. I. Kuzmenko, N. V. Khalipova, I. Y. Lesnikova, 2026

orientation towards the transformation of the world is deterministic, Jonas feels personal responsibility for the future of humanity. For him, the unambiguous connection between a narrow understanding of rationality and the mass extermination of people in concentration camps was axiomatic. It exists, he writes, "as a consequence of the development of 'value-free' rationality – science, technology, economics, politics, etc." (Jonas, 2001, p. 360).

It is important for us that our current situation of AI expansion is analogous to the one that Jonas analyzes. As is easy to see, he is looking for ways to humanize instrumental rationality. The aphoristic definition of modern man is that he must "learn to fear" (Jonas, 1987). It is axiomatic that here Jonas, like other representatives of humanistically oriented philosophy, is looking for ways to go beyond the limits of reduced rationality constructively. A more detailed analysis of his position is beyond the scope of our article.

Originality

This study examines the nature of the threat AI poses to humanity, stemming from the naive acceptance of instrumental rationality as universal, alongside the neglect of inherent spiritual rationality. A defining feature of the modern interpretation of AI is its abstract vision of reality, specifically a focus on quantitative parameters. This shift away from humanistic values and meanings creates a prerequisite for reducing humans to a mere means of servicing AI. Furthermore, the text argues that the works of 20th-century humanists on the phenomenon of technology hold significant heuristic potential for understanding the problem of AI responsibility in those texts of 20th-century humanists devoted to the phenomenon of technology is argued.

Conclusions

The study reveals that the irresponsible interpretation of AI stems from the naive assumption of the universality of instrumental rationality and a nihilistic view of spiritual rationality. Substantive prerequisites for the prevalence of such optimistic interpretations include an abstract vision of reality – focusing exclusively on quantitative parameters – and a conceptualization of man as the center of the universe. Distinguishing the reduced rationality underlying AI from the spiritual rationality inherent in humans allows for the specification of "red lines" that manifest a responsible attitude toward spontaneous AI development. The authors identify prospects for further research in exploring the nature of human freedom and responsibility at this historical stage.

REFERENCES

Annet, K. A. (2025). The ethics of AI: Philosophical perspectives. *Research Invention Journal of Research in Education, 5*(2), 65-71. DOI: https://doi.org/10.59298/RIJRE/2025/526571 (in English)

Barnes, E., & Hutson, J. (2024). A framework for the foundation of the philosophy of artificial intelligence. *International Journal of Recent Engineering Science, 11*(4), 113-126. DOI: https://doi.org/10.14445/23497157/ijres-v11i4p114 (in English)

Bogachov, M. (2021). Artificial intelligence as an anthropotechnology. *Filosofska Dumka (Philosophical Thought),* (3), 180-200. DOI: https://doi.org/10.15407/fd2021.03.180 (in Ukrainian)

Fromm, E. (1973). *The anatomy of human destructiveness*. Holt, Rinehart and Winston. (in English)

Fromm, E. (2020). *To have or to be?* (H. Yanovska, Trans.). Kharkiv: Klub simeinoho dozvillia. (in Ukrainian)

Harari, Y. N. (2025). *Nexus: A brief history of information networks from the Stone Age to AI* (N. Khaietska, Trans.). Kyiv: BookChef. (in Ukrainian)

Jonas, H. (1987). *Technik, Medizin und Ethik: Zur Praxis des Prinzips Verantwortung*. Frankfurt am Main: Insel. (in German)

Creative Commons Attribution 4.0 International
doi: https://doi.org/10.15802/ampr.v0i29.365948
© A. I. Kuzmenko, N. V. Khalipova, I. Y. Lesnikova, 2026

Jonas, H. (2001). *Das Prinzip Verantwortung: Versuch einer Ethik für die technologische Zivilisation* (A. Yermolenko & V. Yermolenko, Trans.). Kyiv: Libra. (in Ukrainian)
Kozlovets, M. (2024). Artificial Intelligence Technologies and Their Impact on Human Self-Being. *Humanities studies*, (19), 55-66. DOI: https://doi.org/10.32782/hst-2024-19-96-06 (in Ukrainian)
Krymskyi, S. (2003). *Zapyty filosofskykh smysliv*. Kyiv: PARAPAN. (in Ukrainian)
Krymskyi, S. (2008). *Pid syhnaturoiu Sofii*. Kyiv: Vydavnychyi dim "Kyievo-Mohylianska akademiia". (in Ukrainian)
Morioka, M. (2023). Artificial intelligence and contemporary philosophy: Heidegger, Jonas, and slime mold. *Journal of Philosophy of Life, 13*(1), 29-43. (in English)
Mumford, L. (1934). *Technics and civilization*. New York. (in English)
Skolimowski, H. (1979). Philosophy of technology as a philosophy of man. In G. Bugliarello & D. B. Doner (Eds.), *The history and philosophy of technology* (pp. 325-336). University of Illinois Press. (in English)

LIST OF REFERENCE LINKS

Annet K. A. The Ethics of AI: Philosophical Perspectives. *Research Invention Journal of Research in Education*. 2025. Vol. 5, Iss. 2. P. 65–71. DOI: https://doi.org/10.59298/RIJRE/2025/526571
Barnes E., Hutson J. A Framework for the Foundation of the Philosophy of Artificial Intelligence. *International Journal of Recent Engineering Science*. 2024. Vol. 11, Iss. 4. P. 113–126. DOI: https://doi.org/10.14445/23497157/ijres-v11i4p114
Богачов М. Штучний інтелект як антропотехнологія. *Філософська думка.* 2021. № 3. С. 180–200. DOI: https://doi.org/10.15407/fd2021.03.180
Fromm E. *The Anatomy of Human Destructiveness*. Holt, Rinehart and Winston, 1973. 521 p.
Фромм Е. *Мати або бути?* / пер. з англ. Г. Яновська. Харків : Клуб сімейного дозвілля, 2020. 303 с.
Харарі Ю. Н. *Nexus. Коротка історія інформаційних мереж від кам'яного віку до ШІ* / пер. з англ. Н. Хаєцька. Київ : BookChef, 2025. 640 с.
Jonas H. *Technik, Medizin und Ethik: Zur Praxis des Prinzips Verantwortung*. Frankfurt am Main : Insel, 1987. 323 s.
Йонас Г. *Принцип відповідальності. У пошуках етики для технологічної цивілізації* / пер. А. Єрмоленко, В. Єрмоленко. Київ : Лібра, 2001. 400 с.
Козловець М. Технології штучного інтелекту та їх вплив на буттєвість людини. *Humanities studies*. 2024. № 19 (96). С. 55–66. DOI: https://doi.org/10.32782/hst-2024-19-96-06
Кримський С. *Запити філософських смислів*. Київ : ПАРАПАН, 2003. 240 с.
Кримський С. *Під сигнатурою Софії*. Київ : Видавничий дім "Києво-Могилянська академія", 2008. 367 с.
Morioka M. Artificial Intelligence and Contemporary Philosophy: Heidegger, Jonas, and Slime Mold. *Journal of Philosophy of Life*. 2023. Vol. 13, No. 1. P. 29–43.
Mumford L. *Technics and Civilization*. New York, 1934. 445 p.
Skolimowski H. Philosophy of Technology as a Philosophy of Man. *The History and Philosophy of Technology* / ed. by G. Bugliarello, D. B. Doner. University of Illinois Press, 1979. P. 325–336.

А. І. КУЗЬМЕНКО[1*], Н. В. ХАЛІПОВА[2*], І. Ю. ЛЕСНІКОВА[3*]

[1*]Університет митної справи та фінансів (Дніпро, Україна), ел. пошта albinakuzmenko03@gmail.com, ORCID 0000-0001-7278-3647
[2*]Університет митної справи та фінансів (Дніпро, Україна), ел. пошта khalipova@ukr.net, ORCID 0000-0001-5605-6781
[3*]Університет митної справи та фінансів (Дніпро, Україна), ел. пошта lesnikova@i.ua, ORCID 0000-0002-2750-6031

ШІ та відповідальність людини

Мета – проаналізувати запит сучасності на вивчення форм відповідального ставлення до розвитку штучного інтелекту (ШІ). Реалізація цієї мети передбачає розв'язання наступних завдань: а) з'ясувати основні складові сучасного тлумачення ШІ, б) вивчити природу загрозливості ШІ для подальшого розвитку людства. **Теоретичний базис.** Теоретико-методологічну базу статті склали принципи гуманізму, світоглядного плюралізму та об'єктивності. **Наукова новизна.** Вивчено природу загрози для людини з боку ШІ, яка

Creative Commons Attribution 4.0 International
doi: https://doi.org/10.15802/ampr.v0i29.365948
© A. I. Kuzmenko, N. V. Khalipova, I. Y. Lesnikova, 2026

пов'язана з наївною рецепцією інструментальної раціональності як універсальної, що супроводжується ігноруванням притаманної людині духовної раціональності. Своєрідність сучасного тлумачення ШІ пов'язана з абстрактним баченням реальності, а саме – зосередженням основної уваги на її кількісних параметрах. Йдеться про відхід від гуманістичних цінностей та смислів, що є передумовою редукування людини до ролі засобу для обслуговування ШІ. Аргументовано наявність евристичного потенціалу для осмислення проблеми відповідальності ШІ в тих текстах гуманістів ХХ ст., які присвячені феномену техніки. **Висновки.** Виявлено, що в основі безвідповідального тлумачення ШІ лежить наївне припущення про універсальність інструментальної раціональності та нігілізм щодо духовної раціональності. З'ясовано, що до числа змістовних передумов поширеності наївно-оптимістичних тлумачень ШІ належить абстрактне бачення реальності, що полягає в зосередженості уваги виключно на кількісних параметрах світу, що доповнюється висуванням людини на перший план як центру світобудови. Осмислення істотної відмінності між редукованим типом раціональності, який лежить в основі ШІ та притаманною людині духовною раціональністю, дозволяє нам конкретизувати уявлення про ті червоні лінії, які є проявою відповідального ставлення людини до спонтанних тенденцій розвитку ШІ в сучасній цивілізації. Перспективи подальших досліджень автори вбачають у осмисленні особливостей існування свободи та відповідальності людини на даному етапі історії.

Ключові слова: людина; штучний інтелект (ШІ); інструментальна раціональність; духовна раціональність; етика; священне

Received: February 20, 2026
Accepted: April 8, 2026
Published: June 30, 2026

Creative Commons Attribution 4.0 International
doi: https://doi.org/10.15802/ampr.v0i29.365948
© A. I. Kuzmenko, N. V. Khalipova, I. Y. Lesnikova, 2026