FUTURITY PHILOSOPHY

DOI: https://doi.org/10.57125/FP.2022.06.30.05

How to cite: Nikolenko, K. (2022). Artificial Intelligence and Society: Pros and Cons of the Present, Future Prospects. *Futurity Philosophy, 1*(2), 54-67. https://doi.org/10.57125/FP.2022.06.30.05

Artificial Intelligence and Society: Pros and Cons of the Present, Future Prospects

Kseniia Nikolenko

PhD in Philosophical sciences, Associate Professor, Department of Social Humanities and Legal Disciplines, Kryvy Rih Educational and Scientific Institute, faculty №1, Donetsk State University of Internal Affairs, Kryvyy Rih, Spivdruzhnosti str., 92 A, Nikolenkoksenia985@gmail.com, 0000-0002-2091-7823

Received: March 28, 2022 | **Accepted:** May 15, 2022 | **Published:** June 30, 2022

Abstract: Aims: The purpose of this scientific research is to provide a philosophical interpretation of the positive and negative characteristics of artificial intelligence that have been formed in the scientific and cultural discourse. It is noted that the research objectives are not focused on the usual analysis to determine the pros and cons of artificial intelligence potential. The philosophical interpretation of the status of artificial intelligence is to look for its negative aspects in the advantages and positive aspects in the disadvantages of this potentially global sociocultural phenomenon. Methodology: The study used general scientific, cultural, and philosophical methods. An obvious feature was the philosophical and methodological principle of reversal - the reverse (and in some cases, inverse proportional) interaction between traditional intellectual research activity and innovative artificial intelligence. Results: According to the results of the study, a format for assessing the advantages and disadvantages of artificial intelligence through the prism of dialectical opposition and synergistic interaction is proposed. The model of scientific and philosophical analysis of artificial intelligence is formed by axiological, epistemological, methodological, and ontological characteristics. Scientific Novelty: A promising area of research is the correlation of artificial intelligence tools with innovative social concepts of the new humanistic paradigm. In the modern scientific picture of the world, the status of artificial intelligence is at the stage of worldview verification, so the prospects of this innovation are

possible only if it is recognised at the level of public consciousness. Conclusion: Philosophy is called upon to form the value, scientific, cognitive, methodological (and in the future, existential) dimensions of artificial intelligence in the human-dimensional worldview paradigm. The process of forming these constants is possible only if the pros and cons of artificial intelligence in the system of social and individual activity and consciousness are interpreted holistically.

Keywords: artificial intelligence, new rationality, axiological dimensions of technologization, innovativeness, philosophy of science, philosophy of technology, scientific and technological picture of the world.

Introduction

The intellectual potential is fundamental in the scientific and worldview paradigm. This problem has always been of dominant interest to philosophers. The study of human intelligence has confirmed the uniqueness of humans who possess this potential and has shaped the anthropocentric worldview.

Given that artificial intelligence is a reflection of human intelligence, philosophy defines two of its positions in the scientific and worldview picture: the weak and strong dimensions (Lee, 2020). Two definitions of artificial intelligence are used in the scientific literature: general artificial intelligence (AGI) and artificial niche intelligence (ANI). Fjelland (2020) notes that AGI is classified as weak AI, ANI, in turn, is strongly pronounced AI and more controversial in terms of its impact on humans and society.

In human-machine interaction: "human-machine" interaction, the fact that technology depends on humans is crucial (Mühlhof, 2019). The reverse dependence is purely hypothetical and poses a threat only as one of the possible scenarios. Currently, the idea that artificial intelligence contrasts with natural intelligence is generally denied (Behrens & Sandis, 2020). Rather, the imitation of intelligence is realised with the highest possible level of naturalness.

Livingston & Risse (2019) ask a pertinent question about human-technology interaction: "What do machines owe to humans, and what do humans owe to superintelligent machines?" To answer this question, it is necessary to realise that a human being in modern philosophical understanding is not a generalized image endowed with standard qualities. The problem of threats from artificial intelligence, described in detail in the works of scientists and thinkers, is generalized to humans. But modern philosophy proposes differentiation of human qualities, which makes it impossible to standardise the impact of AI on humans.

Research Problem

The key research problem in the study is focused on the public perception of the phenomenon of artificial intelligence not only as a scientific and technological tool but also as an element of influence on global consciousness. Nowadays, there is a total impact of information and digital technologies on the individual consciousness of an individual user. This approach generates a large number of individual disparate reactions to the use of artificial intelligence. Philosophical reflection makes it possible to systematize this disparate perception of artificial intelligence in the context of the worldview integrity of this phenomenon. To implement this systematization, it is necessary to define the value, cognitive and methodological characteristics of artificial intelligence.

Ferrario et al. (2020) propose a model of gradual trust in artificial intelligence. This process begins with the interaction between humans and artificial intelligence in scientific, specialised, and technological aspects. Later on, trust extends to the everyday level, where every consumer of information and digital technologies can realize the role (and, in general, the negative manifestations) of artificial intelligence. Finally, the final level of interaction is formed at the fundamental worldview

level, when artificial intelligence is positioned as a generally accepted phenomenon that has a comprehensive perception (positive or negative) on the part of individual and social consciousness.

At present, the scientific interpretation and positioning of artificial intelligence is dominant. However, with the deeper integration of artificial intelligence into the practical and everyday dimension, the need for a philosophical explanation of its nature and functioning will grow. As of today, the mutual coordination of the existence of artificial intelligence in the sociocultural space is at the verge of scientific and technological manifestation and transition to the practical and everyday plane of its use.

Research Focus

The focus of the study is to identify the following aspects:

- the positive potential of artificial intelligence in the socio-cultural sense and a specific philosophical and analytical dimension;
- negative aspects of artificial intelligence on a global civilisation scale and threats to this phenomenon as an innovative element of scientific and technological progress.

Philosophy deals with the issue of artificial intelligence not only to organise this phenomenon in the scientific and worldview paradigm. It is no secret that science has ambitious plans to continue working on the creation of a technological environment. In particular, the issue of creating artificial consciousness is being considered (Chowdhary, 2020). If this scientific and technological project is successful, there will be a need to completely transform the worldview paradigm with new human-dimensional principles.

Research Aim and Research Questions

The purpose of the research is to provide a philosophical analysis of the pros and cons of artificial intelligence in the axiological, epistemological, methodological, and existential dimensions. The purpose of the article is not to make a cultural statement of the positive and negative manifestations of AI in society but to search for the opposite effects of its functioning. This is how the holistic characteristics of artificial intelligence are formed, not only with regard to the practically oriented aspects of its use, but also with regard to the intended purpose of information and digital technologies.

Literature Review/Theoretical Overview

In scientific and philosophical thought, the object of study of actual phenomena and processes is traditionally divided. Science focuses on laws (in the era of innovation - on technologies), while philosophy concentrates on the moments of building worldview perception. This format has been preserved in the issue of artificial intelligence, taking into account the growing role of the technological component (Sloman, 2019). Attempts to avoid a polarized space in the study of artificial intelligence are made in Abedin (2022). Identifying the opposite effects of artificial intelligence activity allows reducing the growing tension over the threats of this phenomenon to society.

The problem of the purpose of artificial intelligence is beginning to prevail over its functionality, which leads to a qualitatively new positioning in the scientific and philosophical discourse (Krkač, 2019). The literature review of the problem of artificial intelligence can be divided into the following clusters:

- scientific and cultural discourse on the establishment of artificial intelligence;
- philosophical and synergistic approach that defines the functionality of artificial intelligence.

Artificial intelligence has gone through several stages in scientific and cultural coverage. At first, the scientific and technological characteristics of the information and digital element prevailed. Later, cultural and ideological interpretations of artificial intelligence gained popularity. Currently, there is a return to the computer science and mathematical coverage of artificial intelligence (Frank et al., 2019).

Science, in solving highly complex interdisciplinary problems, needs to be recognized and agreed upon at the level of public consciousness. That is why science needs a philosophy that has been creating public opinion throughout its development, ensuring the loyalty of individuals and communities to scientific successes and failures (Laplane et al., 2019). In addition, artificial intelligence is used in practice, in the real sector of production, which makes it even more important (Abioye et al., 2021).

Post-digital science involves a transition from technological dominance to the interaction of the natural (human) and technological (digital). The convergence between the information technology and natural-biological domains allows for the formation of a separate ecosystem in which artificial intelligence occupies its unique niche (Peters et al., 2022).

Dreyfus (1992) was perhaps the first in the scientific and philosophical discourse to propose a discussion of the problems associated with the use of artificial intelligence, actualising the issue of imperfections in computer intelligence. The criticism of artificial intelligence was actively taken up by the scientific and philosophical community at the turn of the twentieth and twenty-first centuries, during the period of the rapid spread of information and digital technologies.

The large-scale introduction of artificial intelligence has provoked a significant number of discussions on this phenomenon. Particularly, the author proposes to consider the positive and negative aspects of information and digital activities. It is noted that in terms of culture, all information related to artificial intelligence is perceived. The philosophical interpretation in scientific discourse does not focus on quantitative indicators, acting according to the principle of "analyse problems and remove non-problems" (Müller, 2021).

Feng (2019), notes that philosophy's interest concerning the problem of artificial intelligence is also indicative of a crisis in philosophy itself. The deficit of philosophical thinking in society requires the actualisation of such concepts as "intelligence," "consciousness," and "subjectivity" through the prism of conventionalism.

Islam et al. (2022) emphasise that artificial intelligence (AI) and machine learning (ML) have been thoroughly described in recent years in the context of their direct application. At the same time, there is an obvious lack of secondary research. A promising area for further scientific and philosophical research in the field of AI functioning would be to characterise its strengths and weaknesses (rather than its positive or negative aspects). Such positioning in the literature will allow us to form a holistic view of artificial intelligence in the worldview paradigm.

Research Methodology

General Background

The diversity of the methodological arsenal of the study of the problem of artificial intelligence was ensured through the use of scientific methods of synergistic measurement. Namely, with the help of systematization and modelling, an attempt was made to create a structured system of scientific and ideological ideas that determine the peculiarities of the establishment of artificial intelligence. Consideration of the positive and negative manifestations of artificial intelligence in the scientific and cultural sense is possible due to a clear definition of the status and features of this phenomenon.

Somewhat different methodological approaches are actualised to achieve specific goals of the work. When determining the advantages of negative manifestations of artificial intelligence and the disadvantages of the positive role of this phenomenon, classical methodological approaches lose their relevance. Since the principle of research from the opposite has been actualised in this work, the general scientific methodology was focused primarily on the potential of comparative analysis. In this regard, the study actively uses the philosophical and dialectical method of identifying fundamental contradictions of a scientific and ideological nature regarding the functioning of artificial intelligence.

It should be noted that modern science is increasingly facing problematic methodological manifestations. In particular, the model of science is being updated, in which the rules of substantiation of phenomena and processes are based on the principles of non-contradiction and impossibility of verification of truth (Floridi, 2019). Artificial intelligence in this context is actually an exemplary example of methodological uncertainty and imperfection of scientific and epistemological activity. Dick (2019), systematising the scientific definitions of artificial intelligence, generalised this concept to the reproduction of intelligent human behaviour in a machine.

Taxonomic systematisation is considered to be a relevant methodological approach to the positioning of artificial intelligence, as seen in the implementation of the AI Watch model in the European Community (Samoili et al., 2020). Taxonomy provides flexibility in monitoring the process of artificial intelligence development. This methodological principle is relevant for the presented research since the consideration of positive and negative manifestations of artificial intelligence is more objective in the dynamic dimension.

Gabriel (2020) proposes a methodology for aligning the concepts of general and artificial intelligence. According to this principle, an attempt is made to establish interaction in a society that uses the intellectual arsenal from all possible sources.

Science and philosophy are constantly striving to assess all phenomena and processes that have a global impact on civilizational progress. Usually, their methods intertwine and interact, forming innovative methodological approaches. In the study of artificial intelligence, such a methodological approach has become neurophilosophy, which in the course of research activities has been transformed into neuro-techno-philosophy (Al-Rodhan, 2023).

In the methodological dimension, artificial intelligence is characterized by opacity, which does not allow to fully explaining the essence of its activities (Zednik, 2021). Science is able to explain the technological component of artificial intelligence but will not be able to determine the activation of other aspects of this phenomenon.

Research Results

Throughout the development of civilisation, philosophy has been distinguished by its specific coverage of current processes. In the scientific and worldview picture of the world, every global phenomenon received a comprehensive socio-cultural assessment. At the same time, the scientific and cultural analysis did not provide a complete understanding of the potential and prospects of a particular process or phenomenon. Typically, cultural interpretation identifies fundamental positive and negative qualities. Such a paradigm is appropriate when the development of a cultural and historical epoch is stable and balanced. When the socio-cultural space is in the process of transformation or in times of radical transformations or changes in cultural and historical epochs, the model for assessing and characterizing key social processes undergoes significant changes. The main scientific and methodological feature in times of instability is the flexibility of views and the relativity of assessments.

The modern world is still far from changing the cultural and historical era, but it is characterised by ongoing fundamental transformations of the socio-cultural system. This state of affairs requires flexibility in assessing global trends in the development of modern civilisation. Having a full-fledged scientific and cultural analysis of processes and phenomena, the author states the need for their specific philosophical interpretation.

Artificial intelligence as an element of information and digital technologies is one of the most pressing problems of our time. The potential of artificial intelligence forces the scientific and philosophical community to thoroughly study it as a phenomenon, despite the fact that it is not actually realized by 1% of its practical potential.

The scientific and cultural analysis of artificial intelligence is widely represented in contemporary literature - scientific, journalistic, fiction, or even advertising. These works describe in detail the technological characteristics of artificial intelligence, its scientific and cognitive value, and the practical and everyday benefits of its application. Considerable attention is also paid to the threats associated with the use of artificial intelligence.

The results of the study attempt to characterise artificial intelligence not in terms of the fundamental development of culture, but in the context of flexibility characteristic of a dynamic innovative socio-cultural environment. It is a philosophy that can offer specific characteristics (see Table 1).

Table 1

Philosophical and ideological characteristics of artificial intelligence

1. methodological	1. Artificial intelligence is a model of development that focuses on innovative means of acquiring knowledge and tools. The dominance of the human dimension of scientific and cultural activity is levelled in the cluster of practice-oriented activities. At the same time, the human-centred target vector of socio-cultural development is preserved.
2. gnosiological	2. Cognitive and search activity is centred on the potential of information and digital technologies, in which artificial intelligence is not only a tool but also an organizer (creator).
3. axiological	3. Artificial intelligence gets the opportunity not only to create cultural artefacts but also to influence their further positioning and status in the scientific and ideological paradigm.

4. Existential	4. The existence of artificial intelligence becomes relevant not because of its effectiveness, but, in fact, because of the statement of its activity, which equates it with the uniqueness of a person.

Source: author's own development

The key problem of the methodological and cognitive nature for artificial intelligence is the formality of decision-making (Dreyfus, 1997). In the era of the New European scientific and philosophical rise, a similar dichotomy was traced in the issue of eliminating the sensual from rationalistic scientific knowledge. Páez (2019) proposes to change the methodological principles of artificial intelligence functionality. It is believed that the interpreted decisions of AI should be reformulated into the principles of pragmatic naturalistic understanding.

In its epistemological manifestation, artificial intelligence is clearly expressed in deep learning, which is possible through the use of an expanded range of rationalistic and empirical tools of knowledge (Buckner, 2023). Hehl (2021) sees artificial intelligence (especially its potential ability to surpass natural intelligence) as a chance to prove that chance prevails over fundamentality.

The advantage of natural human intelligence over its artificial counterpart is its increased ability to adapt to rapid changes in the sociocultural environment (Mohamed et al., 2020). Artificial intelligence has not yet had the opportunity to demonstrate its effectiveness in times of global transformation. This raises certain doubts about the realisation of the benefits of artificial intelligence in a crisis, where the natural dimensions of self-preservation, self-awareness, and self-organisation, in general, are actualised.

An important component in the assessment of artificial intelligence in a worldview system is information (Lepri, 2021). Recent years have seen a significant increase in the flow of information in two aspects:

- big data becomes a source for artificial intelligence, expanding its functionality;
- humans and society receive much more information (including practically oriented results) about the activities of artificial intelligence.

The transition to the introduction of artificial intelligence elements into the public system leads to a reorientation of the value norms and rules of this phenomenon (Coeckelbergh, 2019). An important nuance of the characterisation of artificial intelligence is the realisation of the difference between individual and public perception of this phenomenon (List, 2021).

The globalised world is searching for non-standard solutions to the issue of developing public consciousness. In particular, there are attempts to raise the level of morality through the use of biotechnology (Lara & Deckers, 2020). This approach is actually controversial since moral and ethical norms traditionally have their own guidelines for establishing themselves in the socio-cultural paradigm. The use of an information and digital resource for the implementation of the human dimensional function is an innovative solution of modern science. Sutrop (2019) proposes to clearly delineate the principles of AI ethics, prioritising trust over granting this phenomenon a normative status. This approach should protect society from potential conflict between artificial intelligence and humans.

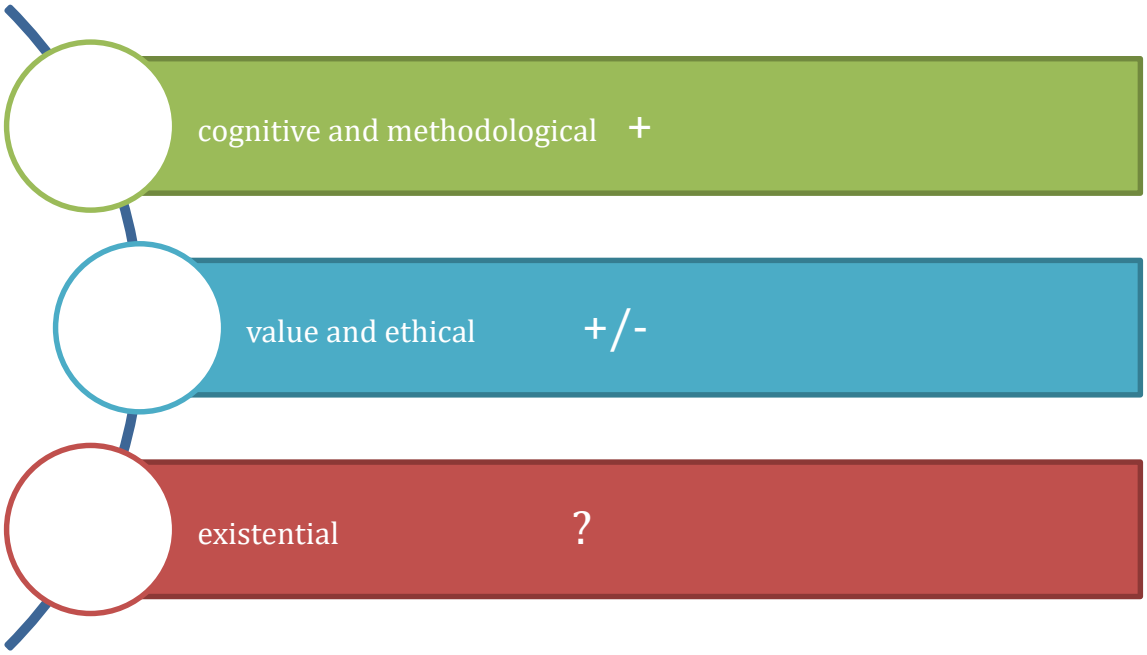
The ability of artificial intelligence to learn and improve itself helps reduce the human influence on its activities (Shaw, 2019). The danger predicted by most experts is related to the potential inability to control AI. However, the ability of humans to deactivate or destroy artificial intelligence at the practical and technological level is not mentioned.

Discussion

Having examined the characteristics of artificial intelligence, the scientific and philosophical community has actualised the discussion on the sequence of actualisation of the key components of information and digital technologies. The discussion is about prioritizing the main philosophical characteristics of artificial intelligence.

Figure 1

An algorithm for the development and establishment of artificial intelligence in the paradigm of public consciousness



Source: author (s) own development

Currently, the philosophical characterisation of artificial intelligence gives it a positive status in the cognitive and methodological dimension. The issue of value and ethical elements is in a state of contradiction between pragmatism and ethics. The existential prospects of AI are generally questionable since granting it ontological status immediately actualizes irreversible changes and threats to humanity on a global scale.

The modern pragmatic world is enthusiastically observing the ability of artificial intelligence to perform highly complex tasks (Brendel et al., 2021). In the philosophical dimension, the primacy of efficiency is gradually beginning to dominate other elements of the world picture, especially the moral and ethical aspect.

The task of contemporary thinkers is to promote the principles of harmonised artificial intelligence (Berberich et al., 2020). Humans are fascinated by the perfection of artificial intelligence, which forms false axiological constants (Shank et al., 2019). However, Schneider (2019) argues that

artificial intelligence does not compete with the natural qualities of humans but leads the epistemological potential to new directions of development. Humanistic guidelines are oriented towards progress (provided by science and technology), but with the preservation of the fundamental principles of humanity (the role of philosophy and culture). That is why the modern scientific and philosophical paradigm is in a state of confrontation over the status of artificial intelligence. The cultural and philosophical analysis states the need for AI and other components of innovative technologies but warns against the uncontrolled spread.

The key contradiction of artificial intelligence from the point of view of philosophical understanding is the expediency of its creation not as a current methodological and epistemological tool, but as a prospect for expanding human mental and physical abilities (Farrow, 2019). In this context, the prospects look doubtful, since science has not yet been able to fully explore the potential of human intelligence. There is a classical quantitative multiplication of entities (according to the principle of Occam's razor) without giving them qualitative characteristics. With the limitless possibilities of human intelligence, the science and technology cluster does not study them but develops another similar entity - artificial intelligence, which will also need to study opportunities and risks, provide prospects and assessments.

Abbass (2019) points to the transformation of technological progress: “from encoding human knowledge into a machine to developing machines that know how to autonomously acquire the knowledge they need”. Artificial intelligence duplicates human qualities in an improved and refined form. While everything is more or less clear with brain activity, the functionality of the physical body raises a number of controversies. It is clear that artificial intelligence is not used for physical labor, but the ability of technology to support vital activity, and with it brain activity, is an important condition for its effectiveness (Su & Luvaanjalba, 2021). The issue of liability for the adverse effects of artificial intelligence also remains uncertain (de Sio & Mecacci, 2021). The problem of liability is exacerbated by the lack of direct application of legal norms to a non-human object (Gellers, 2020). From the perspective of axiological principles, responsibility lies entirely with humans, as they created and coordinate the operation of technology. On the other hand, if we consider the activity of artificial intelligence from the point of view of the scientific and methodological approach, it is autonomous and capable of taking independent steps.

Scientists emphasise the potential risks of AI bias, which is a consequence of the inability to analyse available data from a moral perspective (Lin & Chen, 2022). Such realities will lead to the gradual accumulation of the injustice factor (not de jure, but de facto). It is quite difficult to understand whether artificial intelligence is able to predict the public reaction in its actions, let alone distinguish it from the public response.

Scientists emphasise the need to classify elements of information and digital technologies with a clear division into factors that are limited to practice-oriented influence and those that are capable of existential expression (Zhou & Razak, 2022). Hui (2021) presents a variety of interpretations of the limits and status of artificial intelligence: from the technological modelling of conventional intelligence theory to an autonomous self-sufficient worldview element. The question that philosophers are concerned about is not whether artificial intelligence poses a threat (this potential danger is also understood by scientists), but rather the ability of artificial intelligence to surpass human intelligence, which is an existential threat (Clune, 2019). In this context, it is important to understand the attitude of humans and society towards the most acceptable (or more convenient) status of artificial intelligence.

If artificial intelligence is viewed as an expensive multifunctional toy, then it is not allowed to acquire a philosophical and existential status. If we are guided by the principle of humanistic (or rather

anthropocentric) priorities, then artificial intelligence is not an alternative to a human competitor (Cope et al., 2021). Auernhammer (2020) clearly defines the human-centred nature of artificial intelligence. At the same time, artificial intelligence is intended to take over some of the functional functions. In this context, a relevant question arises: is there a need for such a replacement? The statement of the primacy of human intelligence over artificial intelligence immediately removes the strategic positioning of these elements of philosophical discourse. If considering the pragmatic role of artificial intelligence, then we should completely eliminate the existential dimensions of artificial intelligence from the worldview paradigm. As for the value and ethical cluster, the evaluative and normative vector should be directed to artificial intelligence as a functional element, not a global axiological unit. Thus, in the status in which artificial intelligence and its prospects are positioned in the modern world picture, philosophical interpretation is concentrated on the practically oriented epistemological and methodological level. The axiological philosophical dimension of artificial intelligence is not global but rather applied normative in nature. Existential perspectives of artificial intelligence in the short or long term are not currently considered.

Conclusions and Implications

Thus, artificial intelligence has acquired the status of a global phenomenon, which forms a qualitatively new interpretation of this phenomenon in the scientific and philosophical dimension:

- artificial intelligence needs to be given a clear status in the current dimension of its use and forecasting its prospects in strategic planning for the development of sociocultural space.
- artificial intelligence is classified in a strong and weak format, which has formed two strategies for positioning AI in the scientific picture of the world with corresponding perspectives on the global impact on humans and society: on the one hand, artificial intelligence is not a competitor to human intelligence and complements it as information and digital technological support; on the other hand, artificial intelligence has the potential to be separated into an autonomous structure that can potentially have an unpredictable (both positive and negative) impact on the development of civilisation.
- the potential threat of AI is not in its practical and subject matter functionality, but in the distortion of its intended purpose, according to which AI turns from a tool into an independent (and, thanks to intelligence and information, self-organised) subject of the scientific and ideological paradigm. Such positioning will deprive humans of the ability to influence AI and make it impossible to physically destroy or stop its inhumane or destructive activities.
- scientific characteristics of information and digital technologies serve as a source for the study of the epistemological potential of artificial intelligence and determine the directions of further development of the general theory of intelligence.
- the introduction of artificial intelligence should be based on innovative methodological approaches characterised by flexibility rather than fundamentality since the scientific and technological consideration of AI leaves many issues of the functional nature of this phenomenon unexplored.
- artificial intelligence, according to the philosophical understanding of its status, should remain in the cluster of practice-oriented tools, while making it impossible to give it a global autonomous role (without human interference in its activities), since such an approach would pose existential threats to human existence and civilisation as a whole.
- philosophical understanding of the role and significance of artificial intelligence emphasises the need to correlate innovative information and digital technologies with traditional humanistic principles, which is the key to the harmonious development of society and promotes balanced but progressive civilizational progress.

- the achievements of the socio-cultural space acquired through artificial intelligence are rapidly increasing in quantity and require the formation of qualitative characteristics. Among the achievements of artificial intelligence, the number of controversial and debatable elements is significantly increasing, which will eventually reach a critical point in society and will require the resolution of these contradictions.
- the modern scientific and philosophical picture of the world needs new guidelines, which is associated with the introduction of the achievements of scientific and technological progress into global use. Interdisciplinarity and synergistic principle are innovative philosophical principles that will allow the formation of relevant philosophical systems in the cluster of philosophy of science (philosophy of mind, philosophy of consciousness, neuro-techno-philosophy, etc.).
- philosophy is not able to stop or deny the rapid scientific and technological progress, a key element of which is artificial intelligence, but the task of philosophy remains to maintain a balance between the dichotomous components of the scientific and worldview paradigm (traditionalism vs. innovation, fundamentality vs. flexibility, pragmatism vs. ethics), which preserves the supremacy of humanistic principles and ensures the existence and development of human civilisation. The task of philosophy is to form the value, scientific, cognitive, methodological (and in the future, existential) dimensions of artificial intelligence in a human-dimensional worldview paradigm.

References

Abbass, H. A. (2019). Social Integration of Artificial Intelligence: Functions, Automation Allocation Logic, and Human-Autonomy Trust. *Cognitive Computation*, 11, 159–171. https://doi.org/10.1007/s12559-018-9619-0

Abedin, B. (2022). Managing the tension between opposing effects of explainability of artificial intelligence: a contingency theory perspective. *Internet Research*, 32(2), 425–453. https://doi.org/10.1108/INTR-05-2020-0300

Abioye, S. O., Oyedele, L. O., Akanbi, L., Ajayi, A., Davila Delgado, J. M., Bilal, M., Akinade, O. O., & Ahmed, A. (2021). Artificial intelligence in the construction industry: A review of present status, opportunities and future challenges. *Journal of Building Engineering*, 44, 103299. https://doi.org/10.1016/j.jobe.2021.103299

Al-Rodhan, N. (2023). Transdisciplinarity, neuro-techno-philosophy, and the future of philosophy. *Metaphilosophy*, 54, 73–86. https://doi.org/10.1111/meta.12595

Auernhammer, J. (2020). Human-centered AI: The role of Human-centered Design Research in the development of AI. In Boess, S., Cheung, M., Cain, R. (Eds.), *Synergy - DRS International Conference 2020* (pp. 11–14). https://doi.org/10.21606/drs.2020.282

Behrens, K., & Sandis, C. (2020). Philosophy, In a Sense. *The Philosophers' Magazine*, 91, 11–12. https://doi.org/10.5840/tpm20209184

Berberich, N., Nishida, T. & Suzuki, S. (2020). Harmonizing Artificial Intelligence for Social Good. *Philosophy & Technology*, 33, 613–638. https://doi.org/10.1007/s13347-020-00421-8

Brendel, A. B., Mirbabaie, M., Lembcke, T-B., & Hofeditz, L. (2021). Ethical Management of Artificial Intelligence. *Sustainability*, 13(4), 1974. https://doi.org/10.3390/su13041974

Buckner, C. (2023). Deeply Rational Machines. What the History of Philosophy Can Teach Us about the Future of Artificial Intelligence. Moderate Empiricism and Machine Learning [Preprint], Retrieved from: http://philsci-archive.pitt.edu/id/eprint/21902

Chowdhary, K. R. (2020). *Fundamentals of Artificial Intelligence*. Springer. https://doi.org/10.1007/978-81-322-3972-7

Clune, J. (2019). AI-GAs: AI-generating algorithms, an alternate paradigm for producing general artificial intelligence. *Computer Science*. https://doi.org/10.48550/arXiv.1905.10985

Coeckelbergh, M. (2019). Artificial Intelligence: Some ethical issues and regulatory challenges. *Technology and Regulation*, 31–34. https://doi.org/10.26116/techreg.2019.003

Cope, B., Kalantzis, M., & Searsmith, D. (2021). Artificial intelligence for education: Knowledge and its assessment in AI-enabled learning ecologies. *Educational Philosophy and Theory*, 53(12), 1229–1245. https://doi.org/10.1080/00131857.2020.1728732

Dick, S. (2019). Artificial Intelligence. *Harvard Data Science Review*, 1(1). https://doi.org/10.1162/99608f92.92fe150c

Dreyfus, H. L. (1992). *What Computers Still Can'T Do: A Critique of Artificial Reason*. MIT Press. Retrieved from: https://books.google.com.ua/books?hl=uk&lr=&id=7vS2y-mQmpAC&oi=fnd&pg=PR9&ots=1W39SJszOC&sig=PsPP3E_EirHRAJAOunTbwl1z0ig&redir_esc=y#v=onepage&q&f=false

Dreyfus, H. L. (1997). Intuitive, Deliberative, and Calculative Models of Expert Performance. In Zsambok, C. E., Klein, G. (Eds.), *Naturalistic Decision Making, Psychology,* Press. https://doi.org/10.4324/9781315806129

Farrow, E. (2019). To augment human capacity – Artificial intelligence evolution through causal layered analysis. *Futures*, 108, 61–71. https://doi.org/10.1016/j.futures.2019.02.022

Ferrario, A., Loi, M., & Viganò, E. (2020). In AI We Trust Incrementally: a Multi-layer Model of Trust to Analyze Human-Artificial Intelligence Interactions. *Philosophy & Technology*, 33, 523–539 https://doi.org/10.1007/s13347-019-00378-3

Feng, T. (2019). Artificial Intelligence's Turn of Philosophy. *IOP Conference Series: Materials Science and Engineering*, 646, 012008. https://doi.org/10.1088/1757-899X/646/1/012008

Fjelland, R. (2020). Why general artificial intelligence will not be realized. *Humanities and Social Sciences Communications*, 7(10). https://doi.org/10.1057/s41599-020-0494-4

Floridi, L. (2019). What the Near Future of Artificial Intelligence Could Be. *Philosophy & Technology*, 32, 1–15. https://doi.org/10.1007/s13347-019-00345-y

Frank, M. R., Wang, D., Cebrian, M., & Rahwan, I. (2019). The evolution of citation graphs in artificial intelligence research. *Nature Machine Intelligence*, 1, 79–85. https://doi.org/10.1038/s42256-019-0024-5

Gabriel, I. (2020). Artificial Intelligence, Values, and Alignment. *Minds & Machines*, 30, 411–437. https://doi.org/10.1007/s11023-020-09539-2

Gellers, J. C. (2020). *Rights for Robots. Artificial Intelligence, Animal and Environmental Law*, Routledge. https://doi.org/10.4324/9780429288159

Hehl, W. (2021). Chance in Physics. *Computer Science and Philosophy*, Springer, Wiesbaden. https://doi.org/10.1007/978-3-658-35112-0

Hui, Y. (2021). On the Limit of Artificial Intelligence. *Philosophy Today.* https://doi.org/10.5840/philtoday202149392

Islam, M. R., Ahmed, M. U., Barua, S., & Begum, S. (2022). A Systematic Review of Explainable Artificial Intelligence in Terms of Different Application Domains and Tasks. *Applied Sciences*, 12(3), 1353. https://doi.org/10.3390/app12031353

Krkač, K. (2019). Corporate social irresponsibility: humans vs artificial intelligence. *Social Responsibility Journal*, 15(6), 786–802. https://doi.org/10.1108/SRJ-09-2018-0219

Laplane, L., Mantovani, P., Adolphs, R., Hasok Chang, M., Mantovani, A., McFall-Ngai, M., Rovelli, C., Sober, E., & Pradeu, T. (2019). Why science needs philosophy. *PNAS*, 116(10), 3948–3952. https://doi.org/10.1073/pnas.1900357116

Lara, F., & Deckers, J. (2020). Artificial Intelligence as a Socratic Assistant for Moral Enhancement. *Neuroethics*, 13, 275–287. https://doi.org/10.1007/s12152-019-09401-y

Lee, R. S. T. (2020). *Artificial Intelligence in Daily Life*. Springer, Singapore. https://doi.org/10.1007/978-981-15-7695-9

Lepri, B., Oliver, N., & Pentland, A. (2021). Ethical machines: The human-centric use of artificial intelligence. *iScience*, 24(3), 102249. https://doi.org/10.1016/j.isci.2021.102249

Lin, T.-A., & Chen, P.-H. C. (2022). Artificial Intelligence in a Structurally Unjust Society. *Feminist Philosophy Quarterly,* 8(3/4). Retrieved from https://ojs.lib.uwo.ca/index.php/fpq/article/view/14191

List, C. (2021). Group Agency and Artificial Intelligence. *Philosophy & Technology*, 34, 1213–1242. https://doi.org/10.1007/s13347-021-00454-7

Livingston, S., & Risse, M. (2019). The Future Impact of Artificial Intelligence on Humans and Human Rights. *Ethics & International Affairs*, 33(2), 141–158. https://doi.org/10.1017/S089267941900011X

Mohamed, S., Png, M. T., & Isaac, W. (2020). Decolonial AI: Decolonial Theory as Sociotechnical Foresight in Artificial Intelligence. *Philosophy & Technology*, 33, 659–684. https://doi.org/10.1007/s13347-020-00405-8

Muller, V. C. (2021). Ethics of Artificial Intelligence and Robotics. In The Stanford Encyclopedia of Philosophy Zalta, E. (Ed.). Retrieved from: https://plato.stanford.edu/archives/sum2021/entries/ethics-ai/

Muhlhof, R. (2019). Human-aided artificial intelligence: Or, how to run large computations in human brains? Toward a media sociology of machine learning. *New media & society*, 22(10), 1868–1884. https://doi.org/10.1177/1461444819885334

Páez, A. (2019). The Pragmatic Turn in Explainable Artificial Intelligence (XAI). *Minds & Machines,* 29, 441–459. https://doi.org/10.1007/s11023-019-09502-w

Peters, M. A., Jandrić, P., & Hayes, S. (2022). Biodigital Philosophy, Technological Convergence, and Postdigital Knowledge Ecologies. In Peters, M. A., Jandrić, P., Hayes, S. (Eds.), *Bioinformational Philosophy and Postdigital Knowledge Ecologies. Postdigital Science and Education*. Springer, Cham. https://doi.org/10.1007/978-3-030-95006-4_1

Samoili, S., Lopez, M., Gomez, E., De Prato, G., Martinez-Plumed, F., & Delipetrev, B. (2020). AI Watch. Defining Artificial Intelligence. Towards an operational definition and taxonomy of artificial intelligence. *Technical Report. Joint Research Centre*. Retrieved from: https://eprints.ugd.edu.mk/id/eprint/28047

Santoni de Sio, F., & Mecacci, G. (2021). Four Responsibility Gaps with Artificial Intelligence: Why they Matter and How to Address them. *Philosophy & Technology*, 34, 1057–1084. https://doi.org/10.1007/s13347-021-00450-x

Shank, D., Graves, C., Gott, A., Gamez, P., & Rodriguez, S. (2019). Feeling our way to machine minds: People's emotions when perceiving mind in artificial intelligence. *Computers in Human Behavior*, 98, 256–266. https://doi.org/10.1016/j.chb.2019.04.001

Shaw, J. (2019). Articial Intelligence and Ethics Ethics and the dawn of decision-making machines. *Harvard magazine*, 1(7). Retrieved from: https://www.harvardmagazine.com/2019/01/artificial-intelligence-limitations 8/11

Schneider, S. (2019). *Artificial You: AI and the Future of Your Mind*. Princeton University Press. https://doi.org/10.1515/9780691197777

Sloman, A. (2019). The Computer Revolution in Philosophy: Philosophy. *Science and Models of Mind.* Retrieved from: http://epapers.bham.ac.uk/3227/

Su, B., & Luvaanjalba, B. (2021). The Effect of Hubert Dreyfus's Epistemological Assumption on the Philosophy of Artificial Intelligence. In Nah, F. F., Siau, K. (Eds.), *HCI in Business, Government and Organizations. HCII 2021. Lecture Notes in Computer Science*, 12783, Springer, Cham. https://doi.org/10.1007/978-3-030-77750-0_42

Sutrop, M. (2019). Should we trust artificial intelligence? *Trames*, 4, 499–522. Retrieved from: https://www.ceeol.com/search/article-detail?id=856194

Zednik, C. (2021). Solving the Black Box Problem: A Normative Framework for Explainable Artificial Intelligence. *Philosophy & Technology*, 34, 265–288. https://doi.org/10.1007/s13347-019-00382-7

Zhou, J., & Razak, R. R. (2022). Research on the Philosophy of Science Oriented to Deep Learning under the Ethical Dilemma. *International Journal of Intelligent Systems and Applications in Engineering,* 10(2s), 65–69. Retrieved from https://www.ijisae.org/index.php/IJISAE/article/view/2363