Journal of Digital Economy Research. 2023. 1. P. 109–136
DOI: 10.24833/14511791-2023-1-109-136

Received: 25 September 2022,
Revised 31 October 2022,
Accepted: 20 December 2022

ARTIFICIAL INTELLIGENT SYSTEMS AND THE PROBLEM OF "NATURAL" TRUST

by E. Degteva, O. Kuksova

Abstract. The development of AI technologies has heightened the problem of humanitarian challenges at all levels of social regulations. Ethical issues and, in particular, the problem of trust have become relevant to the field of high technology, given the fact that AI performs increasingly significant managerial functions that previously could only be performed by humans. This issue is directly related to artificial intelligence systems, which have already been embodied in specific extensive projects. In this study, the authors analyze the concept of trust through the prism of technological development. For this purpose, the study presents an overview of historical and contemporary interpretations of the concept of trust and proves that this concept is relevant and necessary to control the risks that arise when integrating AI products into social life. The authors show that a rethinking of the concepts of ethics and morality in the new context is required. This is a necessary requirement for the creation of trusted AI and for the achievement of trust in human interaction with technology products. The authors conclude that it is necessary to build an interdisciplinary dialogue to integrate theory and practice from numerous fields. To do this, it is necessary to create a common knowledge base and a platform for communication between all stakeholders, but it is also important to create favorable conditions for sustainable and constructive interaction. Therefore, trust is a relevant concept that needs to be constructed in a multidimensional frame of reference that targets different stakeholders and also takes into account interaction between human and technology, in other words, at all levels and on all scales.

Keywords: AI, artificial intelligence, artificial intelligence system, ethics, moral, trust

For citation: Degteva E., Kuksova O.,(2023) Artificial Intelligent Systems and the Problem of "Natural" Trust. *Journal of Digital Economy Research,* vol. 1, no 1, pp. 109–136. (in English).
DOI: 10.24833/14511791-2023-1-109-136

Information about the authors:

Degteva Elena – Master of Economics. Deputy Director of the Department for the Development of the Social Sphere and the NPO Sector, Ministry of Economic Development of the Russian Federation.
degteva@list.ru
ORCID: 0000-0002-4137-6317

Kuksova Olga – Master of Economics. Deputy Head of Directorate, Directorate "Analytical Center for the Fuel and Energy Complex" FGBU "REA" of the Ministry of Energy of Russia.
129110, Moscow, st. Gilyarovskogo, d.39, building 1.
kuksovaki@gmail.com
ORCID: 0000-0001-8169-1517

Introduction

Recently, there has been a clear need to "fit" humanitarian values into digital technology products, in particular, within the framework of diverse projects based on AI, which are used in public administration, business, education and other areas. Numerous theorists, experts, practitioners, cultural and artistic figures are conducting a heated and lively discussion about how to identify possible risks and threats in time, who and how should develop and apply risk control tools and how to implement them into the management and public control system with in order to prepare and adapt society for even greater adoption of AI technologies. Today it is ethical issues that are being widely discussed and actively developed in the context of the development and application of AI in the political, economic and social spheres.

These ethical issues are inextricably linked with the actual problem of trust, which today requires attention at all levels: in international relations, in the economy, in the social sphere. The phenomenon of trust has become relevant in the field of high technologies. Moreover, the discussion about human interaction with technology is exacerbated in an environment where AI is beginning to perform more and more managerial functions that previously could only be done by Humans [20]. It deals not only with "tactical" projects, but also with technologies that are used for high-level management. Among such technologies, artificial societies or digital twins of society can be distinguished. Humanity is faced with the task of training neural networks to "understand" and "take into account" ethical values, or to clearly determine that such a result is impossible. This challenge places a huge responsibility not only on developers but to all stakeholders of AI products creation, introduction and usage.

At the same time, the stated problem has its roots in the distant past - both the question of trust as is, and the problem of trust with technology have a long history and complex multifaceted content.

Taking into account the fact that projects of artificial societies already exist, and their solutions are applied in practice to test management decisions, it is obvious that further development and use requires great attention, because such activities affect as society as well as an individual. We believe that trust is a key factor and requirement for testing the acceptability of using AI projects. This is especially relevant for significant projects that are aimed at solving political and social issues using AI and, in particular, artificial intelligence systems, and within which it is vitally important to ensure trust between all interested participants [19].

Literature review

Interaction of technological products with the anthroposphere have attracted the attention of thinkers since Antiquity. Renaissance and the New Age periods were no exception. At the beginning of the 20th century, leading thinkers were concerned with the humanitarian aspects of technological progress and the interaction between Human and technology [16]. We will not make a philosophical and historical essay, but will draw attention to the fact that the diverse aspects of innovative technologies have become the subject of philosophical reflection much earlier than the term "artificial intelligence" appeared. At the same time, the concept of AI is undoubtedly of revolutionary importance for culture, since it has combined the categories of Artificial and Natural into the most important category for Human - in the topic of Intellection. As is known, the term "artificial intelligence" was proposed in 1955 by D. McCarthy, and later M. Minsky defined it as a science whose goal is to implement intellectual human tasks with the help of a machine [14].

The AI problem has become fertile ground on which researchers began to test the possibilities of convergence of natural and artificial, biological and technological - what today seems to be the most promising areas for the development of science and practice.

Right at the time of the birth of the very concept of AI, there were formulated essentially contradictory approaches to its definition. In the 1950s and 1960s, two views stood out in the field of AI. One approach, proposed by D. McCarthy, used the approach of pure logic and the common mathematical basis. McCarthy believed that research in the field of AI would be fruitful when the very concept of intelligence (which still has a number of interpretations) is defined. In his opinion, the definition should form a coordinate system determined by epistemology and heuristics. That is, epistemology is how new knowledge is acquired in terms of its usefulness for solving problems, and heuristics is how these problems are solved: by what procedures and by what means [16].

M. Minsky proposed different approach, which later became known as connectionism (biologically inspired approach). He called for studying how the human brain and psyche work, and on this basis to implement knowledge in a computer model. This approach provides a basis for interpreting AI as something subjective, with individuality, consciousness, and even emotions. [7].

In general, today we can talk about the existence of 3 main approaches to the development of AI:

Artificial Intelligence (AI), or "Symbolical Artificial Intelligence", or "Classical Artificial Intelligence" is a top-down approach aimed at reproducing cognitive abilities without diving into the level of individual neurons.

Generalized artificial intelligence, or Strong artificial intelligence (AGI), is currently an assumed, hypothetical form of AI, in which the machine will receive equal human intelligence, and, accordingly, such AI will have self-awareness and freedom of choice

Biologically inspired artificial intelligence, or biologically inspired cognitive architectures, are projects that, based on the achievements of neuroscience, are aimed at creating artificial systems that reproduce the behavior and thinking of biological beings [12]. In particular, the development of the theory of artificial neural networks correlates and interacts with research on the functioning of the human brain [18, 23].

Thus, the first approach develops the approach of McCarty, and the second and third directions are based on the works of Minsky. We see that today there is no unified approach and no clear understanding of the directions of AI development and, accordingly, there is no unified approach to understanding this issue [26].However, the attempt at comprehension of this general topical topic within the framework of the philosophy of technology began in the early 20th century. One of the greatest thinkers of the last century, M. Heidegger, pointed to the threats posed by such a stage in the development of civilization, in which technology began to play a dominant role. The thinker noted that technology is not neutral in relation to Human, and that Human loses his subjectivity and ceases to determine the development of civilization. Technique itself begins to transform the natural world, and misunderstanding of the meaning of technology and its capabilities is fraught with a mistake, for which Human will pay with his enslavement [16, 22].

In our time, the horizon of interpretations and forecasts is very wide and multidirectional. It should be noted that fundamental thought has become the object of not only professional philosophers, but also of talented inventors, entrepreneurs and businessmen. In particular, the rationalist E. Yudkowsky believes in the possibility of creating a conscious AI, but according to his position first of all it is necessary to learn the secret of the intelligence [17].

Techno-optimist R. Kurzweil believes that AI will definitely be created. Provided that Human himself can preserve values and culture, AI will help to make the world a better place and ensure the prosperity of civilization [24].

Transhumanist N. Bostrom predicts that the creation of AI will lead to the next stage in the evolution of sapient life - the emergence of the Supersanity, which will get to control human existence [9]. A.V. Abramova researches the ethical aspects of the development of AI, studying and systematizing global trends in the development of ethics tools in the field of AI [1].

A.Yu. Alekseev leads the Artificial Personality Project, within which framework he explores the prospects of computer modeling of cultural, political, social, economic, moral and other aspects of public life and Human. The philosopher writes about the relevance of artificial life research as an important component of the development of nano- and biotechnologies in order to create the phenomenon of biological life using digital technologies. In his opinion, artificial life is a key component of artificial societies, which in turn serve as the basis for information and sociocultural technol-

ogies aimed at exploring the entire breadth of social life. This field opened the way for an artificial personality, to which researchers attribute not only consciousness and self-awareness, but also, in particular, freedom of choice and the ability for moral imputation - justifying or condemning one's own thoughts and actions [3].

The study of the features of human intelligence and AI is done by I.M. Dzyaloshinsky, who aims to prove the fundamental difference between Human and artificial intelligence [14]. It should be noted that in recent years not only scientists, but also publicists, journalists and writers have started to deal with this problematic. Numerous cultural figures reflect on the prospects of AI development and the risks and opportunities offered by the products of technologies based on AI. In particular, we can mention the books "Artificial You: Machine Intelligence and the Future of the Mind", written by S. Schneider, and "Our Final Invention: Artificial Intelligence and the End of the Human Era", written by J. Barrat [6, 32].

Recently, a number of publications have appeared devoted to the study of the potential dangers of the abuse of neural networks by technology owners and resistance to them, relying on trusted artificial intelligence. [5]..

It is the development of trusted artificial intelligence that represents the most promising concept of creating and understanding the idea of ethical AI today. However, the concept of "trusted" itself has already moved beyond theory and has begun to be used as a legal term in regulation. In particular, the concept of trusted AI is defined in the Ethical Guidelines for Trusted AI High Level Expert Group on Artificial Intelligence of the European Commission (Ethics guidelines for trustworthy AI, 2019) [29]. According to this document, trustworthy AI meets the following requirements, which at first glance relate more to human activities than to technology:

- Ethical (existing in accordance with ethical values and a code of ethics);
- Legal (not contrary to the law);
- Robust (reliable and sustainable).

In this context, it is necessary to reflect on the understanding of such a concept that from the first glance belongs mainly to the social field, as trust. This concept unites to a certain extent the theme presented and therefore some relevant approaches to the definition of this concept will be proposed in the following. It is trust that acts as a unifying concept and as a practical tool that will enable it to become a unifying link between human values and technological development [28].

This notion has been introduced into theory and social practice since the early stages of civilization, practically at the same time being introduced by Confucius in Ancient China and Pericles in Ancient Greece, describing relations based on trust as a necessary condition for prosperity of the state and society [27].

Subsequently, at different historical periods, depending on the ideology, development strategies and personality traits of the leaders and state ideologists, approaches to the definition and evaluation of the meaning of this concept have repeatedly changed in the most radical way.

Turning to the realities of the 20th century, we note that in the post-World War II period, the problems of trust and the opposite state - distrust in the context of the arms race and nuclear deterrence policy were among the key ones. The school of political realism, developed by H.Kissinger, T. Schelling and a number of other scholars and strategists, emerged in the US within the framework of this policy. This approach implies that the level of trust of political actors (in this case players on the international field) is low, and the stability (predictability) of partners' actions is determined by the balance of power and the system of checks and balances. Only such an approach can ensure predictability in the political game [33].

The neoliberal approach approaches the issue of complex interdependence as an opportunity through multiple channels and instruments to build trust between international partners and thereby ensure a high level of trust in foreign policy discourse [30].

S. Huntington noted that a special value of trust emerges in the periods of transformation and crises, and emphasizes the importance of institutional trust because it is the stable institutions that allow maintaining society and civilization in difficult times [34].

The postmodernist view essentially rejects trust, and J.-F. Lyotard highlights as a trend a global distrust of "metanarratives" that are supposed to justify reality. Z. Bauman points to the ephemerality, instability and unreliability of post-modernity, and argues that the center of gravity is the shift from state and international institutions to the global market with its consumerist values and the replacement of rational knowledge with shallow media exposure [25].

This last-mentioned idea leads us to the problem of technological development because it has enabled the globalization of the economy and culture, the development of mass media and mass culture, and subsequently the social and mobile web, all enabled by AI technologies and products. Thus, the breadth of the discussion, the scope of the problem and the horizon of opinions indicate that this issue is a complex, urgent and promising field of study.

Discussion

The technological basis of many of the categories that shape civilization today influences the transformation of values and the emergence of new ones, as well as the formation of development strategies.

Developing on the topic of artificial intelligent systems, we will try to define this phenomenon through its basic principles: such a system actively interacts with the outside world, perceiving and demonstrating (in accordance with the available properties and resources) a reaction to a directed impact. The degree of autonomy of such a system is the higher, the more independent the system is from information sources and control commands. The less the system depends on information sources and control commands, the higher the degree of its autonomy. Artificial intelligent systems are, in a certain sense, self-sufficient systems that can be entrusted with solving a certain set of applied problems [19].

In particular, the neural network created by Salesforce is testing and developing ideal tax system in a simulated environment [35]. So called AI Economist, it is based on reinforcement learning, which involves applying rewards and punishments to machine algorithms to maximize desired outcomes. By the same principle, for example, Google DeepMind AlphaGo and AlphaZero algorithms were created. The purpose of the experiment is to help governments around the world create more equitable system of taxation. The program is committed to creating a huge number of ecosystems with theoretical workers who trade currencies and build houses. Pay levels and skill sets vary, and AI determines optimal tax rates. The researchers note that this approach will reveal irrational behavior that economists often do not take into account in their models [2].

The relevance of such projects is increasing due to the fact that humanity is now progressing at an accelerated pace, and the unprecedented speed of change is causing global crises on a regular basis. That is why the notion of "black swans" has become popular, regularly emerging large-scale risks and threats that cannot be anticipated, detected or identified in advance because they originate from areas beyond knowledge and are largely (certainly, not entirely) driven by rapid technological development, whose achievements are being integrated into all areas of society by expanding the technosphere and blurring the lines between the humanitarian and technological domains [13].

In this context, classical tools of economic theory, strategic management and even anti-crisis management lose credibility and cease to bring results. To explore a world in which the technosphere has taken an important place, tools are needed that also include the capabilities of digital technology, in particular AI. In this sense, artificial intelligent systems are very promising, as this framework uses agent-based modelling. This method investigates the behavior of decentralized agents and how this behavior determines the behavior of the whole system. Use of agent-based modelling in artificial intelligent systems offers great opportunities not only to predict social developments under ordinary conditions, but also to predict possible crises and minimize the damage of such events. Of course, such an approach is a challenge, as it must take into account a large number of complex factors and be based on the results of multifaceted studies of real-world complex systems [8].

In this context, trust is an important requirement for implementation, as we are dealing with radical uncertainties. In public life, at the individual level, and indeed globally, uncertainty about the future is now coming to the fore, both on a practical, strategic scale and on a philosophical, literally metaphysical level. However, in this environment, AI-based dark projects may just be the solution, because classical tools cannot show effective results today. The pace of change, the emergence of new conditions, the acceleration of development requires all factors to be taken into account, and classical theories and methodologies are unable to account for important but unnoticed and uncalculated elements, such as irrationality and other features of human nature, and all the limitations and opportunities that follow from this.

Under conditions of radical uncertainty, trust is a requirement providing meaning and foundation for further development. Moreover, in a society where development is determined by self-created phenomena (phenomena that were not originally conceived but emerged on their own, in particular when the combined effect of the action of a set of individuals differs from that of individuals) and non-ergodic processes (for non-ergodic processes the probability observed in the past is not applicable to future processes), uncertainty becomes a major characteristic. At the same time, society increases uncertainty through its own incoherence, creativity, and it is under these conditions that artificial intelligent systems are required to function and, moreover, to provide confidence in its activities and in its results [8].Right today we can already predict some consequences of such intelligent systems performance, which have both positive and negative characteristics. On the one hand, these systems are aimed at ensuring security and well-being, on the other hand, the application of solutions of such systems in real life can affect natural values such as freedom, privacy and personal information of real people. And most importantly, it is necessary to make clear and set up who will play the dominant role in the new conditions and what balance will be in the management of social life between the State, technological corporations, society and the individual.

Also, along with the opportunities that intelligent systems provide, it is necessary to take into account the threats emanating from them that have social consequences.

As an example, there are problems in the field of profiling and aggregation of personal data. Social networks, entertainment services, which are intelligent systems, can accumulate large amounts of personal information about the life activities of each user, their personal preferences, views, health, family relationships and other information. In cases where the system has learned to recognize a user by their behavioral factors, identification can occur even when the user appears on affiliate services under a pseudonym or anonymously. Such data profiling may violate human rights and freedoms [11].

Aggregation of user data to improve cooperation with affiliate services is not in itself dangerous, but if sufficient personal information from various sources is accumulated, the system may become attractive to intruders for the purpose of creating a data breach. Aggregation services are therefore obliged to improve the security of all systems involved in a timely manner. The problem, however, is far from being solved globally at the moment. The issue is currently being addressed at the highest level. According to the UN Economic and Social Council, there is a need to clarify who owns the data, to ensure responsible use of data, privacy protection, accountability and fairness in social media, participatory platforms and online commerce portals. The main objective of these activities is to build trust and stability in the digital space, in particular in the field of AI [15].

Another example of the possible negative social consequences when using smart systems is the topic related to big data. Big data is supposed to be depersonalized, but the possibility of extracting personal data from depersonalized data streams (de-personalization) and deriving individual information from big data is a possible work in

progress. In this case, the individual is not protected against unauthorized use of personal information and, of course, the credibility of the intelligence systems is no longer guaranteed. The list of the social consequences could be continued since the presented challenges have no generally valid solutions at this time. That is why further we will talk about the ethical aspects of the functioning of such systems, about the extent to which their behavior can be conditioned by ethical paradigms, norms, and ideas [31].

It is necessary to find clear and generally accepted answers to the questions of what we should and can use intelligent systems for, what powers should be transferred to these systems, what risks this entails, and how they can be controlled. At the same time, we must understand that intelligent systems are already capable of influencing society today, and, noting the need to develop Code of Ethics for AI, we must understand that we are talking about systems that simulate planning, goal setting, choice and implementation of Human behavior [17]. It is necessary to timely analyze the processes taking place in public life, and on the basis of the analysis carried out, to introduce relevant changes in legislation and social norms in advance. This requires the integration of different areas of knowledge, methodologies and practices to make sense of these processes.

Integration is necessary both within the framework of scientific and research, as well as legislative and practical activities between the State, the academic community and the entire society, since this issue covers a much wider area compared to those that individual institutions, organizations, communities and researchers are able to work out.

Given the relevance and need to create and implement Code of Ethics that will become fully applicable to AI technologies, it should be noted that this is not enough[36]. Taking into account the fact that in recent years in the public space the concepts of ethics, morality (in general) and public morals have been often used as synonyms, and moreover, the scope of their application captures the areas of not only philosophy and culture, politics and law, but also technologies, economics and others, let us recall the difference between these terms. If, in the simplest sense, ethics defines and contrasts the concepts of "good" and "evil" in the philosophy of life and the worldview system, and morality regulates social relations through norms and principles [21]. Public morals are the real behavior and specific actions, as well as how this behavior is reflected in the actual experience of the subject, group and society. In other words, by public morals we mean objective assessments of practical actions. Therefore, the general narrative of ethics, according to our understanding, is at least an abstract discussion. It only introduces new simulacrum that threatens Human autonomy and stimulates the polarization of society and increases the degree of social tension. We believe that modern social trends form a single moral knot - the foundation on which the steps of moral progress should be built [15].

It is for this reason that there a lot of speculations around these concepts, and the price of this in relation to the use of artificial intelligent systems may be too high. The problem is that there is no single conceptual apparatus for basic definitions, there is no proper level of digital and humanitarian literacy, and there is often a bias or conflict of interest among interested participants. The task is to create new definitions for

relevant categories that will be multidimensional in nature, and will also allow us to consider phenomena in new conditions from different positions and within different assessment systems. We believe that it is urgent to move to a systematic solution of this problem. After all, the creation and implementation of AI systems in various spheres of life is becoming increasingly important every year, to the point that today in practice there are systems that function autonomously, which are so significant that they carry not only potential, but also real risk [37].

This proves that the study of the principles of ethical behavior of intelligent systems today has a multifaceted significance, both from a theoretical and practical point of view. Intelligent systems, as we noted, will be used more often and will affect society and the individual. Under these conditions, it is critical to ensure transparency and flexibility in creating a digital twin of society. It is undoubtedly important to ensure that the credibility of such projects is strengthened. In scientific and especially public discourse, one can find both fantastic and justified fears in connection with the risks of irresponsible use of AI technologies, up to the threat of loss of autonomy and freedom by Human [10]. At the same time, there is no generally accepted and clear plan for building trust in this area, however, the historical experience of introducing innovations shows that trust is usually born if technologies are useful and safe, and their regulation is transparent.

Speaking about the practical side of the ethical regulation of the artificial society model, it is necessary to solve the problem of formalizing ethics in a format accessible to AI technologies, as well as to implement ethical decision-making in the autonomous system itself based on AI restrictions in accordance with moral norms or by teaching AI to recognize ethical conflicts and accept their individual decisions [5]. This requires increased accountability in the development and use of these systems, as well as the greatest possible operational control in view of the serious consequences of making the wrong ethical choice.

The existing experience of systematizing the main ethical issues related to the creation and development of complex projects based on AI technologies allows us to highlight the problems of confidentiality and data protection, transparency, and, perhaps, as a common denominator, trust in predictive data.

Conclusion

The study allows us to conclude that today it is necessary to study technological and social development in a complex, and this global view will allow us to capture and find a solution to the problems caused by the development and spread of AI and products such as artificial intelligence systems in time.

The authors' research has analyzed the concept of trust through the prism of technological realities. The authors gave an overview of historical and current interpretations of the concept of trust and demonstrated that this concept is very promising and even essential as a tool to control the risks arising in the creation of digital technology products, which is particularly evident in the case of intelligent systems.

At the same time, humanitarian values today require a certain rethinking and new definitions that will be relevant to the new emerging reality. The authors show that the use of notions of ethics and morality without a qualitative revision of their meaning to meet modern requirements is untenable and is a formal intellectual exercise that will not really lead to the creation of trusted AI and to the formation of trust in human interaction with technology. The authors argue that this issue is a challenge because modern civilization in its current stage of development includes a wide range of stakeholders representing political, economic and social spheres, from the global community to the individual, and it is obvious that to build this multidirectional dialogue, a new language and new definitions of terms that define requirements and allow describing values should be formed. According to the authors, the contribution to the development of this issue lies in the fact that considerable research has been carried out into the interpretations of the relevant concepts, and the conclusions presented can be used for the next specific steps.

For a comprehensive understanding and management of ongoing processes, the authors suggest the following activities:

1. It is necessary to build an interdisciplinary dialogue to integrate theory and practice from numerous fields, taking into account the global level of the issues that we are exploring.
2. To do this, it is necessary to create a common knowledge base and a platform for communication between all stakeholders.
3. Under these conditions, sustainable and constructive interaction will become possible, in which new definitions of the concepts under discussion can be formulated and agreed upon in order to further integrate them in practice, both socially and scientifically.

Thus, the authors believe that in modern conditions trust is a category that needs and can be built in multiple dimensions - between the stakeholders, between human and technology, that is, at all levels and on all scales.

References

1. Abramova A.V. Ethics in the field of artificial intelligence - from discussion to scientific justification and practical application: analytical report / A.V. Abramova, A.G. Ignatiev, M.S. Panova; MGIMO MFA of Russia, Center for Artificial Intelligence; XIII Convention of the Russian Association for International Studies (RAMI) (Moscow, October 14–16, 2021). - M.: MGIMO-University, 2021. - 24 p.

2. AI Designed a Fairer Tax System // Hightech.fm URL: https://hightech.fm/2020/05/07/ai-economist (Accessed: 11/14/2022).

3. Alekseev A.Yu. Cognitive-technological projects of artificial personality // Man: Image and Essence. Humanitarian aspects. - 2014. - No. 1. - P. 156-174.

4. Anton, Eduard & Kus, Kevin & Teuteberg, Frank. (2021). Is Ethics Really Such a Big Deal? The Influence of Perceived Usefulness of AI-based Surveillance Technology on Ethical Decision-Making in Scenarios of Public Surveillance. 10.24251/HICSS.2021.261.

5. Avdoshin S.M., Pesotskaya E.Yu. Trusted artificial intelligence as a way of digital protection // Business Informatics. - 2022. - №2. - S. 62-73.

6. Barrat J. (2013). Our final invention: artificial intelligence and the end of the human era (First). Thomas Dunne Books.

7. Boldachev A.V. Essay: myths and dead ends of artificial intelligence // Philosophical problems of information technology and cyberspace. - No. 2 (18). - P. 27-41.

8. Bookstaber, Richard. (2017). The End of Theory: Financial Crises, the Failure of Economics, and the Sweep of Human Interaction. 10.1515/9781400884964.

9. Bostrom, N. Artificial Intelligence. Stages. Threats. Strategies. - St. Petersburg: Mann, Ivanov and Ferber, 2016. - 490 p.

10. Bulavinova M.P. Risks and threats of new technologies based on artificial intelligence. (review) // Social and humanitarian sciences. Domestic and foreign literature. Ser. 8, Science of Science: Abstract Journal. - 2018. - No. 2. - P. 23-41.

11. Comi, Antonio & Nuzzolo, Agostino. (2014). Individual pre-trip path choice modelling in transit advisor tools: theoretical and empirical evidences.

12. Dubrovsky, David. (2021). The Task of the Creation of Artificial General Intelligence and the Problem of Consciousness. Russian Journal of Philosophical Sciences. 64. 13-44. 10.30727/0235-1188-2021-64-1-13-44.

13. Dubrovsky, David. (2022). An Epistemological Analysis of the Social and Humanitarian Significance of Artificial Intelligence Innovations in Context of Artificial General Intelligence. Russian Journal of Philosophical Sciences. 65. 10-26. 10.30727/0235-1188-2022-65-1-10-26.

14. Dzyaloshinskiy I.M. Artificial Intelligence: Humanitarian Perspective // Bulletin of the Novosibirsk State University. Series: History, Philology. - 2022. - №6. – P. 20-29.

15. Ethics And 'Digital' - A Short Summary. Robot Doctor, Robot Teacher, Robot Policeman: Social Risks And Ethical Challenges For Industry: Policy Brief For Volume 2 Of The Report "Ethics And Digital: Ethical Challenges Of Digital Technology". - Moscow: RANEPA, 2020. - 45 c.

16. Finn V.: "Not all functions of natural intelligence can be formalized and automated" // Kommersant URL: https://www.kommersant.ru/doc/4198609?ysclid=lafntgvfcq626184900 (date of access: 11/14/2022).

17. Goertzel, Ben. (2015). Superintelligence: Fears, Promises and Potentials: Reflections on Bostrom's Superintelligence, Yudkowsky's From AI to Zombies, and Weaver and Veitas's "Open-Ended Intelligence". Journal of Ethics and Emerging Technologies. 25.55-87. 10.55613/jeet.v25i2.48.

18. Gurjar, Arunaben & Patel, Shitalben. (2022). Fundamental Categories of Artificial Neural Networks. 10.4018/978-1-6684-2408-7.ch001.

19. Gurov O. (2020). Ethical Interaction with Intellectual Systems. Artificial societies. vol. 15, no. 3 DOI: 10.18254/S207751800010905-4

20. Gurov O. (2021). Artificial Intelligent Systems and Social Problems Solving: Challenge of Trust. artificial societies. vol. 16, no. 1 DOI: 10.18254/S207751800014095-3

21. Gurov O.N. (2022) Metaverse—Flight From Dusk to Darkness? The Art and Science of Television, 18 (1), 11–46. https://doi.org/10.30628/1994-9529-2022-18.1-11-46

22. Heidegger M. Time and being: articles and speeches. - M.: Respublika, 1993. - 447 p.

23. Karpov V. E., Gotovtsev P. M., Roizenzon G. V. On the issue of ethics and systems of artificial intelligence // Philosophy and Society. - 2018. - No. 2 (87). - P. 84-105.

24. Kurzweil, R. The evolution of the mind. How expanding the capabilities of our mind will solve many of the world's problems. - M.: EKSMO, 2015. - 352 p.

25. McHale, Brian. (2013). Afterword: Reconstructing Postmodernism. Narrative. 21. 357-364. 10.1353/nar.2013.0020.

26. Muggleton, Stephen, and Nicholas Chater (eds), Human-Like Machine Intelligence (Oxford, 2021; online edn, Oxford Academic, 19 Aug. 2021), https://doi.org/10.1093/oso/9780198862536.001.0001, accessed 12 Dec. 2022.

27. Munira, Fayzulloeva. (2022). Ethical Ideas of the Ancient World. Annals of Bioethics & Clinical Applications. 5. 10.23880/abca-16000222.

28. Omrani, Nessrine & Rivieccio, Giorgia & Fiore, Ugo & Schiavone, Francesco & Garcia-Agreda, Sergio. (2022). To trust or not to trust? An assessment of trust in AI-based systems: Concerns, ethics and contexts. Technological Forecasting and Social Change. 181. 121763. 10.1016/j.techfore.2022.121763.

29. Palladino, Nicola. (2021). The role of epistemic communities in the «constitutionalization» of internet governance: The example of the European Commission High-Level Expert Group on Artificial Intelligence. Telecommunications Policy. 45. 102149. 10.1016/j.telpol.2021.102149.

30. Ralph, Jason. (2013). The liberal state in international society: Interpreting recent British foreign policy. International Relations. 28. 3-24. 10.1177/0047117813486822.

31. Rawat, Romil & Yadav, Rishika. (2021). Big Data: Big Data Analysis, Issues and Challenges and Technologies. IOP Conference Series: Materials Science and Engineering. 1022. 012014. 10.1088/1757-899X/1022/1/012014.

32. Schneider, S. (2019). Artificial You: AI and the Future of Your Mind. Princeton University Press. https://doi.org/10.2307/j.ctvfjd00r

33. Schultz, Robert A. "Political Realism and the Society of Societies." In Information Technology and the Ethics of Globalization: Transnational Issues and Implications. Hershey, PA: IGI Global, 2010. https://doi.org/10.4018/978-1-60566-922-9.ch006

34. Thanetsunthorn, Namporn. (2021). Organization development and cultural values of trust in international contexts. Review of International Business and Strategy. ahead-of-print. 10.1108/RIBS-04-2021-0054.

35. The AI Economist: Improving Equality and Productivity with AI-Driven Tax Policies // Salesforce URL: https://blog.salesforceairesearch.com/the-ai-economist/ (Accessed 11/14/2022).

36. Kopalle, Praveen & Gangwar, Manish & Kaplan, Andreas & Ramachandran, Divya & Reinartz, Werner & Rindfleisch, Aric. (2021). Examining Artificial Intelligence (AI) Technologies in Marketing Via a Global Lens: Current Trends and Future Research Opportunities. International Journal of Research in Marketing. 39. 10.1016/j.ijresmar.2021.11.002.

37. Vakkuri, Ville & Kemell, Kai-Kristian & Jantunen, Marianna & Abrahamsson, Pekka. (2020). "This is Just a Prototype": How Ethics Are Ignored in Software Startup-Like Environments. 10.1007/978-3-030-49392-9_13.