

UDC 101.1

[https://doi.org/10.30546/ifs.2025.2\(45\).093](https://doi.org/10.30546/ifs.2025.2(45).093)

**PHILOSOPHY OF ARTIFICIAL INTELLIGENCE AND THE COORDINATION
OF ITS CAPABILITIES WITH HUMAN VALUES IN
THE REPUBLIC OF AZERBAIJAN**

Kamina Mamedova

PhD in Philosophy, Associate Professor,
Azerbaijan State Oil and Industry University, **Azerbaijan**
email: kmamedova1962@mail.ru
<https://orcid.org/0000-0002-3859-8859>

Abstract. In the modern period, the issue of artificial intelligence (AI) is the most widely discussed and relevant topics. The mechanisms of human ontological states and their convergence with AI, the leader of knowledge and innovation, are the subject of active discussions among scientists and specialists. Some of them anticipate substantial results from the symbiosis of humans and AI, while others, appealing to the unpredictability of AI, speak of a possible threat it may pose to humanity on a global scale. In accordance with this, the article sets the goals of identifying the risk of uncontrolled development of AI and, using the example of initiatives in the Republic of Azerbaijan, examines the potential for harmonizing of AI with human values.

Keywords: artificial intelligence, philosophical meaning, human values, interaction, threat, Azerbaijan's initiatives.

**SÜNİ ZƏKA FƏLSƏFƏSİ VƏ AZƏRBAYCAN RESPUBLİKASINDA ONUN
QABİLİYYƏTLƏRİNİN İNSAN DƏYƏRLƏRİ İLƏ UZLAŞDIRILMASI**

Kamina Məmmədova

Xülasə. Müasir dövrdə süni intellekt (Sİ) problemi ən aktual və çox müzakirə olunan məsələdir. İnsan ontoloji vəziyyətlərinin mexanizmləri və onun bilik və innovasiyaların lideri olan Sİ ilə yaxınlaşması alim və mütəxəssislər arasında fəal müzakirələrin mövzudur. Onlardan bəziləri insan və süni intellektin simbiozundan böyük nəticələr gözləyir, digərləri isə süni intellektin gözlənilməzliyinə müraciət edərək, onun qlobal bəşəriyyət üçün ola biləcək təhlükəsindən danışıq. Müvafiq olaraq, məqalənin məqsədi Sİ-nin nəzarətsiz inkişafı riskini müəyyən etmək və Azərbaycan Respublikasının təşəbbüsündən istifadə edərək, Sİ-nin ümumbəşəri dəyərlərə mümkün uyğunluğunu müəyyən etməkdir.

Açar sözlər: süni intellekt, fəlsəfi məna, insani dəyərlər, qarşılıqlı əlaqə, təhlükə, Azərbaycanın təşəbbüsləri.

**ФИЛОСОФИЯ ИСККУСТВЕННОГО ИНТЕЛЛЕКТА И СОГЛАСОВАНИЕ ЕГО
СПОСОБНОСТЕЙ С ЦЕННОСТЯМИ В АЗЕРБАЙДЖАНСКОЙ РЕСПУБЛИКЕ**

Камина Мамедова

Резюме. В современный период проблема искусственного интеллекта (ИИ) является одной из наиболее обсуждаемых и актуальных тем научного и философского дискурса. Механизмы онтологических состояний человека и процессы конвергенции человека и ИИ – лидера современных знаний и инноваций - становятся предметом активных исследований среди учёных и специалистов. Одни из них рассматривают ИИ как перспективный инструмент развития, способный существенно повысить уровень знаний и качество жизни человечества, другие – как потенциальную угрозу его существованию. В статье анализируются философские

аспекты взаимодействия человека с ИИ, рассматриваются возможные последствия их симбиоза для глобального общества, а также оценивается вероятность согласования интеллектуальных возможностей ИИ с базовыми человеческими ценностями. Особое внимание уделено инициативам Азербайджанской Республики в сфере развития технологий ИИ и вопросам их соотношения с гуманистическими принципами.

Ключевые слова: искусственный интеллект, философское значение, сенсорная информатика, человеческие ценности, взаимодействие, угроза, инициативы Азербайджана

Introduction. The phenomenon of AI as a subject of philosophy

We are guided by a multitude of conceptual assumptions that we are not always aware of, but they shape our thinking and understanding of ourselves and the world around us. Throughout our lives, we are guided by a multitude of conceptual assumptions, some of which we are not always consciously aware, but which shape our thinking and understanding of ourselves and the world around us. Together, they form a logical structure, or architecture, underlying our consciousness and existence. What makes AI such a profound philosophical development is that it challenges many of the most fundamental, most taken-for-granted concepts or philosophies that have defined the modern era and by which most people still live. AI literally makes them untenable, marking a deep caesura - that is, creating a pause within processes of comprehension [7]. The subjectivity of AI is embedded in philosophical inquiry into its fundamental nature: consciousness and being, ethics, the nature of thought, and the implications of creating intelligent machines. The emergence of new interpretations of consciousness itself leads to new interpretations of AI itself. Many philosophical arguments aim to prove the impossibility of creating AI, a striking example being Daniel Dennett's arguments on intentionality - a way of interpreting objects or systems whereby we view them as possessing consciousness, desires, and beliefs so as to understand and predict their actions. Dennett views intentionality not as an intrinsic property of consciousness, but as an "intentional stance" - a practical approach to predicting the behavior of systems. In his view, intentionality is not inherent to nature or objects, but is ascribed to them by us in the process of interpreting their actions, resulting in us viewing them as thinking and willing entities. For Dennett, to understand why something seems "intentional" to us, we need not look for intrinsic properties, but rather to understand why we ourselves have adopted an intentional stance toward it [6; 212]. Australian philosopher David Chalmers emphasizes the importance of theories consciousness, arguing that they are essential for examining the social, ethical, and philosophical implications of integration superintelligent AI into society - scenarios involving augmented or uploaded human minds." In his book, "The Conscious Mind: A Fundamental Theory," Chalmers seeks to establish a "naturalistic dualism." He develops a thesis about the non-physical nature of consciousness and its dependence on functional circuits in the brain, arguing that his approach opens new perspectives for interpreting quantum mechanics and allows for the possibility of conscious robots [5; 106]. David Chalmers argues against the reduction of consciousness to purely physical processes and proposes a new, non-reductionist theory of consciousness [5; 120]. David Chalmers' book was a significant contribution to the philosophy of consciousness, addressing the "hard problem" of consciousness and offering a new framework for understanding the relationship between physical brain activity and subjective experience [5; 172]. J. Chalmers's book "The Conscious Mind: In Search of a Fundamental Theory" has influenced ongoing debates about the nature of consciousness.

The interdisciplinary field of AI encompasses ontological philosophy, computing and cognitive science [14].

Important philosophical debates about AI address questions such as: Can a computer truly be conscious, or can it merely simulate conscious experience? What are our responsibilities to intelligent machines? If AI systems are sentient or possess sensory information, do we have moral obligations to them? And would shutting down an AI be tantamount to murder? Key areas of research include ethics and responsibility: these include questions of AI safety, accountability for AI actions, the fairness of AI decision-making (for example, due to bias in training data), and how to align AI values with human values. Also, the epistemology of AI in relation to knowledge acquisition and discovery: can it truly understand the world around it, or is it only capable of limited information processing? [12]

The development of AI compels us to reconsider fundamental human concepts from new philosophical perspective. Within this framework, we examine the origins of AI philosophy, debated on consciousness and existence, the risks associated with generative AI, humans – AI interaction and the challenge of aligning of AI capabilities with human values in Azerbaijan.

The origins of artificial intelligence.

Today, we think of artificial intelligence (AI) as a technological product that enables computers and machines to imitate human learning, understanding, problem solving, decision making, creativity, and autonomy. In this form, AI emerged in the 1940s. Among the AI theorists, the British mathematician, logician, and cryptographer Alan Turing, Marvin Minsky, Claude Shannon, and Nobel laureates Herbert Simon and John Nash, Daniel Dennett, David Chalmers are commonly cited. Among Azerbaijani scientists are Lotfi Zadeh, Masud Mahmudov, Rasim Aliguliyev, İlham Mammadzadəh, Arif Bunyatov and a number of others. They explored the ability of machines to use language, form abstractions and concepts, and solve problems that are now reserved for humans. The Presidium of the National Academy of Sciences has established the "Electronic Academy" department, which is conducting important work in the field of robotics. The department's specialists developed the "Sufi Humanoid" robot (Universal Functional Word Processing).

The pioneer of AI is considered to be the British logician and pioneer of computing Alan Mathison Turing. In 1935, he described an abstract computing machine consisting of an infinite memory and a scanner that moves through the memory forward and backward, symbol by symbol, reading the symbols found and writing new ones. The scanner's actions are determined by a program of instructions, which are also stored in memory as symbols. This is Turing's stored-program concept, and it implies that the machine can execute its own program, modifying or improving it. Turing's concept is now known simply as a universal Turing machine. All modern computers are, in fact, universal Turing machines [13; 433]. In 1948, Turing presented his basic concepts of AI in a paper titled “Intelligent Machines.” Many of his ideas were later reinterpreted by other scientists. For example, one of Turing’s original ideas was to train a network of artificial neurons to perform specific tasks - an approach later described within the field of “connectionism.”

True artificial intelligence programs only appeared with the advent of electronic digital computers with stored programs.

The earliest successful AI program was written in 1951 by Christopher Strachey, later director of the Programming Research Group at Oxford University. Strachey's checkers-playing program ran on a Ferranti Mark I computer at the University of Manchester, England. By the summer of 1952, the program could play a full game of checkers at reasonable speed.

The earliest successful demonstration of machine learning was published in 1952. Shopper, a game written by Anthony Ettinger of the University of Cambridge, ran on the EDSAC computer. Shopper’s virtual world was a shopping mall with eight stores. When instructed to buy an item, Shopper searched for it by visiting stores at random until it found the right one. As it searched, Shopper memorized several products available at each store it visited (just as a human shopper would). The next time Shopper went to buy the same item, or another one it had already found, it immediately went to the right store. This simple form of learning is called rote learning.

The first AI program to run in the United States was also a checker’s program, written in 1952 by Arthur Samuel for an IBM 701 prototype. Samuel took Strachey’s checkers program as a starting point and expanded it significantly over the course of several years. In 1955, he added features that allowed the program to learn from experience. Samuel included mechanisms for both rote learning and generalization, improvements that eventually led to his program winning a game against a former Connecticut checkers champion in 1962. Samuel’s checkers program is also notable for being an early attempt at evolutionary computation. (His program “evolved” by comparing a modified copy to the best version of his program at the time, with the winner becoming the new standard.) [8]

Threats and security of generative artificial intelligence.

Thanks to the advancement and widespread use of interconnected devices, AI is gaining momentum in the gradual transformation of the socio-technical system [7]. There are now search engines, translation programs, facial recognition, driverless cars, airplanes...

All of the common applications of AI in modern society also coincide with the growing presence of AI in modern warfare, which is waged not only within physical territorial boundaries but also in the global information field. The need to adapt communication methods reflects a context in which new technologies and social transformations interact in complex ways, influencing both the form and content of public communication. These changes represent both a challenge and an opportunity to adopt more sophisticated and targeted approaches that can account for the new dynamics of participation and interaction in the modern digital and social landscape. The rise of AI in political processes creates previously unknown threats and risks, including:

1. Automated disinformation. The automated creation of disinformation poses a direct challenge to democratic processes, as algorithmically amplified lies systematically distort the political information environment, undermine public trust in institutions, and contribute to polarization – conditions that undermine democratic decision-making. These vulnerabilities are exacerbated by the asymmetry of information regulation between traditional media (historically under public control) and digital platforms.

2. Deepfakes are synthetic media. Deepfakes are synthetic media that replace the image of one person with another. Photo and film manipulation, phone decoders, and image enhancements have been around for a long time. However, the development of generative AI (GenAI) has taken this a step further by making the tools available at low or even free cost, allowing virtually anyone with a smartphone and some basic knowledge to create synthetic media using AI. In the hands of bad actors, these technologies can increasingly be used to create deepfakes for the purposes of fraud, deception, and defamation. Their use blurs the line between fiction and reality, causing the public to question what is real and what is not. We are currently seeing deepfakes used to advance political propaganda and financial fraud. The rise of deepfakes exploits the human tendency to trust visual and audio content, creating significant risks.

3. Advanced cyber-attacks. Advanced cyber-attacks are highly sophisticated malicious activities that go beyond simple scenarios and are aimed at compromising systems, stealing data, and disrupting systems. Advanced cyber-attacks pose a serious threat to individuals, organizations, and even national security, creating a dynamic and ever-changing challenge for cybersecurity professionals that requires constant vigilance and adaptation. These sophisticated attacks use advanced tools such as artificial intelligence (AI) to automate vulnerability detection and develop convincing schemes, as well as techniques such as advanced persistent threats (APTs), ransomware, social engineering, and supply chain attacks to remain undetected, steal data, and disrupt services.

4. Advanced, sophisticated, malicious software (malware). Advanced, sophisticated, malicious software (malware) that uses advanced techniques such as self-modification (polymorphism) and stealth to evade detection, allowing it to persist on a system and perform actions such as data theft or control. Advanced and sophisticated malware poses a significant cyber threat and risk because it is designed to operate stealthily, evade traditional defenses, and maintain long-term access to target systems. These attacks, often orchestrated through advanced persistent threats (APTs), use techniques such as obfuscation and fileless execution to evade detection, leading to risks of data theft, network disruption, and long-term damage to organizations.

5. Phishing and spear phishing attacks. Phishing is a type of social engineering attack in which cybercriminals imitate online conversations, websites, and other sensitive information to trick users into revealing sensitive information such as passwords, banking details, or other personal data. An advanced version of this attack is spear phishing. It is used with personalized information to appear more credible and deceive specific users or even entire organizations. Through phishing, attackers can manipulate public opinion and create a security threat.

Other threats include the exposure of sensitive data through the improper use of AI tools, potential over-reliance on fragile AI-based supply chains, and the concentration of democratically uncontrolled power in the hands of AI developers.

Accordingly, it is necessary to make significant efforts to ensure the security of digital systems. The most effective solutions to security challenges include:

1. Strong access control and authentication.
2. Improved privacy and data protection systems.

3. Robust AI model governance and monitoring.
4. Invest in AI ethics and security training.
5. Collaboration between cybersecurity experts and AI researchers.

Specific measures to protect digital systems include using multi-factor authentication, regularly updating security software (such as antivirus and firewalls), implementing data encryption and frequent backups, controlling access to sensitive information, monitoring systems for suspicious activity, and providing continuous cybersecurity training for employees.

Alignment of AI capabilities with human values in the Republic of Azerbaijan.

Azerbaijan is part of the global world and everything that happens in the world inevitably affects it. In the mid-20th century, when the republics of the former USSR were just beginning to implement AI technologies, Azerbaijan was one of the first to be involved in this problem. In the 1970s, the Azerbaijan Institute of Cybernetics (currently the Institute of Control Systems), as well as other scientific and educational institutions, were already developing projects and conducting research on AI. In those years, in addition to the technological aspects of AI, the Academy of Sciences of Azerbaijan conducted research on philosophical and physiological problems. After the 1980s, Professor Lotfi Zadeh's fuzzy logic theory had a great influence on research in the Republic of Azerbaijan. Thanks to him, fifth-generation computing systems with very large-scale integrated circuits (VLSI) were created using optoelectronic principles (lasers, holographs).

Based on Lotfi Zadeh's theory, the first robot with AI, Alter, was created in Japan in 2016 and was subsequently exhibited at the Tokyo International Science Museum. Thanks to Lotfi Zadeh's theory, scientific and theoretical research on decision-making under conditions of incomplete information, uncertainty, fuzzy decision-making, image and description recognition, machine learning and creation of expert systems was initiated in Azerbaijan. The work of modern Azerbaijani scientist-philosophers İlham Məmmədzadə and Arif Bünaytov “Logical-methodological foundations of the application of artificial intelligence systems in the context of fuzzy set theory” is devoted to the logical-methodological and philosophical-theoretical foundations of the concept of fuzzy logic əsasları” [10; 18-19]. The results of the above research have found application in the oil and gas, chemical industries and other areas. A group of Azerbaijani scientists was awarded the USSR State Prize for their scientific, theoretical and practical achievements in this area [4].

In the post-Soviet period, after the restoration of its independence, based on the understanding that AI is the future of the planet, priority attention was given to the development issues of AI technologies. By the Decree of the President of Azerbaijan İlham Aliyev dated March 19, 2025, the "Strategy for Artificial Intelligence of the Republic of Azerbaijan for 2025-2028" was adopted [1], [2].

The strategy included the following measures:

- identifying government agencies responsible for priority areas of AI development that determine its development;
- clarifying the risks associated with the protection of personal data during the design, implementation and use of AI systems;
- approving risk prevention rules.

The AI strategy reflected the fundamental issues of developing and improving AI governance, as well as issues related to the need to create an AI-based management system that would protect the interests and values of society, stimulating investment attraction. Currently, new technologies in the activities of various government agencies and individual infrastructure projects on AI are widely used in various areas of life: everyday life, consumer goods, such as virtual assistants - influencers, shoppers, shopping consultants, streaming experts, etc.; as well as in critical industries: healthcare, transport, finance and manufacturing. There is no denying the role of AI in government services, where it helps detect fraud and ensure public safety; in education through personalized learning and in scientific research in the field of environmental monitoring and precision medicine; in the military – in defense systems and, if necessary, in the offensive system. In accordance with the obvious role of AI as a virtual assistant, as well as with the “Strategy for the Development of AI in the Republic of Azerbaijan for 2025-2028”, the work on the implementation of its standardization is ongoing. The development, examination and approval of artificial intelligence standards is carried out by the Technical Committee

for Standardization “Information and Communication Technologies” (TK-05), which includes various relevant structures based on the Ministry of Digital Development and Transport.

As a result of the activities of the Technical Committee in recent years, measures have been taken to identify numerous international standards relevant in the field of information and communication technologies and implement them as national standards, 8 of which are in the field of AI. The management and members of the Technical Committee strive to adapt the national standardization system of Azerbaijan to international requirements by actively participating in various international conferences and seminars [11].

The Republic of Azerbaijan has made a significant step in the development of the digital economy and innovation: a new state standard was adopted, which is part of Azerbaijan's long-term strategy to strengthen its position in the field of high technology and digital transformation. The new standard defined key requirements for risk management, compliance with ethical principles and ensuring transparency in the development and application of AI technologies. A book by academician, Doctor of Physical and Mathematical Sciences Asaf Hajiyev, co-authored with Chinese professor Jiuping Xu, “High-regression analysis and Artificial Intelligence” (Springer Publishing House) was published on the introduction of AI in various fields, in compliance with norms and values. According to the book "High-regression analysis and Artificial Intelligence", artificial intelligence is developing at a very fast pace. But, we must be very careful here, because artificial intelligence, in addition to being useful, can cause harm if it is not handled properly. Therefore, standards are really needed in AI so that it does not harm, but brings benefit. Currently, in Azerbaijan, AI developers and users are moving towards the management of clear rules that are aimed at minimizing potential threats, protecting data privacy and ensuring algorithmic fairness. It is expected that targeted work in this area will contribute not only to the progress of science, but also to the significant development of industry, where intelligent technologies will find the widest application [11].

In the near future, it is planned to form a management system that supports and stimulates the activities of innovators, researchers and entrepreneurs, implementing public involvement in AI technologies; measures are being taken for state support of work on AI, and setting tasks for the relevant structures and supporting innovative projects. Azerbaijani scientists, who occupy leading positions in the Caucasus, were the initiators of the AI problem: Stanford University of the United States of America included them in the ranking of "The most influential scientists in the world", which highlights the contributions of the National Academy of Sciences of Azerbaijan, the Ministry of Science and Education, as well as higher education institutions.

In the modern period, Azerbaijani scientists are faced with the task of studying numerous problems related to AI, including the construction of “smart” infrastructures.

The Republic successfully implements programs for training personnel in AI engineering specialties: machine learning, data analysis, robotics, natural language processing, computer vision, as well as non-technical specialties: AI product managers, AI ethics specialists and operational information engineers. All specialties included in the Program play a key role in the ethical and effective development, management and deployment of AI systems.

The following initiatives are welcomed in the Republic of Azerbaijan:

- “Scientists of Tomorrow” competitions organized by the Ministry of Science and Education using the “STEAM” approach, which stipulates integrated learning, development of critical thinking, creativity and problem-solving skills related to AI;

- TEKNOFEST aviation, space and technology festivals.

Such events provide great support to increasing the interest of young people in the field of AI. In the near future, the Academy of Artificial Intelligence Technologies will be established in Azerbaijan, which will be produced in technology parks and industrial zones of Azerbaijan.

Conclusion

The nature of AI is dualistic: on the one hand, there are advantages, on the other, there are disadvantages. The advantages of AI include efficiency achieved through automation of tasks, data analysis for decision-making, assistance in various areas of life, including progress in the field of self-

driving cars. The disadvantages of AI include the loss of jobs, ethical issues related to bias and privacy, security risks associated with hacking, and the lack of human creativity and empathy. Since people can be manipulated, AI without ethical constraints can be extremely dangerous. Therefore, the priority is to create a higher entity to ensure global security and prevent harm to humanity. Absolute security is impossible. It is quite possible to imagine that a properly designed AI could even help eliminate the shortcomings of human activity. Everything else is self-destruction.

However, it is impossible to erase AI from life; it is like trying to stop a speeding train. Therefore, we must learn to reconcile the capabilities of AI with human values. In this regard, the Republic of Azerbaijan demonstrates an example of optimal interaction with AI.

Reference

1. Azərbaycan Respublikası Prezidentinin Sərəncamı. Azərbaycan Respublikasının 2025–2028-ci illər üçün süni intellekt Strategiyası”nın təsdiq edilməsi haqqında/ Bakı şəhəri, 19 mart 2025-ci il № 530/ <https://e-qanun.az/framework/59218>
2. Azərbaycan Respublikasının 2025–2028-ci illər üçün süni intellekt/ <https://e-qanun.az/framework/59218/2025.07.30>
3. Azərbaycan süni intellektdən istifadə üzrə milli standartların qəbuluna başlayıb/ <https://mincom.gov.az/ru/media-ru/novosti/v-azerbaidzane-nacalos-priniatie-nacionalnyx-standartov-po-primeneniiu-iskusstvennogo-intellekta>
4. Bu gün müasir “ağıllı” qurğular Lütfi Zadənin nəzəriyyəsinə əsaslanır. Müsahibə, çıxışlar Akademik Rasim Əliquliyev/ <https://ict.az/news/2721/03-oktyabr-2017>
5. David J. Chalmers. The conscious Mind in search of a fundamental theory. New York: Oxford University Press, 1996, 414 p.
6. Dennet D. The Intentional Stance. XI. Cambridge: MIT Press, 1987, 388 p.
7. Diakopoulos, Nicholas. 2019. Automating the news: How algorithms are rewriting the media. USA: Harvard University Press
8. Copeland B.J. History of artificial intelligence (AI)/ <https://www.britannica.com/science/history-of-artificial-intelligence/> Last Updated: Jul 30, 2025
9. Hauser L. Internet Encyclopedia of Philosophy IEP. Artificial Intelligence/ Alma College U.S.A/ <https://iep.utm.edu/artificial-intelligence/>
10. Məmmədzadə İ., Buniyatov A. Qeyri-səlis məntiq konsepsiyasının elm fəlsəfəsi kontekstində məntiqi-metodoloji əsasları. Bakı: Bilgi MMC, 2021, 136 c.
11. Rzayeva S. CBC TV Azerbaijan/ <https://www.youtube.com/watch?v=oRpsDu-Vy6o>/ Jul 28, 2025
12. The Cambridge Handbook of the Law, Ethics and Policy of Artificial Intelligence. Part I - AI, Ethics and Philosophy Publisher: Cambridge University Press: 06 February 2025, pp.15-130/ Ed. Smukha N.A./ <https://www.cambridge.org/core/books/cambridge-handbook-of-the-law-ethics-and-policy-of-artificial-intelligence/ai-ethics-and-philosophy/EFEFAE5CD989984FD6C3FA64BF B25658/>
13. Turing A.M. Computing Machinery and Intelligence. Mind 49, 1960, p. 433-460.
14. Tobias Rees. Why AI Is A Philosophical Rupture. The symbiosis of humans and technology portends a new AIXIAL Age/ <https://www.noemamag.com/why-ai-is-a-philosophical-rupture/> February 4, 2025
15. Xalq qəzeti. Akademik Rasim Əliquliyevin Xalq TV kanalına müsahibəsi/<https://science.gov.az/az/news/open/26538/10/10/2023>

Rəyçi: dos. Liliya Mirzəzadə