SOCIO-PHILOSOPHICAL ANALYSIS OF THE INTERRELATIONSHIP BETWEEN ARTIFICIAL INTELLIGENCE AND HUMAN VALUES

Soliyev Jasurbek Isroilovich
History Teacher at School No. 70, Namangan District
E-mail: jasur1986@gmail.com

Abstract: This article examines the interrelationship between artificial intelligence and human values from a socio-philosophical perspective. It argues that artificial intelligence should be understood not merely as a technological instrument, but as a socio-cultural phenomenon capable of influencing human autonomy, dignity, justice, responsibility, trust, equality, and social solidarity. The development and widespread application of AI systems are transforming traditional mechanisms of decision-making, communication, knowledge production, professional activity, and social interaction. In this context, the problem of preserving the priority of human values acquires particular philosophical significance.

Keywords: artificial intelligence, human values, social philosophy, axiology, human dignity, autonomy, justice, responsibility.

Introduction. The rapid development of artificial intelligence has become one of the most significant factors transforming contemporary society. Artificial intelligence systems are increasingly used in education, healthcare, finance, public administration, scientific research, communications, security, industry, and other spheres of social life. As a result, AI is no longer merely a technical tool designed to automate individual operations. It is gradually becoming an important component of the social environment in which people make decisions, acquire knowledge, communicate, work, and construct their understanding of reality.

The expansion of artificial intelligence raises not only technological and economic questions but also fundamental philosophical problems. The central issue is no longer limited to what artificial intelligence can technically accomplish. The more significant question concerns what AI systems should be permitted to do, which human purposes they should serve, what values should guide their development, and who should remain responsible for the consequences of algorithmically mediated decisions.

From an axiological perspective, values constitute fundamental orientations that determine what individuals and societies consider desirable, meaningful, legitimate, and worthy of protection. Values influence human

choices, social institutions, moral norms, and collective development. They are not merely abstract ideals but function as important motivational and normative principles of social action [1]. Consequently, any technology that significantly influences human choices and social relations inevitably enters into interaction with existing systems of values.

Artificial intelligence intensifies this interaction because algorithms increasingly participate in processes that traditionally required human judgment. AI systems may recommend medical treatments, rank job applicants, assess financial risks, select educational content, moderate information, generate texts and images, and support administrative decisions. Such practices inevitably affect freedom, equality, privacy, dignity, fairness, responsibility, and trust.

The philosophical problem therefore lies in determining whether artificial intelligence remains an instrument controlled by human values or whether algorithmic systems gradually begin to restructure the very conditions under which those values are interpreted and realized. The relationship between AI and human values is thus dialectical: humans design technologies in accordance with particular social priorities, while technologies subsequently influence human behavior, institutions, expectations, and value orientations.

Main Part. Artificial intelligence should first be considered as a socio-technical phenomenon. Although its functioning is based on mathematical models, algorithms, computational infrastructures, and large amounts of data, the creation and application of AI take place within specific social, political, economic, and cultural contexts. Data are selected by people, objectives are defined by organizations, models are developed according to particular institutional priorities, and technological outputs are interpreted within existing social relations. Therefore, artificial intelligence cannot be regarded as completely independent of human values.

Contemporary research on AI ethics emphasizes that technological innovation has both opportunities and risks for human dignity and social development. Artificial intelligence may expand individual capabilities, increase access to knowledge, support scientific discovery, improve healthcare, reduce routine labor, and make certain public services more efficient. At the same time, AI may intensify inequality, reproduce discrimination, facilitate manipulation, increase surveillance, weaken privacy, and reduce meaningful human control over important decisions [2].

This dual character demonstrates that technology itself does not automatically determine social progress. The social consequences of AI depend on the goals, institutions, values, and power relations within which it is developed and applied. Thus, technological capability and social value cannot be considered identical concepts. An AI system may be technically effective

while remaining socially undesirable or ethically problematic.

The relationship between artificial intelligence and human values can be analyzed at several interconnected levels. The first is the **axiological level**, which concerns the values that should guide technological development. The second is the **institutional level**, which concerns the organizations and rules responsible for implementing these values. The third is the **behavioral level**, which relates to the way AI influences human choices and patterns of conduct. The fourth is the **anthropological level**, which raises questions concerning human identity, agency, responsibility, and the meaning of being human in an increasingly automated society.

Among these dimensions, the axiological dimension is fundamental because it establishes the criteria by which technological progress should be evaluated. The question "Can an AI system perform this function?" is technological. The question "Should an AI system perform this function?" is axiological and philosophical.

One of the central values affected by artificial intelligence is **human dignity**. Human dignity implies that the individual possesses intrinsic value and must never be reduced merely to an object of calculation, classification, manipulation, or economic optimization. In the context of AI, this principle becomes particularly relevant because algorithmic systems operate by converting human characteristics and behavior into data.

Data-driven technologies can classify individuals according to predicted preferences, risks, productivity, creditworthiness, educational performance, or other measurable characteristics. Such classifications may be useful for particular purposes, but they create a philosophical danger when the complexity of human personality is reduced to a numerical profile. A human being cannot be completely identified with data generated about him or her.

For this reason, international approaches to AI governance increasingly place human dignity at the center of technological development. UNESCO's Recommendation on the Ethics of Artificial Intelligence identifies respect for human rights and human dignity as a foundational value and links it to justice, inclusiveness, diversity, environmental well-being, transparency, and human oversight [7]. This approach reflects an important philosophical principle: technology must be assessed according to its contribution to human flourishing rather than requiring human beings to adapt uncritically to technological logic.

Another fundamental value is **human autonomy**. Autonomy implies the capacity of individuals to formulate goals, make informed choices, exercise judgment, and take responsibility for their actions. Artificial intelligence can strengthen autonomy by providing people with better information and additional opportunities. However, it can also weaken autonomy when

algorithmic systems begin to determine choices without sufficient awareness or control on the part of the individual.

Recommendation systems provide a clear example. Digital platforms increasingly determine which news, entertainment, products, political messages, or educational materials users are likely to encounter. Formally, individuals retain the right to choose. Yet the environment in which their choices are made is already structured by predictive algorithms. This creates a distinction between formal freedom and substantive autonomy.

From a socio-philosophical perspective, autonomy requires more than the existence of multiple options. It requires the ability to understand the conditions under which choices are offered and to exercise reflective judgment. If individuals continuously delegate decisions to automated systems, there is a risk that decision-making skills, critical thinking, and moral responsibility may gradually weaken.

Technology ethics therefore cannot focus exclusively on preventing direct physical or economic harm. It must also consider how technologies shape moral character, habits, attention, and human capabilities. The virtue-ethical approach to technology emphasizes that emerging technologies transform not only external conditions but also the qualities required for individuals to live well in technologically mediated societies [3]. Values such as honesty, justice, courage, empathy, care, humility, self-control, and practical wisdom acquire new forms in digital environments.

This idea is especially important in the age of generative artificial intelligence. When AI systems can produce texts, images, analyses, and recommendations, the question arises as to which activities should remain meaningful expressions of human judgment and creativity. The philosophical issue is not that machines perform intellectual operations but whether reliance on machines changes the way individuals understand knowledge, authorship, effort, originality, and responsibility.

Artificial intelligence may support human cognition, but assistance should not be confused with substitution of human agency. A human-centered model of AI requires maintaining a meaningful role for human deliberation, especially in decisions involving moral, legal, medical, educational, or political consequences.

The value of **justice** constitutes another central dimension of the relationship between AI and society. Algorithmic systems frequently rely on historical data. Yet historical data reflect existing social structures and may contain patterns of inequality or discrimination. If such data are incorporated into AI systems without appropriate critical evaluation, algorithms may reproduce existing injustices while giving them the appearance of technological

neutrality.

This problem demonstrates why the idea of a completely value-neutral algorithm is philosophically problematic. Every AI system involves choices concerning what should be measured, what data should be included, how outcomes should be evaluated, and which errors should be considered acceptable. Such choices inevitably have normative implications.

A system optimized primarily for efficiency may produce outcomes different from one designed to prioritize equality or protection of vulnerable groups. Consequently, technological design itself becomes a form of practical value selection.

Global analyses of AI ethics guidelines demonstrate broad convergence around values such as transparency, justice, fairness, non-maleficence, responsibility, privacy, and human autonomy [5]. However, agreement at the level of general principles does not automatically guarantee ethical outcomes in practice.

The difference between declaring values and institutionalizing them constitutes an important philosophical and governance problem. Ethical principles can remain symbolic unless they are translated into concrete technical standards, organizational responsibilities, evaluation procedures, and mechanisms of accountability. It has therefore been argued that principles alone cannot guarantee ethical artificial intelligence [6].

This distinction is highly relevant to the philosophy of values. A value becomes socially effective only when it is embodied in practices and institutions. Justice, for example, cannot remain merely an abstract declaration; it must influence how datasets are constructed, how algorithms are tested, how affected individuals can challenge decisions, and how responsibility is allocated when harm occurs.

Thus, the relationship between AI and values can be represented through the following mechanism:

human value → ethical principle → technological requirement → institutional mechanism → practical application → social consequence → reassessment of the value.

This model also demonstrates that the process is not linear. Social consequences generated by technology may cause societies to reinterpret existing values. For example, the widespread use of digital technologies has expanded traditional understandings of privacy. Privacy no longer concerns only protection from physical intrusion; it also includes control over personal data, profiling, automated inference, and behavioral prediction.

The value of **responsibility** becomes particularly complex in AI-mediated environments. Traditional moral responsibility is usually associated with an

identifiable human agent capable of understanding actions and their consequences. Artificial intelligence introduces distributed forms of agency involving developers, data providers, organizations, users, managers, and automated systems.

When an algorithmic decision causes harm, identifying responsibility may therefore become difficult. Was the problem caused by the programmer, the organization that deployed the system, biased training data, inappropriate use, insufficient human supervision, or an unpredictable interaction between these factors?

This situation creates what may be described as a potential "responsibility gap." Nevertheless, attributing autonomous moral responsibility to AI itself does not necessarily solve the problem. AI systems do not possess moral agency in the same sense as human beings simply because they generate complex outputs. Moral responsibility remains embedded in social institutions and human decisions concerning design, deployment, oversight, and use.

AI ethics therefore requires the preservation of meaningful human accountability. The principle of explicability is particularly significant in this context. Ethical frameworks for a "Good AI Society" have emphasized beneficence, non-maleficence, autonomy, justice, and explicability, where explicability incorporates both intelligibility and accountability [2]. If an algorithm affects an individual's rights or opportunities, society must be able to ask not only how the decision was produced, but also who is responsible for its consequences.

Transparency and trust are closely related to responsibility. Trust is a fundamental social value that makes cooperation possible. Individuals rely on medical institutions, educational institutions, financial organizations, government agencies, and information systems partly because they expect those institutions to operate according to legitimate norms.

Artificial intelligence changes the structure of trust. In traditional interpersonal interaction, trust may be based on personal experience or institutional authority. In algorithmic environments, individuals often interact with systems whose internal operations they cannot directly understand.

Absolute technical transparency is not always possible or meaningful. Modern AI models may contain extremely complex internal processes. Therefore, transparency should not be reduced to exposing every technical detail. From a social-philosophical standpoint, transparency means providing sufficient information for affected individuals and institutions to understand the purpose, limitations, risks, and consequences of an AI system.

This is why contemporary frameworks increasingly connect transparency with explainability, accountability, and human oversight. The OECD AI

Principles, adopted in 2019 and updated in 2024, emphasize human rights, democratic and human-centered values, including dignity, autonomy, privacy, equality, fairness, social justice, transparency, and mechanisms for human agency and oversight [8].

The NIST Artificial Intelligence Risk Management Framework similarly treats trustworthy AI as a socio-technical issue and identifies reliability, safety, security, accountability, transparency, explainability, privacy protection, and fairness as interconnected characteristics [9]. This indicates that trust cannot be created by technical accuracy alone. A highly accurate system can still be socially unacceptable if it violates privacy, systematically discriminates against particular groups, or prevents individuals from challenging decisions.

The relationship between artificial intelligence and **equality** is equally complex. AI can increase access to education, healthcare, public information, and professional opportunities. At the same time, unequal access to advanced technologies may deepen existing social divisions.

A new form of inequality may arise between individuals and communities that possess the skills, infrastructure, computing resources, data, and economic capacity to benefit from AI and those that do not. Consequently, the problem of digital inequality is transforming into a broader problem of algorithmic inequality.

This process raises questions of distributive justice. Who receives the benefits created by artificial intelligence? Who bears its risks? Which groups participate in defining technological priorities? Who owns the data and computational infrastructure that make AI possible?

These questions demonstrate that AI ethics cannot be separated from political economy and social philosophy. The development of artificial intelligence takes place within systems of ownership, labor relations, institutional authority, and global inequality. Therefore, ethical analysis must consider not only individual interactions with AI but also the distribution of technological power within society.

Artificial intelligence may also transform the meaning of **work**. Automation can free individuals from repetitive and dangerous forms of labor, potentially creating opportunities for more creative and meaningful activity. At the same time, it may reduce demand for certain professions, restructure professional identities, and alter the relationship between human skills and economic value.

The philosophical importance of work extends beyond income. Work can provide social recognition, identity, participation, responsibility, and a sense of contribution to society. Consequently, discussions about AI-driven automation should not be limited to productivity indicators. They must also address human

flourishing and the social meaning of productive activity.

The ethical challenges of artificial intelligence have increasingly become the subject of institutional regulation. The Council of Europe Framework Convention on Artificial Intelligence, opened for signature in September 2024, seeks to ensure that activities throughout the lifecycle of AI systems remain consistent with human rights, democracy, and the rule of law [10]. The European Union's Artificial Intelligence Act similarly establishes a human-centered approach and explicitly links trustworthy AI with fundamental rights, democracy, the rule of law, safety, and human well-being [11].

These developments demonstrate a broader transition from a technology-centered paradigm toward a human-centered paradigm. Under the technology-centered approach, the primary question is how rapidly and effectively AI can be developed. Under the human-centered approach, technological progress is evaluated in relation to its social purpose and its compatibility with human values.

A similar issue is becoming increasingly relevant in Uzbekistan. The Strategy for the Development of Artificial Intelligence Technologies until 2030, approved by Presidential Resolution No. PQ–358 of October 14, 2024, defines objectives related to expanding AI-based products and services, developing research laboratories and computing infrastructure, strengthening the regulatory framework, increasing international cooperation, and improving the population's knowledge and skills in the field of artificial intelligence [12].

From a socio-philosophical perspective, implementation of such strategies requires the technological dimension to be supplemented by an axiological dimension. Training specialists capable of creating AI systems is important, but it is equally necessary to develop a culture of responsible use of AI. Technological competence should be accompanied by ethical competence.

This means that education related to artificial intelligence should not be limited to programming, data analysis, and technical skills. It should also include questions of human dignity, privacy, fairness, responsibility, intellectual integrity, critical thinking, and the social consequences of technology.

The need for such an approach becomes especially evident in education. Generative AI can support teachers and students by providing explanations, generating exercises, translating materials, and facilitating access to information. However, excessive dependence on automated generation may weaken independent thinking, academic integrity, and personal intellectual effort.

The central philosophical task is therefore not to prohibit technological assistance but to establish a meaningful balance between technological support and human self-development. AI should extend human capabilities rather than

replace the processes through which human beings develop judgment, competence, creativity, and moral responsibility.

M. Coeckelbergh argues that AI ethics should move beyond abstract fears about intelligent machines and focus on concrete ethical questions concerning responsibility, privacy, bias, transparency, human relationships, and social power [4]. This orientation is methodologically important because it shifts philosophical analysis from speculative opposition between "human" and "machine" toward examination of the actual social relations created through AI.

Conclusion. The socio-philosophical analysis of the interrelationship between artificial intelligence and human values demonstrates that AI development is not a value-neutral technological process. Artificial intelligence emerges within society, reflects existing social priorities and contradictions, and subsequently influences the conditions under which individuals exercise freedom, make decisions, communicate, work, and participate in social institutions.

REFERENCES

1. Schwartz S.H. An Overview of the Schwartz Theory of Basic Values // Online Readings in Psychology and Culture. – 2012. – Vol. 2. – No. 1. DOI: 10.9707/2307-0919.1116.
2. Floridi L., Cowls J., Beltrametti M., Chatila R., Chazerand P., Dignum V., Luetge C., Madelin R., Pagallo U., Rossi F., Schafer B., Valcke P., Vayena E. AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations // Minds and Machines. – 2018. – Vol. 28. – P. 689–707. DOI: 10.1007/s11023-018-9482-5.
3. Vallor S. Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting. – New York: Oxford University Press, 2016.
4. Coeckelbergh M. AI Ethics. – Cambridge, MA: The MIT Press, 2020. – 248 p.
5. Jobin A., Ienca M., Vayena E. The Global Landscape of AI Ethics Guidelines // Nature Machine Intelligence. – 2019. – Vol. 1. – P. 389–399. DOI: 10.1038/s42256-019-0088-2.
6. Mittelstadt B. Principles Alone Cannot Guarantee Ethical AI // Nature Machine Intelligence. – 2019. – Vol. 1. – P. 501–507. DOI: 10.1038/s42256-019-0114-4.
7. UNESCO. Recommendation on the Ethics of Artificial Intelligence. – Paris: United Nations Educational, Scientific and Cultural Organization, 2021.
8. OECD. Recommendation of the Council on Artificial Intelligence. OECD/LEGAL/0449. – Adopted 22 May 2019; amended 3 May 2024.

9. Tabassi E. Artificial Intelligence Risk Management Framework (AI RMF 1.0). – Gaithersburg, MD: National Institute of Standards and Technology, 2023. – NIST AI 100-1. DOI: 10.6028/NIST.AI.100-1.
10. Council of Europe. Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law. – Strasbourg: Council of Europe, 2024.
11. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) // Official Journal of the European Union. – 12 July 2024.
12. Resolution No. PQ–358 of the President of the Republic of Uzbekistan dated October 14, 2024, “On Approval of the Strategy for the Development of Artificial Intelligence Technologies until 2030.”