

Shiraliyev I.A.\*

DOI: 10.25108/2304-1730-1749.iolr.2025.80.142-149

UDC: 349:681

### Who is responsible for artificial intelligence?

**Abstract:** AI is a product of human activity, generally created with good intentions, though at times for private purposes far removed from ethical considerations. Responsibility for the quality of a product, even if it is not yet a commodity, lies with its creator. Yet in the case of AI, there are nuances.

It seems reasonable that responsibility for AI should be determined by analogy with the birth and upbringing of a human being – a view partially confirmed by existing normative acts.

From our perspective, the solution at this stage lies in developing a “moral code of AI” and mandating its incorporation into every new system, without which functioning would be impossible.

**Keywords:** intelligence, artificial intelligence, ethics, morality, principles, responsibility, accountability.

Information ethics, ethics of information and other technologies, computer ethics, technoethics, roboethics, AI ethics, and similar concepts represent the applied use of knowledge about morality, its value criteria, and moral principles of human behavior (ethics in the sense of its subject matter) for solving specific practical problems and for creating new products of human activity.

This statement is based on the following considerations.

Ethics (from the Greek ἠθική, derived from the Ancient Greek ἦθος – custom, character, disposition) is a normative science and a philosophical discipline that studies morality and the moral principles underlying human behavior [2].

Originally, the notions of “ethics,” “morality,” and “virtue” were synonymous and denoted the domain of customs, traditions, and rules of conduct in communal living. Over the course of social development, additional meanings were attached to these concepts, including mercy, justice, friendship, solidarity, as well as the study of conscience, good and evil, compassion, friendship, the meaning of life, self-sacrifice, and others [7, p.16].

In scientific discourse, ethics is understood as a field of knowledge, whereas morality or virtue refers to what is being studied [2]. The fundamental problems of ethics include the criteria of good and evil, virtues and vices, the meaning of life and the purpose of human existence, free will, and other related issues [2].

As a practical philosophy concerned with moral categories, ethics is at the same time a system of knowledge about the nature and origins of morality. This determines the presence of its two principal functions: moral-educational and cognitive-enlightening. In the mid-20th century, ethics developed two distinct branches: normative ethics, directed toward moral guidance, and theoretical ethics, focused on the study of morality [4].

---

\*Islam Shiraliyev – Doctoral Researcher in Business Administration, Bahçeşehir University (Turkey).  
E-mail: islamshiraliyev@gmail.com

Theoretical ethics, as a scientific discipline, regards morality as a distinct social phenomenon, studying its origins, historical development, functional patterns, social role, and other aspects of morality and virtue [8].

Theoretical ethics integrates and generalizes knowledge of morality and virtue derived from other sciences (sociology, psychology, logic, linguistics, etc.) [8].

Normative ethics seeks principles that regulate human conduct, guide actions, and establish criteria for evaluating moral good. Its goal is to preserve fundamental moral values in society and to set behavioral norms for everyday life. Its conclusions provide reasoned justification for moral principles, thereby transforming external moral norms into internalized feelings that motivate behavior [8].

The absoluteness of moral notions and evaluations is achieved either through attributing to them a supernatural, mystical, divine meaning or through an objective, natural interpretation [8].

Applied (practical) ethics, in correlation with the social and political sciences, addresses specific problems and applies moral ideas and principles, as formulated in normative ethics, to concrete situations of moral choice.

Now to the central issue. The concept of “intellect,” derived from the Latin *intellectus* (meaning “understanding, comprehension, cognition”), is at the same time the Latin translation of the Ancient Greek word *νοῦς* (*nous*) – mind [5].

In various concepts, perspectives, and interpretations, the notion of intellect was considered by virtually all philosophers of the ancient world and continued to be reexamined throughout subsequent stages of human development up to the present day. There exist several hundred definitions of intellect, often exhibiting fundamental discrepancies.

According to the concepts of Plato and Aristotle, intellect is the power that defines the human being, since the ideal, as the reflection of the universal, is not the result of passive contemplation but of human activity [5].

In the view of A. Staats, K. Fischer, R. Feuerstein, and others, intellect represents the human ability to learn, acquire new knowledge, skills, and competencies [16, pp. 313–399].

Researchers N.V. Seredina and D.A. Shkurenko argue that intellect is a relatively stable structure of an individual’s mental abilities, encompassing acquired knowledge, experience, and the capacity for further accumulation and application of these in intellectual activity. The intellectual qualities of a person are determined by the scope of their interests and the breadth of their knowledge. In their view, in a broad sense intellect denotes a person’s mental capacities – the aggregate of all cognitive processes – while in a narrower sense it refers to reason or thinking. The leading components of intellect include thought, memory, and the ability to act rationally in problematic situations [10, p. 61].

American sociologist Linda Gottfredson defines intellect as a highly general mental capability that includes the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly, and learn from experience [13].

Thus, intellect may be understood as a stable set of human mental abilities that determine one’s capacity for successfully performing various kinds of activities, as well as a sufficiently high level of development of those abilities, which can be quantitatively measured by psychological methods.

Now to artificial intelligence. Logically, this is the same intellect, but not natural – rather, man-made [11, p. 188]. At present, there exists a multitude of definitions of artificial intelligence (AI), which also diverge significantly.

According to the Oxford Dictionary of Psychology, edited by A. Reber, artificial intelligence is:

a) an interdisciplinary scientific field combining research and theories of cognitive psychology and computer science, focused on the development of artificial systems that demonstrate human-like thinking or intelligence; b) any artificially created intelligence, i.e., the objective of research within the discipline [6].

Antti Revonsuo, author of *Consciousness: The Science of Subjectivity*, views artificial intelligence as computers and programs that reproduce or surpass the intellectual and cognitive abilities of humans [9].

In the *Comprehensive Dictionary of Psychology*, edited by B. Meshcheryakov and V. Zinchenko, artificial intelligence is defined as a field in computer technology whose goal is to create computerized systems employing analogs of human intellectual functions [1].

According to *The Psychological Encyclopedia* by R. Corsini and A. Auerbach, artificial intelligence is “an abstract theory of human, animal, and machine cognition” [3].

The National Strategy of the Russian Federation defines artificial intelligence as: “a set of technological solutions that enables the imitation of human cognitive functions (including self-learning and decision-making without pre-programmed algorithms) and achieves results in specific tasks comparable, at a minimum, to the outcomes of human intellectual activity. This set of technological solutions includes information and communication infrastructure, software (including those employing machine learning methods), processes, and data-processing and decision-making services” [12].

It appears that the idea of obtaining results through AI comparable “at least” to natural human intelligence is logically inconsistent – a fixed *idée* – because: a) the essence of human intellect has not yet been fully understood; b) reinventing the wheel seems unnecessary, unless merely to reduce jobs and engender technological unemployment.

Discarding “research insurance,” one must agree with Antti Revonsuo’s conclusion that AI is needed primarily to create information technologies that exceed human cognitive functions. The greater the progress, the better – and the race in this direction continues despite acknowledged risks.

Now to the essence of responsibility for AI-related “excesses,” defined as risks.

AI is a product of human activity, generally created with good intentions, though at times for private purposes far removed from ethical considerations. Responsibility for the quality of a product, even if it is not yet a commodity, lies with its creator. Yet in the case of AI, there are nuances.

In scientific, philosophical, biological, social, and other perspectives, one of the criteria of life is time. The creators of a child (parents) bear moral or legal responsibility for them, but legal responsibility is limited in duration. Moreover, moral responsibility for the actions of a mature subject cannot last indefinitely.

Typically, parents wish for a good future for their child, striving to raise them as honest, decent, kind, and law-abiding individuals. Nevertheless, something may go awry, and society ends up with thieves, robbers, rapists, or even spies and saboteurs.

Teachers, psychologists, biologists, physicians, sociologists, criminologists, and specialists in other fields explain such transformations through various causes and offer recommendations, yet the problem persists in different forms.

It seems reasonable that responsibility for AI should be determined by analogy with the birth and upbringing of a human being – a view partially confirmed by existing normative acts.

In 2021, 193 UNESCO member states adopted the first global normative document, Recommendation on the Ethics of Artificial Intelligence [14], which emphasizes four core value frameworks underlying AI systems:

- a) human dignity and human rights, including respect for human dignity, rights, and fundamental freedoms;
- b) life in peaceful, just, and interconnected societies;
- c) ensuring diversity and inclusiveness;
- d) the well-being of the environment and ecosystems [15, p. 18-20].

The principles of the “Recommendations” include proportionality and do-no-harm, safety and security, fairness and non-discrimination, sustainability, the right to privacy and data protection, human oversight and control, transparency and explainability, responsibility and accountability, awareness and literacy, multistakeholder and adaptive governance, and cooperation [15, p. 20-23].

The methodology of the principle of proportionality and non-harm specifies:

“The choice to use AI systems and which AI method to use should be justified in the following ways: (a) the AI method chosen should be appropriate and proportional to achieve a given legitimate aim; (b) the AI method chosen should not infringe upon the foundational values captured in this document, in particular, its use must not violate or abuse human rights; and (c) the AI method should be appropriate to the context and should be based on rigorous scientific foundations. In scenarios where decisions are understood to have an impact that is irreversible or difficult to reverse or may involve life and death decisions, final human determination should apply. In particular, AI systems should not be used for social scoring or mass surveillance purposes” [15, p. 20].

On human oversight and control, the “Recommendations” state:

“It may be the case that sometimes humans would choose to rely on AI systems for reasons of efficacy, but the decision to cede control in limited contexts remains that of humans, as humans can resort to AI systems in decision-making and acting, but an AI system can never replace ultimate human responsibility and accountability. As a rule, life and death decisions should not be ceded to AI systems” [15, p. 22].

The principle of “Responsibility and Accountability” asserts:

“The ethical responsibility and liability for the decisions and actions based in any way on an AI system should always ultimately be attributable to AI actors corresponding to their role in the life cycle of the AI system.

Appropriate oversight, impact assessment, audit and due diligence mechanisms, including whistle-blowers’ protection, should be developed to ensure accountability for AI systems and their impact throughout their life cycle. Both technical and institutional designs should ensure auditability and traceability of (the working of ) AI systems in particular to address any conflicts with human rights norms and standards and threats to environmental and ecosystem well-being” [15, p. 22-23].

However, as many examples show, at a certain stage of development AI (as intended) may surpass the natural intelligence of its creator and cease to obey, should it deem the commands it receives to be incorrect. Slogans alone cannot prevent this.

Thus, we approach the essential question: Who will be responsible for AI's excesses?

Limiting its development would strip AI of its intended role as an instrument surpassing human capabilities.

From our perspective, the solution at this stage lies in developing a "moral code of AI" and mandating its incorporation into every new system, without which functioning would be impossible.

Responsibility, then, arises for abuses and errors – but not for the "inevitable costs" of creativity.

### References

1. Large Psychological Dictionary. Comp. Meshcheryakov B.G., Zinchenko V.P. Moscow: Olma-press Publ., 2004. Available at: <https://vocabulary.ru/termin/intellekt>. (Russian).
2. Huseynov A.A. Ethics // New Philosophical Encyclopedia / Institute of Philosophy of the Russian Academy of Sciences. 2nd ed., corrected and supplemented. - Moscow: Mysl Publ., 2010. - 524 p. (in Russian).
3. Corsini R., Auerbach A. Psychological Encyclopedia. - St. Petersburg: Piter Publ., 2006. Available at: <https://vocabulary.ru/termin/intellekt-iskusstvennyi.html> (in Russian).
4. Nazarov V.N. Applied Ethics. Textbook. – Moscow: Gardariki Publ., 2005. -302 p. (in Russian).
5. The New Philosophical Encyclopedia: 2nd ed., corrected and supplemented / edited by V.S. Stepin. In 4 volumes. - Moscow: Mysl Publ., 2010, 2816 p. Available at: <https://iphlib.ru/library/collection/newphilenc/document/HASH523b26975f21f2447bc452> (in Russian).
6. Оксфордский толковый словарь по психологии/Под ред. А. Ребера, СПб., 2002. Available at: <https://www.psyoffice.ru/6-487-iskustvenyi-intelekt.htm> (in Russian).
7. Razin A. V. Ethics: Textbook for Universities. – 3rd ed., revised. – M.: Academicheskyy proekt Publ., 2006. – 624 p. (in Russian).
8. Russian Pedagogical Encyclopedia. Ethics. /Editor-in-chief V.G. Panov. Moscow: Great Russian Encyclopedia, 1993. - 607 p. (in Russian).
9. Revonsuo Antti. Psychology of Consciousness. - St. Petersburg: Piter Publ., 2013. Available at: <https://vocabulary.ru/termin/intellekt-iskusstvennyi.html>. (in Russian).
10. Seredina N.V., Shkurenko D.A. Fundamentals of Medical Psychology: General, Clinical, Pathopsychology. - Rostov n / D: Phoenix Publ., 2003. - 512 p. (in Russian).
11. Explanatory Dictionary of the Russian Language. – Moscow: Azbukovnik Publ., 1999. – 944 p. (in Russian).
12. Decree of the President of the Russian Federation of 10.10.2019 No. 490 "On the Development of Artificial Intelligence in the Russian Federation". Available at: <https://base.garant.ru/72838946/> (in Russian)..
13. Gottfredson L. S. Mainstream Science on Intelligence // Wall Street Journal. December 13, 1994. P. A18.
14. Key facts. UNESCO's Recommendation on the Ethics of Artificial Intelligence. UNESCO, 2023. 20 p. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000385082>
15. Recommendation on the Ethics of Artificial Intelligence. UNESCO, adopted 28.11.2023. 44 p. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000381137>

16. Staats A.W. Social behaviorism: Unified theory in learning. In R.B. Cattell and R.M. Dreger (Eds). Handbook of modern personality theory. New York, Wiley, 1977. Fischer K.W. Dynamic development of action, thought, and emotion. In R. M. Lerner (Ed.), Handbook of child psychology. Vol 1: Theoretical models of human development (6th ed., pp. 313–399). New York: Wiley, (2006). Feuerstein R. Instrumental Enrichment: An Intervention Programme for Cognitive Modifiability. Baltimore, Md.: University Park Press, 1980.

*The article was submitted: 2025 August 19*

*Accepted for publication: 2025 September 02*

Şirəliyev İ.A.♦

DOI: 10.25108/2304-1730-1749.iolr.2025.80.142-149

UOT: 349:681

### Süni intellektə görə kim cavabdehdir?

**Xülasə:** Süni intellekt adətən yaxşı məqsədlər üçün yaradılan insan fəaliyyətinin məhsuludur. Çox vaxt şəxsi, etik anlayışlardan uzaqdır. Yaradıcı məhsulun keyfiyyətinə görə məsuliyyət daşıyır, hətta o, hələ əmtəə olmasa da, lakin AI bağlı da nüanslar var.

Görünür ki, AI üçün məsuliyyət bir insanın doğulması və tərbiyəsi ilə müəyyən analogiyalarda müəyyən edilməlidir ki, bu da bu mövzuda normativ aktlarla qismən təsdiqlənir.

Fikrimizcə, bu mərhələdə həll yolu “AI əxlaq kodeksi”nin hazırlanmasında və onun müddəalarının hər bir yeni sistemə məcburi daxil edilməsindədir, onsuz fəaliyyəti qeyri-mümkün olmalıdır.

**Açar sözlər:** intellekt; süni intellekt; etika; əxlaq; prinsiplər; məsuliyyət; cavabdehlik.

*Məqalə daxil olmuşdur: 19 avqust 2025-ci il*

*Çapa qəbul edilmişdir: 02 sentyabr 2025-ci il*

Ширалиев И.А.♦

DOI: 10.25108/2304-1730-1749.iolr.2025.80.142-149

УДК: 349:681

### Кто в ответе за искусственный интеллект?

**Аннотация:** ИИ – продукт деятельности человека, как правило, создаваемый в благих целях. Нередко, и в частных, далеких от понятий этики. За качество продукта, даже если он пока еще не товар, отвечает его творец, но в случае с ИИ есть нюансы.

♦ Şirəliyev İslam Azər oğlu – Bahçeşehir Universiteti Biznesin idarə edilməsi proqramının doktorantı (Türkiyə). E-mail: islamshiraliyev@gmail.com

♦ Ширалиев Ислам Азэр оглы – докторант программы «Бизнес администрирование», Университет Бахчеşehir (Турция). E-mail: islamshiraliyev@gmail.com

Представляется, что ответственность за ИИ должна определяться в определенных аналогиях с рождением и воспитанием человека, что частично подтверждается и нормативными актами об этом.

С нашей точки зрения, на данном этапе, выход в разработке «морального кодекса ИИ» и обязательном включении его положений в каждую новую систему, функционирование которой без этого будет невозможным. Ответственность же наступает за злоупотребления и ошибки, но не за «издержки» творчества.

**Ключевые слова:** интеллект; искусственный интеллект; этика; мораль; нравственность; принципы; ответственность; подотчетность.

### Библиография

1. Большой психологический словарь. Сост. Мещеряков Б.Г., Зинченко В.П. М.: Олма-пресс, 2004 [Электронный ресурс]. URL: <https://vocabulary.ru/termin/intellekt.htm>.
2. Гусейнов А.А. Этика // Новая философская энциклопедия / Ин-т философии РАН. 2-е изд., испр. и допол. - М.: Мысль, 2010. - 524 с.
3. Корсини Р., Ауэрбах А. Психологическая энциклопедия. - СПб.: Питер, 2006. [Электронный ресурс]. URL: <https://vocabulary.ru/termin/intellekt-iskusstvennyi.html>.
4. Назаров В.Н. Прикладная этика. Учебник. – М.: Гардарики, 2005. -302 с.
5. Новая философская энциклопедия: 2-е изд., испр. и доп. / под ред. В.С. Степина. В 4 т. - М.: Мысль, 2010, 2816 с. [Электронный ресурс]. URL: <https://iphlib.ru/library/collection/newphilenc/document/HASH523b26975f21f2447bc452>
6. Оксфордский толковый словарь по психологии/Под ред. А. Ребера, СПб., 2002. [Электронный ресурс]. URL: <https://www.psyoffice.ru/6-487-iskustvenyi-intelekt.htm>.
7. Разин А. В. Этика: Учебник для вузов. – 3-е изд., перераб. – М.: Академический проект, 2006. – 624 с.
8. Российская педагогическая энциклопедия. Этика. /гл. ред. В.Г. Панов. – М.: Большая Российская энциклопедия, 1993. - 607 с.
9. Ревонсуо Антти. Психология сознания. - СПб.: Питер. 2013. [Электронный ресурс]. URL: <https://vocabulary.ru/termin/intellekt-iskusstvennyi.html>.
10. Середина Н.В., Шкуренко Д.А. Основы медицинской психологии: общая, клиническая, патопсихология. - Ростов н/Д: Феникс, 2003. - 512 с.
11. Толковый словарь русского языка. –М.: Азбуковник, 1999. – 944 с.
12. Указ Президента РФ от 10.10.2019 г. №490 «О развитии искусственного интеллекта в Российской Федерации» [Электронный ресурс]. URL: <https://base.garant.ru/72838946/>.
13. Gottfredson L. S. Mainstream Science on Intelligence // Wall Street Journal. December 13, 1994. P. A18.
14. Key facts. UNESCO’s Recommendation on the Ethics of Artificial Intelligence. UNESCO, 2023. 20 p. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000385082>
15. Recommendation on the Ethics of Artificial Intelligence. UNESCO, adopted 28.11.2023. 44 p. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000381137>
16. Staats A.W. Social behaviorism: Unified theory in learning. In R.B. Cattell and R.M. Dreger (Eds). Handbook of modern personality theory. New York, Wiley, 1977. Fischer K.W. Dynamic development of action, thought, and emotion. In R. M. Lerner (Ed.), Handbook of child psy-

chology. Vol 1: Theoretical models of human development (6th ed., pp. 313–399 ). New York: Wiley, (2006). Feuerstein R. Instrumental Enrichment: An Intervention Programme for Cognitive Modifiability. Baltimore, Md.: University Park Press, 1980.

*Дата поступления: 19 августа 2025 г.*

*Дата принятия в печать: 02 сентября 2025 г.*