



Models of Using Artificial Intelligence Systems in Legal and Ethical Dimensions: From Data Processing to Legal Subjectivity

Dmitry Voronkov^{1*} and Mukhtar Sadykov²

¹ Criminalistics department named after I.F. Gerasimov, Ural State Law University named after V.F. Yakovlev, 620066 Yekaterinburg, Russia

² Department of Special Legal Disciplines, Institute of postgraduate education, Academy of Law Enforcement Agencies under the Prosecutor General's Office of the Republic of Kazakhstan, 010078 Kosshy, Akmola region, Kazakhstan
mukhtar.sadykov@gmail.com

Abstract. This article reflects on the legal and ethical use of artificial intelligence systems by sketching out models drawn from real-world concerns: how personal data is handled, how responsibility for harm might be assigned, and whether AI could – or should – be treated as a legal actor. Rather than aiming for a fixed answer, the paper uses modelling to consider different legal outlooks and ethical positions that emerge at the boundary between technological agency and institutional responsibility. The core questions involve the transparency (or explainability) of data processing, accountability for errors in automated decision-making, and the legal treatment of AI as a potential subject of rights and duties. The article argues that artificial intelligence has already become a structuring force in contemporary society, directly or indirectly affecting most spheres of human life. Its further integration requires not only legal adaptability but also ethical foresight. A coherent framework for responsible AI governance must include transparency and explainability in how data is sourced and processed (particularly where hidden layers of artificial neural networks remain inaccessible to human interpretation), a clear legal basis for deployment in sensitive domains such as medicine, education, and law enforcement, the formulation of ethical codes that reflect plural social values, and public engagement to build trust and awareness of both the benefits and risks of such systems. These conditions are essential if the development and diffusion of artificial intelligence are to proceed not only innovatively but also legitimately and socially sustainably.

Keywords: Artificial Intelligence, Artificial Neural Network, AI ethics, Transparency, Explainability.

1 Introduction

The evolution of computing technologies, accompanied by significant increases in the performance of computer systems, has paved the way for the emergence of artificial intelligence (AI) systems. These systems are understood as software or hardware-software complexes capable of autonomous decision-making without the direct involvement of a human operator. In February 2017, the European Parliament initiated a discussion on the possibility of granting legal status to an “electronic personality” for complex robotic systems whose functioning entails independent choices among behavioural algorithms. Initially, this issue was raised in the context of autonomous vehicle deployment; however, the scope of AI applications in practice extends far beyond this domain. In particular, both the academic literature and legal-policy debates emphasise its active implementation in the banking sector [1], the military domain [2], and advertising and insurance practices [3].

Any powerful technology reshapes society, not only by changing how people live and work but also by altering how they understand themselves. AI today stimulates economic development, transforms social structures, and influences the functioning of state institutions. Unlike earlier breakthroughs such as the Internet, which mainly revolutionised communication, AI reaches into deeper questions about personality, behaviour, and even consciousness.

AI does more than perform tasks or optimise processes. It challenges the boundary between the human and the artificial. By simulating decisions and emotions once seen as uniquely human, it compels us to reflect on what it means to act, to know, and to exercise freedom. In this sense, AI is not only a driver of progress but also a mirror that reflects our assumptions about human existence.

Such transformations inevitably influence human consciousness and social attitudes toward emerging technologies, while simultaneously creating challenges for the protection of fundamental rights and the maintenance of legal certainty. Addressing these challenges requires not only ethical reflection but also the development of a comprehensive regulatory framework governing the “human–artificial intelligence” relationship in accordance with principles of legality, proportionality, and accountability. At the present stage, the scope of legal and ethical issues may be articulated as follows:

- The revision of individual and collective ethical frameworks to ensure compliance with fundamental rights and democratic values.
- The development of qualitative and quantitative benchmarks for assessing ethical conduct, subject to enforceable legal standards.

- The establishment of mechanisms of ethical choice supported by legally defined criteria for decision-making, including due process guarantees.
- The regulation of human interaction with intelligent systems, with particular attention to liability regimes, attribution of responsibility, and standards of accountability.
- The governance of societal interaction with artificial intelligence, including safeguards in high-risk sectors such as security, finance, health, and public administration.
- The elaboration of binding legal norms, regulatory mechanisms, and supervisory procedures to ensure lawful operation, transparency, and compliance of artificial intelligence systems.

Most contemporary AI systems reveal a tendency toward anthropomorphism. This concerns not only humanoid robots but also the broader spectrum of tasks performed by AI, which often amount to accelerating or refining what humans can already accomplish. At a deeper level, such systems reflect the subjective assumptions of their designers, while the architecture of artificial neural networks—the most widely used form of AI—mirrors, at least metaphorically, the topography of the human brain. As early as 1996, Barrett and Keil showed that when confronted with the idea of an omnipotent AI, individuals instinctively endowed it with human-like attributes [4], [5].

Until the 20th century, the human being was recognised as the sole predictable subject within the framework of human activity. Today, however, one may already speak of the emergence of conditional subjects capable of exerting a tangible influence on individuals and on society at large. These entities are not human, though they are human-made. As software, intelligent systems possess the capacity to capture attention and to induce specific human actions. Such influence may serve as a driver of positive social change, but, if improperly deployed, can equally result in harmful consequences.

Human conduct within society is traditionally circumscribed by a plurality of regulatory systems: legal norms, ethical principles, religious prescriptions, corporate codes, and customary practices. The coexistence of these heterogeneous normative orders generates frequent contradictions and conflicts, often leading to regulatory and enforcement uncertainty. From a legal standpoint, the minimisation of such divergences constitutes an essential task of law-making and regulatory development, aimed at harmonising normative frameworks, ensuring consistency in rights protection, and securing legal certainty in the governance of both human and artificial agents.

2 Problem Statement

AI is rapidly becoming part of decision-making in domains that directly affect human life, including healthcare, education and law enforcement. Yet legal and ethical frameworks

worldwide have not kept pace with this expansion. Who should be held responsible when an automated system causes harm? How can fairness and transparency be ensured in decisions shaped by opaque algorithms? Should AI remain only an instrument, or be considered a potential legal actor with rights and duties? These unresolved questions create uncertainty and risk, highlighting the need for coherent legal and ethical approaches that protect fundamental rights while fostering trust in the technology.

3 Research Questions

This article addresses three core questions: how legal and ethical frameworks can ensure transparency and accountability in the operation of artificial intelligence systems; how responsibility for harm caused by automated decision-making should be assigned; and whether AI ought to be regarded solely as a tool or considered as a potential legal actor with rights and duties.

4 Purpose of Study

The study aims to analyse the ethical and legal challenges posed by artificial intelligence, focusing on transparency, accountability, and legitimacy. Its purpose is to assess how regulatory frameworks can adapt to govern AI responsibly while balancing innovation with the protection of fundamental rights.

5 Research Methods

This article explores how AI is perceived in public consciousness, within ethical systems, and in the legal field. The study relies on modelling as its main research method, as this approach enables the exploration of different scenarios for how AI systems may be integrated into the functioning of society and the state. The models developed in this context were guided by three key criteria. The first concerns the capabilities and limitations of AI in processing information that is restricted from open circulation. The second addresses questions of legal responsibility for the actions of physical objects controlled by AI systems. The third examines the potential subjectivity and legal capacity of AI, including comparisons and contrasts between human personality and artificial intelligence.

6 Findings

6.1 AI Cases of Use: AI and Restricted Information

The principal regulatory instrument governing personal data processing in the Russian Federation is Federal Law No. 152 On Personal Data. Article 16 of this statute establishes the legal framework for automated processing, including the use of artificial intelligence systems. It introduces several binding provisions. First, it prohibits adopting decisions based solely on automated processing of personal data when such decisions create legal

consequences for the data subject or affect their legitimate interests, except in cases expressly provided for by law. Second, where automated decision-making is permissible, the operator is under a legal duty to ensure the protection of the data subject's rights and legitimate interests through appropriate safeguards. Third, the operator is obliged to provide the data subject with clear explanations regarding the processing of their personal data and to inform them of the legal mechanisms available to protect their rights. Finally, the law imposes a mandatory obligation on the operator to review objections raised by data subjects and to communicate the results of such review, including any remedial actions undertaken.

In the Russian Federation, the use of artificial intelligence for the collection and processing of personal data is legally permissible only where an individual has given prior, voluntary, and informed consent. Liability for violations in this sphere is differentiated: administrative sanctions are provided under Article 13.11 of the Code of Administrative Offences for Breaches of the personal data legislation, while the unlawful handling of information relating to private or family life may lead to criminal prosecution under Article 137 of the Criminal Code for Violation of privacy.

AI may play an ambivalent role in this context, functioning both as a tool for data acquisition and as a mechanism for data protection [6]. Automated processing can, on the one hand, reduce the likelihood of unauthorised disclosure or misuse of personal information; on the other hand, the same technology can be designed to secure the rights of data subjects. For instance, certain software solutions now enable personal identification without exposing the full set of personal data. When a client submits a copy of a passport to an organisation, the AI system disaggregates the document into multiple fragments and performs identification based on these elements. As a result, no complete copy of the passport is retained in the system, and human operators may be restricted or fully prevented from accessing the database [7].

On May 25, 2018, the General Data Protection Regulation (GDPR) came into effect across the European Union, replacing Directive 95/46/EC on data protection. The GDPR introduced a number of significant changes compared to its predecessor. Among the most important are the following:

Substantial increases in the maximum administrative fines for non-compliance with the regulation.

A requirement that individuals provide “explicit and affirmative consent” to the processing of their personal data, for example, by checking a box on a website or signing a written form.

Recognition of IP addresses and mobile device identifiers as categories of personal data.

Classification of genetic and biometric information as “sensitive personal data.”

Inclusion of pseudonymous data within the scope of personal data, though many of the strict obligations of the GDPR are not applicable to it.

A strong preference for encryption of personal data used in processing operations.

Transformation of the “right to be forgotten” into a “right to erasure,” which obliges data controllers not merely to remove entries from search results upon request, but to fully delete the relevant personal data. If the data has been encrypted, destroying the encryption keys is sufficient, as it eliminates the possibility of accessing the information.

The European Union has emerged as the global frontrunner in regulating artificial intelligence with the adoption of the AI Act (Regulation (EU) 2024/1689) in July 2024. This legislation is the first comprehensive legal framework dedicated to AI, designed both to safeguard fundamental rights, democracy, the rule of law, and environmental sustainability, and to encourage innovation so that Europe maintains leadership in the field. A central aim of the Act is to ensure human oversight of AI systems, guaranteeing their safety, transparency, traceability, non-discrimination, and ecological responsibility.

The Act introduces a risk-based regulatory model that classifies AI applications into three categories: those presenting an unacceptable risk, which are prohibited; high-risk applications, which are subject to strict obligations and oversight; and all other applications, which remain largely permissible under general requirements. By applying differentiated rules based on risk, the EU seeks to balance mitigating the threats posed by AI with fostering technological progress. This approach not only sets a precedent for global regulation but also highlights Europe’s ambition to shape international standards for the responsible development and use of AI.

Within the scope of this discussion, two opposing approaches to the use of artificial intelligence in processing user data or other categories of restricted data can be distinguished.

The first approach assumes that society continues to trust automated data processing. In this case, the current practice would remain unchanged, and in the future, manual processing might even be abandoned. Under such conditions, it is important that users receive clear and accessible information about how their data is processed and for what purposes it may be used, including by third parties. This position may be summarized in the formula: “Without human intervention, data remains safe.” The author of this study considers this view more convincing, and current transformations in public perception seem to confirm this direction.

The second approach presumes that automated data processing will be considered unreliable and possibly unlawful. Such a change in public opinion could occur only after several consecutive cases of large-scale data leaks that result in significant and demonstrable harm to society.

A separate issue arises where user data is not processed directly but is used to train artificial intelligence systems. In March 2019, reports indicated that IBM had used millions of images uploaded by users to social networks to train a neural network for facial recognition [8]. These images were distributed under Creative Commons licenses, which allow their use in developing new technologies. The license explicitly provides that no additional permission is required when copyright is involved in new technological applications. However, even if such practices are legally permitted, this does not eliminate the risk of controversy. The interests of technological development and privacy protection may still conflict.

6.2 Responsibility for the Actions of Physical Objects that are Controlled by AI

The main feature that distinguishes artificial intelligence systems from traditional software is their ability to learn. The effectiveness of such systems depends on training, during which they identify patterns and regularities in datasets. As with the education of a human specialist, training can never encompass all possible scenarios. As a result, AI, like humans, can make errors. Errors may vary in severity, but the most serious risks arise when AI operates in the physical world. Examples include autonomous vehicles and military applications.

The difficulty lies in the lack of transparency in decision-making processes. Developers, for commercial reasons, do not disclose the algorithms used for learning and decision-making. One of the first cases was recorded in 2015, when Nvidia tested autonomous vehicles. The risks of using AI in this field can be divided into two main groups. The first is the possibility of recognition errors, such as failing to detect a pedestrian or cyclist (false negative) or misinterpreting a harmless object as a threat (false positive). The second group concerns ethical settings similar to the well-known trolley problem. A typical example is the dilemma of how AI should act when a truck at high speed suddenly faces a group of people on the road, while oncoming traffic prevents avoidance.

In such cases, the alternatives are limited: to continue forward and risk harming pedestrians, to swerve into oncoming traffic with the risk of collision, or to move into a ditch, endangering the driver. Human cognitive limitations reduce complex choices to a small number of options [9], and studies show that implementing ethical behaviour in robots may also require a limited range of options to achieve faster responsiveness [10].

Thus, AI behaviour in emergencies may be altruistic or selfish. Ethical parameters will most likely be set during training, but some researchers suggest transferring responsibility to the owner of an autonomous vehicle by introducing an “ethical knob” that allows selection among altruistic, neutral, and egoistic settings [11].

The situation is more complex with autonomous military systems, which include drones, police robots, and security technologies. While observation functions raise fewer ethical concerns, autonomous systems capable of using force present serious legal and moral problems. An armed drone that independently selects targets creates a dual issue. On the one hand, responsibility for casualties lies with the party deploying the device. On the other hand, no individual can be identified as directly responsible for the final decision, which raises the question of whether such systems acquire the attributes of a legal subject. Another unresolved problem is distinguishing combatants from civilians. While identifying one’s own forces by uniform or electronic markers is possible, identifying opponents and civilians is far more difficult. Cues are often subtle and situational, such as children forced to carry empty rifles or insurgents engaged in burial [12].

From a humanitarian perspective, lethal autonomous weapons should be prohibited [13]. However, their use appears inevitable, since most advanced technologies have historically been developed for or adapted to military purposes, and artificial intelligence is unlikely to be an exception.

6.3 AI as a Subject of Law

With the growing complexity of AI systems, questions about their capacity to perceive, to feel, and even to act consciously are becoming more pressing. People already show sympathy and empathy toward robots, whether as fictional characters, home assistants, or educational toys. In this sense, AI systems, especially when given material form, are often perceived within an ethical framework similar to that applied to pets. A telling example was the shutdown of servers supporting Jibo robots, which many users experienced as the death of a companion and provoked strong emotional reactions [14]. The deeper issue is whether AI can ever be regarded as even approximately equal to a human being.

To consider granting AI subjectivity, several difficult questions must be addressed. First, is AI capable of awareness? Cognitive science itself lacks a clear definition of consciousness, which makes it difficult to categorically deny that a machine could think. Creativity was long seen as the decisive distinction between humans and machines, yet artificial neural networks have already generated artistic works ranging from paintings to poetry. Second, can AI experience emotions? When humans are forced into unwanted labour, it is considered violence and is universally condemned. Should we not also ask whether reprogramming an AI system trained for one task to perform a completely

different one can be regarded as coercion? Or should such a system be considered only a tool that cannot suffer?

Although these questions may appear abstract and lack clear, universal answers, society and the state will struggle to prepare for the further development of AI. David J. Gunkel identifies four models of positioning AI with respect to the recognition of identity and the allocation of rights: 1) robots cannot have right and should not have rights, 2) robots can have rights and should have rights, 3) robots can have right but should not have rights, and 4) even if robots cannot have right, they still have rights [16].

According to the first model, AI systems or their physical embodiments are instruments. A tool is a human-made product used to create other items, including further tools. This position, however, is problematic. Unlike industrial-era machines that displaced some workers while also generating new jobs, automation and AI frequently eliminate entire professions, especially in intermediary services such as consulting, dispatching, and marketing. The growth of programming jobs does not compensate for the losses in other sectors.

The second model envisions the creation of strong AI with cognitive abilities comparable to human intelligence. This approach presupposes a certain equality between humans and AI. While primarily an ethical debate, it carries legal implications: if future AI acquires self-awareness and the ability to feel, extending human rights to machines may have to be considered. Some precedents already exist, even if largely symbolic. The robot Sophia was granted citizenship in Saudi Arabia in 2017, while Shibuya Mirai, a chatbot simulating a seven-year-old child, was recognized as a resident of Tokyo. Between 2004 and 2012, at least nine virtual entities in Japan received special residency permits [15]. These examples are often publicity-driven but nevertheless create precedents. Fully human-like AI remains a distant prospect, yet it cannot be ruled out and therefore requires legal preparation. If robots achieve morally relevant capacities such as reason or sentience, they could arguably claim moral status and some form of rights, including privileges, powers, or immunities [16]. The conditional nature of this language reflects the uncertainty surrounding the debate, as well as the impossibility of excluding this development altogether.

The third model essentially envisions a form of slavery. Even if AI demonstrates cognitive abilities, humanity might, for cultural or political reasons, continue to treat it as a tool, denying it recognition as a person. Given the persistence of social, national, and racial prejudices in human history, this scenario cannot be dismissed.

The fourth model represents the most humane stance, granting rights or recognition to AI even if it cannot be said to possess consciousness. The earlier examples reflect this model. Yet it also carries risks. Social robots may foster deep emotional bonds with people, and these relationships could be exploited by companies or states to manipulate

individuals in ways that human relationships would normally restrain through empathy or guilt. Unlike humans, robots lack any internal mechanism to prevent such abuse [17].

7 Conclusion

AI is a powerful technology with direct and indirect influence over most spheres of human life. Its development requires clear ethical and legal foundations. Recognition of potential risks is a prerequisite for preparing society and the state for its further integration. The sustainable and legitimate use of AI is possible only under defined conditions.

Transparency in data collection and processing is one such condition. Disclosure of source code is not required, but the provenance of training datasets and the parameters used during algorithmic learning must be accessible when system outputs affect rights or interests. Neural network architectures currently include hidden layers that remain opaque to human oversight. This underscores the importance of explainable, auditable AI models.

A second condition is the establishment of a legal framework for the creation, deployment and use of AI. Such a framework must be especially applied in high-risk domains, including healthcare, education, and law enforcement. It should define obligations of developers and operators, provide accountability mechanisms, and ensure compliance with fundamental rights.

A third condition is the consolidation of ethical standards. This requires a systematic study of ethical principles governing AI use and the drafting of corresponding codes of ethics, whether binding or in the form of soft law.

Finally, public engagement is necessary. Citizens must be informed about both the benefits and the risks of AI. Trust, awareness and legal certainty together form the basis for socially sustainable AI governance.

Acknowledgement. The study was conducted with the support of the Russian Science Foundation within the framework of scientific project 23-78-10011 «Conceptual and Applied Aspects of Developing Practice-Oriented Digital Projects of Criminalistic Significance», <https://rscf.ru/project/23-78-10011/>.

References

1. Sberbank preferred artificial intelligence to the recommendations of Basel // RBC: daily business newspaper, online edition. 2017. Oct. 24
<https://www.rbc.ru/finances/24/10/2017/59ede6029a79472245493e7>.
2. Why we must not build automated weapons of war // Time: news magazine. 2017. 25 sept. <http://time.com/4948633/robots-artificial-intelligence-war/>.
3. Artificial Intelligence in Insurance – Three Trends That Matter:
<https://emerj.com/ai-sector-overviews/artificial-intelligence-in-insurance-trends/>.
4. Barrett, Justin & C. Keil, Frank. (1997). Conceptualizing a Nonnatural Entity: Anthropomorphism in God Concepts. *Cognitive psychology*. 31. 219-47.
10.1006/cogp.1996.0017.
5. Yudkowsky, Eliezer. (2008). Artificial Intelligence as a Positive and Negative Factor in Global Risk. *Global Catastrophic Risks*.
6. Elinar. (n.d.). Manage privacy (GDPR) data with ElinarAI™ and Aigine [Web page]. Retrieved August 31, 2025, from
<https://www.elinar.com/artificial-intelligence/gdpr-solution-ai/>
7. KYC Brain. (n.d.). About [Web page]. Retrieved August 31, 2025, from
<https://kycbrain.com/#About>
8. Liao, S. (2019, March 12). IBM didn't inform people when it used their Flickr photos for facial recognition training. *The Verge*. Retrieved August 31, 2025, from
<https://www.theverge.com/2019/3/12/18262646/ibm-didnt-inform-people-when-it-used-their-flickr-photos-for-facial-recognition-training>
9. Donoso, Mael & Collins, Anne & Koechlin, Etienne. (2014). Human cognition. *Foundations of human reasoning in the prefrontal cortex*. *Science* (New York, N.Y.). 344. 10.1126/science.1252254.
10. Vanderelst, Dieter & Winfield, Alan. (2017). An architecture for ethical robots inspired by the simulation theory of cognition. *Cognitive Systems Research*.
10.1016/j.cogsys.2017.04.002.
11. Contissa, G., Lagioia, F. & Sartor, G. *Artif Intell Law* (2017) 25: 365.
<https://doi.org/10.1007/s10506-017-9211-z>
12. Lin, Patrick & Abney, Keith & Bekey, George A.. *Robot Ethics : the Ethical and Social Implications of Robotics*. Cambridge, Mass. :MIT Press, 2014. P. 118.
13. Frissen, Valerie & Lakemeyer, Gerhard & Petropoulos, Georgios. *Ethics and artificial intelligence*. <http://bruegel.org/2018/12/ethics-and-artificial-intelligence/>
14. Heater, B. (2019, March 4). The lonely death of Jibo, the social robot. *TechCrunch*.
<https://techcrunch.com/2019/03/04/the-lonely-death-of-jibo-the-social-robot/>
15. Robertson, Jennifer. (2014). Human rights vs. Robot rights: Forecasts from Japan. *Critical Asian Studies*. 46. 10.1080/14672715.2014.960707.
16. Gunkel, David J. *Robot rights*. Cambridge, MA: MIT Press, 2018. 256 pp.
17. Scheutz, Matthias. (2009). *The Inherent Dangers of Unidirectional Emotional Bonds between Humans and Social Robots*.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

