European Proceedings of
Social and Behavioural Sciences
EpSBS

www.europeanproceedings.com
e-ISSN: 2357-1330

DOI: 10.15405/epsbs.2022.06.76

AMURCON 2021
AmurCon 2021: International Scientific Conference

ETHICAL AND LEGAL ASPECTS OF DEVELOPMENT AND IMPLEMENTATION OF ARTIFICIAL INTELLIGENCE SYSTEMS

Yulia S. Matyuk (a)*
*Corresponding author

Vladivostok State University of Economics and Service, 41 Gogolya St., Vladivostok, Russia, 5080mm@mail.ru

Abstract

The subject of research is ethical norms and legal regulation of social relations connected with the application of artificial intelligence technologies. To date, some experience has been accumulated in the regulation of artificial intelligence technologies and related areas. In this connection, one of the key tasks of this research is the analysis and brief generalization of the current experience of regulation. The methodological basis of the study is general scientific (dialectical, system-structural methods, analysis, synthesis), as well as private legal methods of knowledge (formal-legal and comparative-legal methods). Research of the current acts, recognizing ethical bases of artificial intelligence regulation, shows that this issue is very actual both for separate states, international associations and professional communities. Many stakeholders have developed recommendations, ethical charters, principles, conceptual and programmatic documents, and some have already laid the foundation for the legal regulation of the development, introduction and application of artificial intelligence. Ethical comprehension of possibilities and limits of artificial intelligence technology use seems to be the primary option of its social regulation. As a result of the research, the following key principles are defined: safety, clearness, and responsibility, the priority of human rights, privacy and confidentiality of personal data. Ethical norms elaborated by science and tested in practice can and, if necessary, should form the basis of legal and regulatory acts.

Keywords: Artificial intelligence, ethics, law, principles

https://doi.org/10.15405/epsbs.2022.06.76
Corresponding Author: Yulia S. Matyuk
Selection and peer-review under responsibility of the Organizing Committee of the conference
eISSN: 2357-1330

1. Introduction

The tremendous promise of Artificial Intelligence (AI) technology is already evident today. AI has not yet caused significant problems, but its subject matter is already well outlined, posing a challenge to the whole society and the legal community in particular.

Early examples include fatal car accidents involving Tesla and Uber self-driving cars, protests of artificial intelligence developers at Google company against participation in US DoD military projects, cases of manipulation of information availability, sexism and racism in facial recognition algorithms and targeted advertising using artificial intelligence. Large-scale ethical issues arise in using artificial intelligence by public services to control citizens (Baranov et al., 2020).

Crucially, an AI system can make human-related decisions on its own, to analyze data so much and so quickly that humans cannot do (hence, humans cannot verify the validity of decisions).

In addition to their capacity to operate without direct human oversight and control, these technologies have several other responsibility-relevant and ethical characteristics, including their inscrutability and opacity. The combined effect of the inscrutability and opacity of algorithms results in their characterisation as ‘black boxes’, and these properties have direct implications for the transparency, explainability and accountability of and for the applications that utilise them (Burrell, 2016)

Accordingly, the main challenge is to determine whether the decisions made by an intelligent autonomous system are ethical, i.e. how ethical it is. (Gotovtsev et al., 2020).

In the future humanity will not reject artificial intelligence; it is more likely that artificial intelligence will be much more, more widely and more actively used in a variety of fields. It is important to set the ethical lines within which AI will be developed as quickly as possible, to limit the possibilities for its unethical use and to channel the energy of developers and the ideas of legislators into a direction that provides maximum safety and benefit to the society.

It should be noted that artificial intelligence is an extremely powerful technology that directly or indirectly affects most areas of human life, so it is essential to develop and study the ethical basis for the operation of such systems. Understanding the risks involved will ensure that society and the government are prepared for the development and wide usage of AI technology. However, legal regulation should be preceded by serious scientific research to study the ethical foundations of this technology, including the risks, their consequences and possible feedback of the society (Bakhteev, 2019).

2. Problem Statement

The expansion of artificial intelligence systems causes several technical, philosophical, legal and ethical questions, both about the permission of such systems and the need for ethical standards in their design (Malyshkin, 2019). The global community has been increasingly recently discussing the topic of artificial intelligence, robots and cyber-physical systems. Experts and ordinary people think over the times when a robot will be able to make independent decisions. Such changes will likely be not long in coming, which means that robots need to be regulated (Mamychev et al., 2020)

Up today, most documents are recommendations, charters of all kinds, etc. However, what many documents have in common is that they all record some kind of ethical and moral principles that should

https://doi.org/10.15405/epsbs.2022.06.76
Corresponding Author: Yulia S. Matyuk
Selection and peer-review under responsibility of the Organizing Committee of the conference
eISSN: 2357-1330

be a foundation of future legislation on robots. Jan Kleijssen (2020), Director of the Council of European Office on Cybercrime, notes that:

> contemporary legal instruments, such as the European Convention on Human Rights, certainly provide a basis for the protection of human rights and the rule of law. But it is also clear that there are now so many applications using artificial intelligence that are not covered by legal regulation. The Council of Europe is preparing proposals on a number of these issues based on the ECHR and other instruments as well as over 200 ethical charters and acts developed by industry, governments and non-governmental organizations. These instruments are not obligatory, for example, an ethical charter cannot be appealed to in court. Thus, a normative foundation for effective legal protection is necessary.

Artificial Intelligence is becoming one of the highest priorities on the agenda of various entities, both at national and international levels. Many countries have developed conceptual and policy documents, and some have found the legal development, introduction and application of Artificial Intelligence (Dremliuga et al., 2019).

3. Research Questions

Ethical and legal risks in using artificial intelligence systems are determined, besides technological problems, by the uncertain legal status of such systems and the resulting vague system of responsibility. Therefore, it is important to investigate an ethical-legal component of modern regulation of applying artificial intelligence technologies.

The subjects of research are ethical norms and legal regulation of social relations connected with applied artificial intelligence technologies.

4. Purpose of the Study

An analytical study of the issue of artificial intelligence regulation seems impossible without examining the regulatory documentation circulated. Nowadays, some experience in the regulation of artificial intelligence technologies and related areas has been accumulated. In this regard, one of the key tasks of this study is to review and summarize the current regulatory experience in brief.

At present, most documents are recommendations, charters of all kinds, etc. However, what many documents have in common is that they all validate some kind of ethical and moral principles that should serve as a basis for future legislation on robots. This study will touch upon some of them.

5. Research Methods

The scientific and methodological basis of the study was a complex of various methods of cognition, research, description and explanation of the studied thematic area. Let's list some of them. Dialectics is the main research method, since it has a great heuristic potential, allows for the identification of various facets of the phenomenon under study, its relationship with other phenomena and processes.

https://doi.org/10.15405/epsbs.2022.06.76
Corresponding Author: Yulia S. Matyuk
Selection and peer-review under responsibility of the Organizing Committee of the conference
eISSN: 2357-1330

The essence of the current legislation and ethical standards in the field of AI technologies application cannot be understood without taking into account the socio-economic, political and other factors that affect it at present. This is necessary to assess the possible ethical and legal risks of using AI. The use of the general scientific research method of induction (from the particular to the general) made it possible to draw several generalizing conclusions and assessments of the studied phenomena, processes, and relations. The comparative method was used to study the world experience in the use of AI technologies. In addition, the work used formal legal, specific sociological, as well as political and legal modelling methods necessary for assessing risks in the processes of using AI technologies.

6. Findings

A study of the current ethical acts for the regulation of Artificial Intelligence has shown that this issue is highly relevant to both different states and international associations and the professional community. Recommendations, ethical charters, principles, conceptual and policy documents have been developed by many stakeholders and some have already laid the foundation for the legal regulation of AI development, implementation and application (Gaivoronskaya et al., 2021).

In June 2017, the German Federal Ministry of Transport and Digital Infrastructure developed ethical guidelines for unmanned vehicles. The guidelines developed by the Ministry's Ethics Commission, provide 15 rules for programmed solutions embedded in unmanned cars. The Commission has studied ethical issues in detail, including whose lives are to be saved first (the so-called 'carriage issue'). The guidelines call for that unmanned cars must be programmed in such a way that all human lives are considered equal. If choices must be made between people, unmanned cars must make choices that cause as little damage as possible, regardless of age, race or gender. The Commission also further argues that there should be no obligation on people to sacrifice themselves for others (Automated ..., 2017

The South Korean Robot Ethics Charter 2007 is a code of ethics developed for those involved in the development, production and use of intelligent robots to prevent various kinds of negative consequences (such as disruption of public order) that may result from the development of the functions and intelligence of intelligent robots and to provide that intelligent robots contribute to the quality of human life.

According to The Japanese Society for Artificial Intelligence Ethical Guidelines, ethical standards for AI developers should form the basis of activities when building artificial intelligence systems. These ethical guidelines should provide a moral basis for members of The Japanese Society for Artificial Intelligence (JSAI) to understand their responsibilities to society better. JSAI members should implement and follow these principles, such as humanity, legality, respect for privacy, honesty, safety, transparency, social responsibility, etc. (The Japanese ..., 2017)

In April 2016, the IEEE Institute of Electrical and Electronics Engineers launched the Global Initiative on the Ethics of Autonomous and Intelligent Systems. It aims to advance public debate on the adoption of artificial intelligence technologies and to identify priority values and ethics. The IEEE published Version 2 of its Ethically Aligned Design principles in December 2017, (Ethically ..., 2017). Table 1 presents the main principles of the Ethically Aligned Design principles.

https://doi.org/10.15405/epsbs.2022.06.76
Corresponding Author: Yulia S. Matyuk
Selection and peer-review under responsibility of the Organizing Committee of the conference
eISSN: 2357-1330

Table 1. General Principles contained in the IEEE's Ethically Aligned Design (version 2)

Principles	Objectives
Human rights	Ensure autonomous and intelligent systems (AISs) do not infringe on internationally recognised human rights
Prioritising well-being	Prioritise metrics of well-being in the design and use of AISs because traditional metrics of prosperity do not take into account the full effect of AI systems technologies on human well-being
Accountability	Ensure that designers and operators of AISs are responsible and accountable
Transparency	Ensure AIS operate in a transparent manner
AIS* technology misuse and awareness of it	Minimise the risks of misuse of AIS technology

* *Automatic Identification System*

The Asilomar Principles for Artificial Intelligence were developed and adopted following a conference of developers and researchers in the field of artificial intelligence held in January 2017 in Asilomar, USA. So far, more than 3,500 scientists, developers, entrepreneurs and experts have signed on. Among them are Elon Musk, Stephen Hawking and Raymond Kurzweil, representatives of Google, Apple, Facebook, IBM, Microsoft, etc.

The Asilomar Principles for Artificial Intelligence are 23 principles for the safe and socially beneficial development of artificial intelligence in the near and long term. The principles are grouped into three sections. Research issues require funding for research into useful artificial intelligence, which includes complex issues of Computer Science; Economics, Law and Social Science; a constructive 'science-policy connection'; and a culture of cooperation, trust and transparency in technical research. The Ethics and Values section requires that the development and operation of artificial intelligence systems be safe and secure, transparent and accountable, protecting freedom, privacy, human dignity, rights and cultural diversity, broad opportunity and shared benefits. Longer-term objectives caution against assumptions about the upper limits of future AI capabilities and plan carefully for the possible development of general artificial intelligence (AGI) (Asilomar ..., 2017). Table 2 presents a list of the Asilomar Principles of Artificial Intelligence.

Table 2. Asilomar AI Principles (excerpt titles of the principles)

Research issues	Ethics and values	Longer-term issues
-Research goal	- Safety	- Capability caution
- Research funding	- Failure transparency	- Importance
- Research funding	- Judicial transparency	- Risks
- Science-policy link	- Responsibility	- Recursive self-improvement
- Research culture	- Value alignment	- Common good
- Race avoidance	- Human values	

https://doi.org/10.15405/epsbs.2022.06.76
Corresponding Author: Yulia S. Matyuk
Selection and peer-review under responsibility of the Organizing Committee of the conference
eISSN: 2357-1330

- Personal privacy
- Liberty and privacy
- Shared benefit
- Shared prosperity
- Human control

In September 2016, Amazon, DeepMindGoogle, Facebook, IBM and Microsoft launched the Partnership on Artificial Intelligence to Benefit People and Society (PAI). Its purpose is to explore and define best practices in Artificial Intelligence technology, to deepen public understanding of artificial intelligence. The partnership also serves as an open platform for discussion and interaction on artificial intelligence and its impact on people and society. From the beginning, PAI has evolved into a multi-disciplinary community of stakeholders with over 80 members ranging from commercial technology companies to civil society, academic and research institutes and start-ups (Pshava et al., 2020)

The Information Technology Industry Council (ITI) is a business association of technology companies based in Washington DC with over 60 members. The ITI published the Artificial Intelligence Policy Principles in October 2017 (Table 3). These principles define responsibilities in certain areas and call for public support for AI research and public-private partnerships (AI Policy Principles..., 2017).

Table 3. ITI AI Policy Principles

Responsibility: Promoting responsible development and use	**Opportunity for governments: Investing and enabling the AI ecosystem**	**Opportunity for public-private partnerships: Promoting lifespan education and diversity**
- Responsible design and deployment	- Investment in AI research and development	- Democratising access and creating equality of opportunity
- Safety and controllability	- Flexible regulatory approach	- Science, technology, engineering and mathematics education
- Robust and representative data	- Promoting innovation and the security of the Internet	- Workforce
- Interpretability	- Cybersecurity and privacy	- Public-private partnership
- Liability of AI systems due to autonomy	- Global standards and best practices	

As can be seen, many of today's conventional documents on ethical and legal principles for the development and introduction of artificial intelligence attach similar basic ethical principles. As a result of this synthesis, we have identified the following key principles.

- Safety. AI systems should be safe and reliable throughout their lifetime and should be controlled to the extent possible and applicable.
- Transparency. If an AI system causes harm, it should always be possible to understand the reason for this. The decision-making process of the artificial intelligence system should be

accompanied by convincing explanations that can be cross-checked by the professionals from the competent authorities.

- Responsibility. Developers of advanced AI systems take a key part in shaping the moral consequences of AI use, AI misuse and AI actions; they have an opportunity and responsibility to influence such consequences.
- Priority of human rights. AI systems should be designed and operated in a way that is compatible with the ideals of human dignity, human rights and freedoms and multiculturalism.
- Confidentiality of personal data and privacy. Given the ability of AI systems to analyze and use personal data, individuals should have the right to access, manage and control their personal data. The application of AI to personal data should not unreasonably restrict people's actual or perceived freedom.

7. Conclusion

Ethical reflection on the possibilities and limits of using Artificial Intelligence technology seems to be the primary option for its social regulation. Ethical norms developed by science and tested in practice can and, if necessary, should become the basis of legal acts. However, legal regulation should be preceded by the serious scientific study of the ethical foundations of this technology, including the risks and consequences and possible feedback of society. For artificial intelligence technologies to be fully approved, several conditions need to be met, such as transparency and accountability in receiving and using big data, inclusion and application of an appropriate code of ethics, public education, etc.

Acknowledgments

This research work was supported by the Grant of the President of the Russian Federation No. NSh-2668-2020.6 "National-cultural and digital trends in the socio-economic and political-legal development of the Russian Federation in the XXI century".

References

AI Policy Principles, Information Technology Industry Council. (2017). https://www.itic.org/resources/AI-Policy-Principles-FullReport2.pdf

Asilomar AI Principles. (2017). https://futureoflife.org/ai-principles/

Automated and connected driving. (2017). https://www.bmvi.de/SharedDocs/EN/publications/report-ethics-commission.pdf?__blob=publicationFile

Bakhteev, D. V. (2019). Riski i etiko-pravovyye modeli ispol'zovaniya sistem iskusstvennogo intellekta [Risks and ethical and legal models of using artificial intelligence systems]. *Legal Research*, (11), 1-11.

Baranov, P. P., Mamychev, A. Yu., Plotnikov, A. A., Voronov, D. Yu., & Voronova, E. M. (2020). Problems of Legal Regulation of Robotics and Artificial Intelligence from the Psychological Perspective. *Propósitos y Representaciones*, *8*(2). https://doi.org/10.20511/pyr2020.v8n2.511

Burrell, J. (2016). 'How the machine 'thinks': Understanding opacity in machine learning algorithms.' *Big Data & Society, 3*(1), 1-12.

Dremliuga, R. I., Mamychev, A. Yu., Dremliuga, O. A., & Matyuk, Yu. S. (2019). Artificial intelligence as a subject of law: pros and cons. *Revista Dilemas contemporáneos: Educación, Política y Valores*, *7*(1), 127.

Ethically Aligned Design (Version 2), IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, (2017). https://standards.ieee.org/wp-content/uploads/import/documents/other/ead_v2.pdf

Gaivoronskaya, Y. V., Dremliuga, R. I., Mamychev, A. Y., & Miroshnichenko, O. I. (2021). Ethical issues of digitization. *Laplage In Review*, *7*(3). 539-547. https://doi.org/10.24115/S2446-62202021731340

Gotovtsev, P. M., Karpov, V. E., Nestik, T. A., & Potapova, E. G. (2020). *Etika iskusstvennogo intellekta. Etika i tsifra: eticheskiye problemy tsifrovykh tekhnologiy [Ethics of Artificial Intelligence. Ethics and Digital: Ethical Issues of Digital Technologies]*. https://ethics.cdto.center/3_1

Kleijssen, J. (2020). *More inter-governmental cooperation is needed using Artificial Intelligence to fight Covid-19 Coronavirus: the interview of Jan Kleijssen, Director of Information Society Action against Crime (Council of Europe)*. https://www.coe.int/en/web/portal/covid-19-artificial-intelligence

Malyshkin, A. V. (2019). Integratsiya iskusstvennogo intellekta v obshchestvennuyu zhizn': nekotoryye etiko-pravovyye voprosy [Integration of artificial intelligence into public life: some ethical and legal issues]. *Bulletin of St. Petersburg University. Pravo*, *10*(3), 444-460.

Mamychev, A. Y., Gaivoronskaya, Y. V., & Rusanova, I. O. (2020). Typology of risks and threats caused by digitalization. *Turismo-Estudos E Praticas, 5,* 12.

Pshava, V. V., Babenko, I. A., & Primak, Y. S. (2020) The possibility of using artificial intelligence in legal proceedings of contemporary Russia. *Revista Inclusiones, 7*, 396-405.

The Japanese Society for Artificial Intelligence Ethical Guidelines. (2017). http://ai-elsi.org/wp-content/uploads/2017/05/JSAI-Ethical-Guidelines-1.pdf