UDC 343.2
DOI https://doi.org/10.32782/2524-0374/2024-8/77

ARTIFICIAL INTELLIGENCE AND THE LAW: THE CHALLENGES OF ESTABLISHING CRIMINAL LIABILITY IN THE DIGITAL AGE

ШТУЧНИЙ ІНТЕЛЕКТ І ПРАВО: ПРОБЛЕМИ ВСТАНОВЛЕННЯ КРИМІНАЛЬНОЇ ВІДПОВІДАЛЬНОСТІ В ЕПОХУ РОЗВИТКУ ЦИФРОВИХ ТЕХНОЛОГІЙ

Dvornikova P.A., Fourth-year student
Yaroslav Mudryi National Law University

Osadchaya A.S., PhD of Legal Sciences,
Assistant at the Department of Criminal and Legal Policy
Yaroslav Mudryi National Law University

The article provides a detailed analysis of the essence, definition, and classification of artificial intelligence (AI), which is an extremely relevant topic in the context of rapid technological development. The author emphasizes the lack of a unified opinion not only regarding the definition of the concept of artificial intelligence but also concerning its capabilities and potential threats. This issue sparks numerous discussions among scientists, legal experts, and technologists, as different approaches to interpreting AI can significantly impact the legal and ethical aspects of its use. The paper also notes that the development of technologies, particularly autonomous systems, presents new challenges for society related to the legal status of artificial intelligence. Difficulties arise not only in the context of defining the legal status of AI but also in determining its liability under the law. The author highlights that traditional legal frameworks may not always adequately regulate the new realities associated with autonomous systems capable of making decisions without human intervention. By examining legislative and doctrinal norms in various jurisdictions, such as the USA, EU, Brazil, China, and Ukraine, the author seeks to clarify who is responsible for the actions of autonomous systems with artificial intelligence. This question is extremely important because in cases of harm or wrongdoing caused by an autonomous system, there is a need to identify the subject of responsibility. Whether it will be the software developer, the system's owner, or the system itself is a matter that requires detailed analysis. The article also explores the possibilities of holding artificial intelligence criminally liable. The author investigates whether it is possible to recognize AI as a subject of criminal law or whether responsibility will always rest with physical or legal persons. Different models of liability are examined, including the concept of "collective responsibility," which may be applicable in the context of actions taken by autonomous systems. Thus, the paper highlights the complexity and multifaceted nature of the issues surrounding artificial intelligence in the context of law and ethics. The author calls for the development of new legal norms and ethical standards that could adequately regulate relationships between humans and autonomous systems in the face of rapid technological advancement.

Key words: artificial intelligence, machine learning, criminal liability, criminal subject.

У роботі детально аналізується суть, дефініція та класифікація штучного інтелекту (ШІ), що є надзвичайно актуальною темою в умовах швидкого розвитку технологій. Автор підкреслює, що існує відсутність єдиної думки не лише щодо визначення концепції штучного інтелекту, а й щодо його можливостей та потенційних загроз. Це питання викликає численні дискусії серед науковців, правознавців та технологів, оскільки різні підходи до трактування ШІ можуть суттєво вплинути на правові та етичні аспекти його використання. У роботі також зазначається, що розвиток технологій, зокрема автономних систем, ставить перед суспільством нові виклики, пов'язані з юридичним статусом штучного інтелекту. Складнощі виникають не лише в контексті визначення правового статусу ШІ, але й у питанні його відповідальності перед законом. Автор акцентує увагу на тому, що традиційні правові рамки не завжди можуть адекватно регулювати нові реалії, пов'язані з автономними системами, які здатні приймати рішення без людського втручання. Досліджуючи законодавчі та доктринальні норми в різних юрисдикціях, таких як США, ЄС, Бразилія, Китаї та Україна, автор намагається з'ясувати, хто несе відповідальність за дії автономних систем зі штучним інтелектом. Це питання є надзвичайно важливим, оскільки в разі завдання шкоди або вчинення правопорушення автономною системою виникає необхідність у визначенні суб'єкта відповідальності. Чи буде це розробник програмного забезпечення, власник системи чи сама система — це питання потребує детального аналізу. У роботі також розглядаються можливості притягнення штучного інтелекту до кримінальної відповідальності. Автор досліджує, чи можливо визнати ШІ суб'єктом кримінального права, або ж відповідальність завжди буде лежати на фізичних чи юридичних особах. Розглядаються різні моделі відповідальності, включаючи концепцію "колективної відповідальності", яка може бути застосована в контексті дій автономних систем. Таким чином, робота висвітлює складність і багатогранність проблематики штучного інтелекту в контексті права та етики. Автор закликає до необхідності розробки нових правових норм та етичних стандартів, які могли б адекватно регулювати відносини між людьми та автономними системами в умовах стрімкого розвитку технологій.

Ключові слова: штучний інтелект, машинне навчання, кримінально-правова відповідальність, суб'єкт злочину.

In the 21st century, we have witnessed unprecedented advancements in technology that fundamentally alter our daily lives, communication, work, and even modes of thinking. One of the most striking and simultaneously controversial innovations is artificial intelligence (AI). Such inventions replicate certain types of human cognitive activity, particularly interpretation, evaluation, and decision-making. On one hand, there is a strengthening of artificial intelligence's position, which contributes to technological progress. On the other hand, artificial intelligence creates a new level of risks for social relations and opens new horizons not only in science and industry but also in the legal realm. In particular, the high adaptability and general accessibility of AI provide a basis for the "involvement" of these technologies in various forms of criminal activity, up to and including the independent commission of specific criminal acts by robots. As this autonomy develops, the areas in which crimes are committed with their participation will inevitably expand. New challenges and risks arise that require the development of appropriate legal norms to minimize them without hindering technological advancement. Traditional legal systems often struggle to keep pace with the rapid development of technologies, creating gaps in regulating new realities. While some countries are attempting to adapt their laws to these new conditions, others are only beginning to recognize the necessity for such changes. It is crucial not only to define the legal status of AI but also to understand how we can protect human rights and ensure the ethical use of technology. Specifically, given the fact, the behavior of artificial intelligence entities can sometimes lead to harm to individual or collective interests protected by criminal law [1, p. 179]. Therefore, it can be recognized that there is a necessity for criminal law regulation of relations concerning the development, production, and use of artificial intelligence due to a whole range of criminal risks.

Machine learning methods can be used for cyberattacks, while facial recognition algorithms and data analysis can invade people's privacy. The ability of artificial intelligence systems to effectively search and process information, particularly personal information about health and psychological traits, makes them attractive to certain categories of offenders. The use of neural networks expands the capabilities of criminals, facilitating their activities; for example, instead of a person creating and disseminating false information, this task can be handed over to a neural network, which can accomplish it much faster thanks to deep fake technology. As a result of the development of artificial intelligence, the following questions have arisen:

- Who is responsible for actions carried out with the help of AI? If AI makes decisions that lead to unlawful actions or violations of the law, who should bear responsibility: the software developer, the system operator, or the AI itself? Can a "smart" robot become a subject of crime, including being an accomplice alongside a human?
- Does society require criminal law protection, when it comes to the use of artificial intelligence, and which actions need to be criminalized today?
- Another issue related to holding AI accountable is that it can be trained on data that contains biases and discrimination. For instance, if AI algorithms are used in policing, they may replicate and amplify systemic harassment and discrimination present in society, whom should we consider accountable in that case?
- Furthermore, in cases of crimes committed with the help of AI, there is the question of how to determine guilt and hold parties accountable. Can AI be considered guilty and subjected to punishment?

Unfortunately, as of now, Ukrainian legislators cannot provide answers to these questions, as there is effectively no regulation of AI. This is largely due to the lack of a specific definition for both the concept of AI and directly related phenomena, which is quite natural given the novelty of the technology and the variety of implementations and approaches to it in different countries [2, p. 12].

In 2017, the European Union proposed a comprehensive approach to defining current and prospective legislation regarding robotics, as articulated in the European Parliament Resolution on Robotics (European Parliament Resolution, 2017). This resolution delineates various applications of artificial intelligence (AI), addresses issues of accountability and ethics, and establishes fundamental behavioral guidelines for developers, operators, and manufacturers within the robotics sector. These guidelines are grounded in Isaac Asimov's three laws of robotics (1942).

In 2021, the European Commission published a draft regulation on artificial intelligence, which aims to establish harmonized rules in this domain. The EU presented a document encompassing a broad spectrum of problematic aspects related to the application of AI systems. These EU documents underscore the significance of upholding human rights and maintaining oversight over AI systems. Subsequently, the European Commission introduced the "White Paper on Artificial Intelligence," which proposes a risk-oriented approach to AI regulation. This White Paper emphasizes the necessity of respecting human rights and exercising control over AI systems.

In March 2024, the European Parliament approved legislation regulating the use of artificial intelligence (Artificial Intelligence Act) [3]. This initiative represents the first comprehensive set of regulations of its kind globally. The aim of this initiative is to "protect fundamental rights, democracy, the rule of law, and environmental sustainability from high-risk artificial intelligence, while also fostering innovation and establishing Europe as a leader in this field," as stated in the European Parliament's announcement. The Act stipulates that AI-based technologies will be categorized by risk levels: *unacceptable (in which case the technology will be banned), high, medium, and low.* Under the new law, certain practices will be prohibited, including real-time facial recognition (RBI), emotion recognition in workplaces and schools, social scoring, and AI technologies that manipulate human behavior or exploit vulnerabilities.

General Purpose AI Systems (GPAI) will be required to adhere to specific transparency standards. All images, audio recordings, or video footage processed using AI ("deepfakes") must carry appropriate labeling. European Parliament President Roberta Metsola stated that this legislation will promote innovation while simultaneously safeguarding the fundamental rights of EU citizens. "Artificial intelligence has already become an integral part of our daily lives. Now it is being regulated by our legislation."

In the United States, the National AI Initiative was established to strengthen and coordinate research, development, demonstration, and training in the field of AI across all departments and agencies after the U.S. Congress passed the National AI Initiative Act in January 2021. The law engaged numerous U.S. administrative agencies, including the Federal Trade Commission (FTC), the Department of Defense, the Department of Agriculture, the Department of Education, and the Department of Health and Human Services, and created new offices and working groups aimed at implementing a national strategy for legislation and regulation in the field of artificial intelligence. *The Algorithmic Accountability Act of 2022 noted that within the common law system, artificial intelligence, in its external manifestations, could correspond to such mandatory elements of a crime as mens rea and actus reus, and the model of "direct liability" may be applied in the future to its capabilities to perform actions that have characteristics similar to those of a crime.*

The proposed law will create rules requiring "covered entities," which include companies meeting certain criteria, to conduct impact assessments when using automated decision-making processes. This is a response to reports that artificial intelligence systems can lead to biased and discriminatory outcomes, particularly those created using AI or machine learning. The use of AI in the criminal justice system will be included in such a Bill of Rights. Additionally, to create a "voluntary risk management system for trustworthy AI systems," the National Institute of Standards and Technology (NIST), a division of the U.S. Department of Commerce, is collaborating with stakeholders. The final product of this initiative may be comparable to the voluntary regulatory framework proposed by the EU.

The Brazilian bill on AI provides a detailed description of the rights of users interacting with AI systems and offers guiding principles for categorizing different types of AI based on the risk they pose to society. It is important to note that this law clearly states that AI developers are required to conduct risk analyses before releasing an AI product to the market. The bill's focus on user rights places the responsibility for information about AI systems on the manufacturers. Users have the right to know that they are interacting with AI, as well as the right to an explanation of how the AI arrived at a particular decision or recommendation. Users can also challenge AI decisions or request human intervention, especially in cases where AI decisions may significantly impact them, such as in autonomous vehicle management systems, hiring criteria, credit assessments, or biometric identification. The highest risk category relates to any AI systems used in areas related to safety or that can affect human life and health; such systems are prohibited in principle. All AI developers are responsible for harm caused by their AI systems; however, developers of "high-risk" products must adhere to higher standards of accountability.

The "Next Generation Artificial Intelligence Development Plan" from China in 2017 states that generative AI models should reflect "the core values of socialism." In its current form, the draft regulations indicate that developers are "responsible" for the outcomes produced by their AI [4]. The rules also establish restrictions on the use of data for training, and developers bear legal responsibility if their data infringes on anyone's intellectual property. The rules require that AI services generate only "true and accurate" content. It is noteworthy that China has one of the most developed regulatory frameworks in the world regarding AI regulation, which complements existing laws related to deepfakes, recommendation algorithms, and data security, giving China an advantage over other countries that are beginning to write new laws from scratch. Additionally, the Chinese internet regulator announced restrictions on the use of facial recognition technology in August.

Ukraine is also moving in this direction: in October 2023, the Ministry of Digital Transformation presented a roadmap for AI regulation; individual initiatives for the development and application of AI are emerging at the local level as well. One of the goals of this strategy is stated as "building Ukraine's brand as a digital nation in the field of AI." Ukraine is already applying AI systems in medicine, education, business, and public administration. A particularly important area for AI use is military technology. In this regard, the government plans to create significantly more opportunities for both developers and users of AI. Ukraine's ambitions regarding the use and development of AI raise questions about proper legal and managerial regulation of these technologies. In 2020, the Cabinet of Ministers of Ukraine approved the Concept of Artificial Intelligence Development in Ukraine, which outlined the goals and main tasks of AI development. The main objectives of the Concept are to define the directions for AI development to meet the needs and interests of people, build a competitive economy, and improve public administration.

The Concept identified nine priority areas for the implementation of state policy in the field of AI, including education, science, economy, cybersecurity, defense, information security, state governance, legal regulation, and ethics. One of the priority directions for implementing the Concept is to "occupy leading positions in global ratings (AI Readiness Index by Oxford Insights, AI Index by Stanford University) by 2030."

In the field of justice, the main tasks of transformation were defined as:

- Further development of technologies already used in justice;
- Introduction of AI-based consulting programs;
- Conducting activities to resocialize convicted individuals using AI technologies;
- Issuing court decisions in cases of minor complexity based on results of analysis using AI.

In 2021, the Cabinet of Ministers approved the Plan of Actions for Implementing the Concept of Artificial Intelligence Development in Ukraine for 2021–2024. The plan includes measures to introduce legal regulation, information campaigns, scientific cooperation, and the implementation of AI technologies in various spheres, including justice.

One of the main formal shortcomings of the Concept and Plan is that they do not provide for conducting wide consultations or discussions with stakeholders in those areas where AI technologies will be introduced (in justice, this includes judges, prosecutors, lawyers, etc.). Such discussions could have formed a more realistic view of needs, opportunities, boundaries, and risks of implementing AI in these spheres and enabled further development based on specific data and context.

Currently, there are two main strategies for developing legal regulation:

- Developing a national legal framework;
- Implementing (full or gradual) the EU AI Act (EU AI Regulation).

To adapt European law to the needs and challenges of the AI industry in Ukraine, active cooperation with regulatory bodies is necessary. The introduction of European legislation will undoubtedly face technical and institutional challenges, but European legal and regulatory standards in this area should serve as a reference point for Ukraine, given that Ukraine's course on European integration is enshrined in its Constitution.

A thorough examination of the current domestic and foreign doctrine on artificial intelligence has led to the conclusion that AI refers to a field of computer science that studies the development of software and hardware systems capable of performing tasks that typically require human intelligence. AI can encompass technologies such as machine learning, image recognition, natural language processing, and autonomous systems. As a result, the system not only operates strictly according to the algorithm designed by its creator, but also has the ability to modify this algorithm within certain boundaries to optimize decision-making.

After studying the concept of artificial intelligence and its significance, it is necessary to enumerate the possible types of AI development. In information technology, depending on complexity, three levels of AI development are identified: 1) narrow artificial intelligence (Artificial Narrow Intelligence), focused on specific tasks or a limited range of tasks; 2) general artificial intelligence (Artificial General Intelligence), universal and capable of solving a wide range of tasks at the level of human intelligence; 3) superintelligence (Artificial Superintelligence), exceeding the intellectual abilities of both an individual human and humanity as a whole.

In terms of perception of the surrounding environment, AI systems can be divided into four types. The first type: reactive systems (capable of reacting to the surrounding environment). The second type: systems with limited memory (correcting behavior based on experience). The third type: intelligent systems (capable of recognizing thoughts and emotions). The fourth type: systems with artificial self-awareness (capable of forming self-representations and having cognitive abilities at the level of human beings).

Such a division suggests a predictable growth in the capabilities of AI. Therefore, today's recommendations by the European Parliament emphasize that robot autonomy introduces new aspects in determining their status – whether they can be considered as physical persons, legal entities, animals, or other subjects of rights, or if a new category with unique characteristics and consequences for the distribution of rights and responsibilities is necessary.

In the context of criminal law, the characterization of any criminal offense involves the examination of its subjective features (*subject and subjective side*). Crimes committed using AI systems are no exception. The subject is one of the four essential elements of a crime, without which any consideration of criminal responsibility becomes meaningless.

In Ukraine's Criminal Code, the concept of "subject" is defined in Section IV of the General Part, which is titled "Person liable to criminal responsibility (subject)". According to Article 18, Section 1 of the Criminal Code, the subject of a crime is a physical person who has committed the crime at an age where, according to this Code, criminal responsibility may be established [5].

The proliferation of AI systems raises questions about defining the subject of a criminal offense committed by such a system. It is essential to note that AI system activity is still linked to human beings and is subject to their control (either directly or indirectly), and therefore, AI often serves as a means of inflicting significant harm [6, p. 113].

However, as autonomous AI systems become more prevalent, it becomes increasingly difficult to consider them solely as instruments in the hands of other subjects (manufacturers, operators, owners, or users). This, in turn, may lead to the commission of socially dangerous acts that have severe consequences. Therefore, it is possible to identify several types of situations where AI may commit actions with characteristics of criminal offenses:

- A mistake was made in the creation of an AI algorithm or certain parts of it, which resulted in the commission of a criminal offense. A variant of this situation is cases where an AI system "went out of control" due to the failure to consider certain factors during algorithm development;
- The adoption of an AI system capable of self-learning, which made decisions that can be qualified as criminal-contrary;
- An AI system suffered external interference, as a result of which a criminal offense was committed;
- An AI system was intentionally created with the goal of committing a criminal offense.

By identifying these potential scenarios, it is possible to develop a more comprehensive understanding of the role of AI in criminal justice and to address the challenges posed by this emerging technology.

In 3rd and 4th scenarios, the use of AI as a tool for perpetrating criminal offenses is effectively implied. The individual responsible for using AI for illegal purposes would be liable for punishment.

However, the situation where an error occurs, leading to criminal consequences, is particularly problematic. Traditional criminal responsibility assumes the presence of guilt and awareness of committing a crime. In the case of AI, determining responsibility is challenging due to its algorithm-based operation and lack of consciousness [7, pp. 269–279].

For instance, in 2018, a self-driving car developed by Uber Technologies Inc. struck a girl in Arizona due to a programming error [8, pp. 412–436]. Similarly, in 2019, a self-driving bus called Navia collided with a pedestrian in Vienna due to an algorithmic error. In these cases, the company took responsibility and stated:

- The manufacturer is accountable if an accident occurs due to a software malfunction. For example, Uber stated that they would be responsible for accidents involving their self-driving cars;
- The owner of the vehicle is accountable if the self-driving car is not insured, and the level of automation is low.

This approach is entirely justified and may establish a precedent for addressing similar cases in common law countries. Furthermore, this example highlights the need to include the following individuals in the list of those responsible for crimes related to AI: the creator of the AI model; the manufacturer of AI systems; the seller of products equipped with AI; the user of products equipped with AI; other physical persons involved in AI usage and/or invention. Even if developers, manufacturers, and users are not fully aware of their responsibility, the fact that they deploy systems they do not understand or control cannot excuse them from liability [9].

The emergence of the second scenario suggests that, as predicted by the founders of AI, such as Alan Turing, future AI systems will possess two primary attributes. Firstly, AI will be capable of accumulating experience and learning from it. Secondly, AI will be able to operate independently of human intervention and make individual decisions autonomously. In essence, such a subject of AI would possess cognitive abilities, enabling it to select between alternative possible solutions to problems [10, pp. 58–67].

Therefore, it can be argued that if such an AI system commits a criminal offense, it may be held liable for criminal responsibility. In this context, legislators should develop a new system of penalties that can be applied to artificial intelligence. Considering the impossibility of applying traditional imprisonment to software and hardware systems, an alternative approach would be to temporarily disable the AI system for a specified period. During this time, the AI system would be restricted and temporarily deprived of its virtual freedom. This approach would allow for the development of a new framework for holding AI systems accountable for their actions, which would be essential in addressing the ethical and legal implications of autonomous AI systems.

It is evident that the need for regulating AI through criminal law is becoming increasingly pressing. However, it is also important to acknowledge that AI will, in turn, have an impact on the criminal legal sphere. For instance, the development of autonomous vehicles in the near future will require criminal law regulation of the use of AI systems in the field of traffic safety, including the creation of a new criminal offense – interference with autonomous vehicle software.

Furthermore, AI systems may be involved in the commission of crimes such as causing harm to health or life, invasion of privacy, illegal disclosure of protected information, fraud, traffic violations, and inadequate public transportation services, as well as terrorist acts and computer-related crimes.

At this stage of human kind development and technological advancement, it is premature to attribute elements of legal subjectivity to AI. However, domestic and international lawmakers will need to confront the changing legal status of AI, which requires departing from the traditional understanding of "subject" and recognizing AI as a subject of crime. The recognition of AI as a subject of crime would effectively necessitate defining its rights and obligations and ensuring their guarantee and protection by the state. The decision to recognize AI as a subject of law ultimately depends on constitutional law rather than criminal law. Therefore, this process will require reviewing laws and changing legal culture.

However, it is clear that society must already anticipate potential risks and develop a concept of responsibility, including criminal liability, when creating and using AI. This will be facilitated by the expansion of judicial practice in this area, which will likely lead to the adaptation of legislation to new realities and the creation of new precedents in countries with common law systems that address the questions raised in this work.

REFERENCES

1. Hallevy G. The Criminal Liability of Artificial Intelligence Entities – from Science Fiction to Legal Social Control. *4 Akron Intellectual Property Journal.* 2010. № 12. P. 179. URL: http://dx.doi.org/10.2139/ssrn.1564096. (Last accessed: 09.06.2024).
2. Rofi Aulia Rahman, Rizki Habibulah. The criminal liability of artificial intelligence: is it plausible to hitherto indonesian criminal system. *Journal Ilmiah Hukum*. 2019. № 4. P. 12.
3. Artificial intelligence act : Regulation (EU) of European Parliament and of the Council of 13.06.2024. № 2024/1689. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32024R1689. (Last accessed: 09.06.2024).
4. The Next Generation Artificial Intelligence Development Plan of China and Latest Progress 中国新一代人工智能发展规划和最新进展: Government policy of the China's State Council of 08. 07.2017. №1741. URL: https://www.gov.cn/xinwen/2017-07/20/content_5212064.htm. (Last accessed: 09.06.2024).
5. Кримінальний Кодекс України : Закон України від 05.04.2001. № 2341-III. URL:https://zakon.rada.gov.ua/go/2341-14 (дата звернення: 09.06.2024).
6. Mindaugas Naucius. Should Fully Autonomous Artificial Intelligence Systems Be Granted Legal Capacity. *Teisės apžvalga*. 2018. № 1. P. 113. URL: https://www.vdu.lt/cris/bitstream/20.500.12259/36040/1/ISSN2029-4239_2018_N_1_17.PG_113-132.pdf. (Last accessed: 09.06.2024).
7. Kingston J.K. Artificial Intelligence and Legal Liability. *Research and Development in Intelligent Systems XXXIII :* conference paper of International conference Incorporating Applications and Innovations in Intelligent Systems. Cambridge: SEG Publish, 2016. Pp. 269–279.
8. Gless S., Silverman E., Weigend T. If Robots cause harm, Who is to blame? Self-driving Cars and Criminal Liability. *New Criminal Law Review: In International and Interdisciplinary Journal*. 2016. №. 19. Pp. 412–436.
9. Радутний О.Е. Кримінальна відповідальність юридичної особи стане кроком до закріплення віртуальності життєвого простору. *Електронне наукове фахове видання Національного університету «Юридична Академія України ім. Ярослава Мудрого»*. 2011. № 1. URL: http://nauka.jur-academy.kharkov.ua (дата звернення: 09.06.2024).
10. Радутний О.Е. Стан інформаційно-законодавчої діяльності на прикладі Кримінального кодексу України. *Інформація і право.* 2016. № 3(18). С. 58–67.