ПРИМЕНЕНИЕ НОРМ ПРАВА ОРГАНАМИ ГОСУДАРСТВЕННОЙ ВЛАСТИ

THE LAW ENFORCEMENT BY THE PUBLIC AUTHORITIES

УДК 343.4
DOI 10.52468/2542-1514.2025.9(4).36-46

ПРАВОВЫЕ АСПЕКТЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ОБЕСПЕЧЕНИИ НАЦИОНАЛЬНОЙ БЕЗОПАСНОСТИ: ВЫЗОВЫ И УГРОЗЫ*

А.К. Жарова
Институт государства и права Российской академии наук, г. Москва, Россия

Информация о статье
Дата поступления –
9 марта 2025 г.
Дата принятия в печать –
20 сентября 2025 г.
Дата онлайн-размещения –
20 декабря 2025 г.

Ключевые слова
Доверенные технологии, искусственный интеллект, принципы, национальная безопасность, информационная безопасность, недискриминация, объективность, этические нормы, система оценки соответствия

Исследуется использование искусственного интеллекта (ИИ) в обеспечении национальной безопасности с позиций правового регулирования. Отмечается, что обеспечение национальной безопасности зависит от алгоритмов, реализованных в технологиях ИИ. Применение ИИ открывает широкие возможности для укрепления национальной безопасности, но также вызывает ряд правовых и этических проблем, которые требуют анализа и формирования методики их решения. Сделаны выводы о том, что в областях, где существует угроза национальной безопасности, на законодательном уровне определено требование о применении доверенных технологий ИИ. Эти технологии определяются как соответствующие стандартам безопасности и разработанные на основе принципов объективности, недискриминации и этичности. Однако в законодательстве и документах технического регулирования раскрыты только два принципа – недискриминации и объективности, и несмотря на то, что в законодательстве в области ИИ используется термин «этичность норм», этичность не определена как самостоятельный принцип.

LEGAL ASPECTS OF ARTIFICIAL INTELLIGENCE IN ENSURING NATIONAL SECURITY: CHALLENGES AND THREATS**

Anna K. Zharova
Institute of State and Law of the Russian Academy of Sciences, Moscow, Russia

Article info
Received –
2025 March 9
Accepted –
2025 September 20

The subject. Artificial intelligence (AI) opens up wide opportunities for strengthening national security, but also raises a number of legal and ethical issues that need to be carefully analyzed and resolved. The implementation of a government strategy aimed at improving tools and approaches to information protection using AI is a key aspect of ensuring national security and maintaining law and order.

* Публикация подготовлена в рамках научного проекта «Создание российской историографической модели политико-правовых знаний и ее применение для разработки перспективных средств противодействия идеологическим искажениям цивилизационного развития России», осуществляемого Федеральным государственным бюджетным учреждением науки «Институт государства и права Российской академии наук» при финансовой поддержке Министерства науки и высшего образования Российской Федерации (Соглашение от 12 июля 2024 г. № 075-15-2024-639).

** The publication was prepared within the framework of the scientific project "Creating a Russian historiographical model of political and legal knowledge and its application to develop prospecting means of countering ideological distortions of Russia's civilizational development", implemented by the Federal State Budgetary Institution of Science “Institute of State and Law of the Russian Academy of Sciences”, with financial support from the Ministry of Science and Higher Education of the Russian Federation (Agreement dated July 12, 2024 No. 075-15-2024-639).

ISSN 2658-4050 (Online)

Available online –
2025 December 20

Keywords
Trusted technologies, artificial intelligence, principles, national security, information security, non-discrimination, objectivity, ethical standards, conformity assessment system

The aim of the article was to confirm the hypothesis that ensuring national security depends on algorithms implemented in artificial intelligence (AI) technologies.
Methodology. The criteria formulated in the President Decree of the Russian Federation No. 124 are fundamental on the basis of which trusted AI systems are developed. The analysis of legal acts, acts of technical regulation and scientific literature was made.
Main results. In areas where there is a threat to national security, the use of trusted AI technologies is becoming mandatory. In accordance with Presidential Decree No. 124, trusted AI technologies are defined as meeting safety standards and developed based on the principles of objectivity, non-discrimination and ethics. Their use should exclude the possibility of harming a person, violating his rights and freedoms, as well as harming the interests of society and the state. Only two principles formulated in that Decree are disclosed – this is the principle of non-discrimination and objectivity. Despite the use of the term "ethics of norms", "ethics" in the legislation in the field of AI, ethics is not defined as an independent principle. Ethics in the field of AI can be understood as a system of views and principles aimed at protecting the moral and social consequences associated with the use of various data processing models contained in AI technologies. The principle of objectivity or impartiality is related to the ability of the system to inspire confidence and exclude unjustified bias of the formed estimates. The principle of objectivity is closely intertwined with the principles of transparency of AI systems and data protection and security. To achieve the objectivity of data processing by AI systems, it is necessary to use high-quality and representative datasets obtained from reliable sources. The implementation of the principle of non-discrimination contains certain contradictions that manifest themselves in the conditions of ensuring national security. Since, despite the fact that it is intended to exclude the possibility of using AI algorithms for processing and analyzing data containing discriminatory criteria, if necessary, the state can apply other criteria to ensure the national security.
Conclusions. The implementation of the above principles, the creation of a trusted environment for the use of AI technologies, and the development of trusted AI technologies all ensure information security, the effectiveness of the system for protecting the interests of citizens, society, information, and information infrastructure, thereby ensuring national security.

1. Введение

Искусственный интеллект (далее – ИИ) активно внедряется во все сферы нашей жизни. Его использование открывает большие перспективы для укрепления национальной безопасности (далее – НБ), но в то же время ставит перед нами новые правовые и этические вопросы, которые требуют осмысления и решения.

Высказанная И.Л. Бачило более десяти лет назад идея о том, что не только граждане и государство нуждаются в защите от информации, но и сама информация требует защиты [1], актуальна и в контексте отношений, возникающих при использовании технологии ИИ. Необходимо не только обеспечить информационную безопасность (далее – ИБ) пользователей ИИ, но и защиту самого ИИ от его противоправного применения [2], так он как обладает огромным потенциалом [3; 4]. Например, чтобы обеспечить безопасность ИИ, были разработаны стандарты защиты данных при обработке с помощью технологий ИИ. Кроме того, были сформулированы политики использования и требования, которым должен соответствовать ИИ, чтобы гарантировать ответственное применение этих технологий [5; 6].

Ключевой целью совершенствования законодательства в сфере развития и применения технологий ИИ на период до 2030 г. является создание в России благоприятных правовых условий для разработки, внедрения и использования этих технологий и основанных на них решений. Так, с 1 июля 2020 г. в Москве проводится эксперимент по установлению экспериментального правового режима в целях создания необходимых условий для разработки и внедрения технологий ИИ сроком на пять лет[1].

[1] Федеральный закон от 24 апреля 2020 г. № 123-ФЗ «О проведении эксперимента по установлению специального регулирования в целях создания необходимых условий для разработки и внедрения технологий ИИ в субъекте Российской Федерации – городе федерального значения Москве и внесении изменений в статьи 6 и 10 Федерального закона “О персональных данных”» // Собрание законодательства Российской Федерации. 2020. № 17. Ст. 2701.

ISSN 2542-1514 (Print)

При этом разработка технологий ИИ не должна осуществляться в отрыве от правовых принципов, обеспечивающих защиту прав и свобод человека, а также безопасность Российской Федерации [7; 8]. В таком случае закономерен вопрос: как этапы разработки и использования ИИ связаны с обеспечением НБ?

2. Национальная безопасность и искусственный интеллект

В 2019 г. на своей большой пресс-конференции В.В. Путин заявил, что развитие ИИ является вопросом НБ[2]. ИИ – это «комплекс технологических решений, имитирующий когнитивные функции человека (включая самообучение и поиск решений без заранее заданного алгоритма) и позволяющий при выполнении задач достигать результаты, как минимум сопоставимые с результатами интеллектуальной деятельности человека»[3]. Поскольку ИИ – это комплекс технологических решений, связь между ИИ и НБ кроется в этапах разработки и применения ИИ, т. е. на протяжении всего жизненного цикла технологических решений ИИ. Вопросами обеспечения безопасности личности, общества и государства от внутренних и внешних информационных угроз занимается ИБ[4]. Соответственно, технологии ИИ связаны с обеспечением НБ через обеспечение ИБ, которое является ключевым аспектом обеспечения НБ. Этот тезис подтверждается многочисленными научными исследованиями (см., напр.: [9–11]), законодательством, а также практиками, которые предложили объединить эти понятия и сформировать понятие «национальная информационная безопасность»[5]. Тем самым практики подчеркивают важность обеспечения ИБ в целях достижения НБ.

В 2021 г. в соответствии со Стратегией национальной безопасности Российской Федерации (далее – Стратегия НБ)[6] ИБ отнесена к национальному приоритету. Причем, как указано в Стратегии НБ, реализация государственной политики, направленной на совершенствование средств и методов обеспечения ИБ на основе применения технологий ИИ, является одной из задач, решение которой позволит достичь цели обеспечения ИБ.

В Национальной стратегии развития искусственного интеллекта на период до 2030 г. (далее – Стратегия ИИ)[7] среди прочих указана такая цель развития ИИ, как обеспечение НБ и правопорядка.

Указом Президента РФ от 15 февраля 2024 г. № 124 «О внесении изменений в Указ Президента Российской Федерации от 10 октября 2019 г. № 490 “О развитии искусственного интеллекта в Российской Федерации” и в Национальную стратегию, утвержденную этим Указом» (далее – Указ Президента РФ № 124)[8] определена необходимость не только внесения изменений в законодательство, обеспечивающее регулирование взаимодействия человека с технологиями ИИ в целях стимулирования активного развития и использования этих технологий, но и формирования «перечня областей использования технологий ИИ, в которых может быть нанесен ущерб безопасности Российской Федерации».

Таким образом, исходя из анализа положений законодательства, очевиден вывод, что реализация государственной политики, направленной на совершенствование средств и методов обеспечения ИБ с использованием технологий ИИ, в свою очередь является ключевой целью обеспечения НБ и правопорядка.

Причем государство определило наиболее чувствительные области использования технологий ИИ, в которых факт их применения может представлять опасность, в том числе для НБ. Этот вывод подтверждается разработанным Министерством экономического развития РФ и утвержденным его Приказом от 26 ноября 2024 г. № 752 (далее – приказ Минэкономразвития России № 752) регламентом обеспечения компенсации вреда, причиненного в результате

[2] Самтынова Е. Владимир Путин: «Искусственный интеллект – вопрос национальной безопасности» // Гарант.ру. 2019. 19 дек. URL: https://www.garant.ru/news/1310389/.

[3] Дорожная карта развития «сквозной» цифровой технологии «Нейротехнологии и искусственный интеллект» // СПС «КонсультантПлюс».

[4] Указ Президента РФ от 5 декабря 2016 г. № 646 «Об утверждении Доктрины информационной безопасности Российской Федерации» // Собрание законодательства Российской Федерации. 2016. № 50. Ст. 7074.

[5] Информационная безопасность в системе национальной безопасности // Falcongaze. 2024. 18 дек. URL: https://falcongaze.com/ru/pressroom/publications/informacionnaya-bezopasnost-v-otraslyah/informacionnaya-bezopasnost-v-sisteme-nacionalnoj-bezopasnosti.html.

[6] Указ Президента РФ от 2 июля 2021 г. № 400 «О Стратегии национальной безопасности Российской Федерации» // Собрание законодательства Российской Федерации. 2021. № 27 (ч. II). Ст. 5351.

[7] Указ Президента РФ от 10 октября 2019 г. № 490 «О развитии искусственного интеллекта в Российской Федерации» // Собрание законодательства Российской Федерации. 2019. № 41. Ст. 5700.

[8] Собрание законодательства Российской Федерации. 2024. № 8. Ст. 1102.

ISSN 2658-4050 (Online)

применения технологий ИИ в рамках проведения экспериментальных правовых режимов[9].

На основании анализа положений приказа Минэкономразвития России № 752 мы можем сделать вывод, что наиболее уязвимыми перед применением ИИ сферами общественной жизни являются личная жизнь, здоровье и имущество людей, а также имущество юридических лиц. В случае «причинения вреда жизни, здоровью или имуществу человека либо имуществу юридического лица при реализации экспериментального правового режима в результате использования решений, разработанных с применением технологий ИИ» эти ситуации будут рассматриваться на заседании комиссии для установления обстоятельств, при которых при реализации экспериментального правового режима в сфере цифровых инноваций в результате использования решений, разработанных с применением технологий ИИ, причинен вред жизни, здоровью или имуществу человека либо имуществу юридического лица.

Таким образом, анализ правовых актов приводит нас к выводу о том, что уровень уязвимости различных сфер общественной жизни и, как следствие, НБ и правопорядка напрямую зависит от того, в чьих руках находятся технологии ИИ, какие методы обработки данных реализованы в этих технологиях.

В Указе Президента РФ № 124 отмечается что «распространение соответствующих этических норм» является необходимым для достижения цели «создания в Российской Федерации благоприятных нормативно-правовых условий для разработки, внедрения и использования технологий ИИ и решений, разработанных на их основе, с учетом обеспечения защиты прав и свобод человека и безопасности Российской Федерации» (п. 51 (9)).

Но само использование термина «этические нормы» в вопросах применения технологи ИИ вызывает дискуссии [12–14]. Так, Л.Ю. Василевская считает, что точка зрения некоторых авторов о согласии с тем, что возможны этические правила взаимодействия человека с ИИ, юридически некорректна [15]. С одной стороны, мы согласны с точкой зрения Л.Ю. Василевской, поскольку ИИ – это совокупность технологий. С другой стороны, если исходить из понятия этики как «философской науки, объектом изучения которой является мораль, нравственность как форма общественного сознания и как вид общественных отношений»[10], то этические нормы можно рассматривать как связь между нравственностью и развитием технологий. Как отмечает Ф. Ларюэль, такое «соотношение не является простым. Нередко этичность трактуется так: развертывание науки и техники рождает новые ценностные установки и преобразует прежние» [16, с. 49].

Итак, учитывая использование термина «этичные нормы» в Указе Президента РФ № 124, мы будем рассматривать этичность как процесс формирования общественных ценностей у разработчиков и пользователей технологий ИИ на протяжении всего жизненного цикла технологий ИИ и систем ИИ.

Жизненный цикл системы ИИ в соответствии с ГОСТ Р 71476–2024[11] представлен на рисунке.

В Указе Президента РФ № 124, помимо «распространения соответствующих этических норм», отмечены следующие основные направления создания комплексной системы нормативно-правового регулирования общественных отношений в области развития и использования технологий ИИ, и обеспечения безопасности применения таких технологий: «формирование механизма оценки рисков нарушения этических норм при внедрении технологий ИИ в отраслях

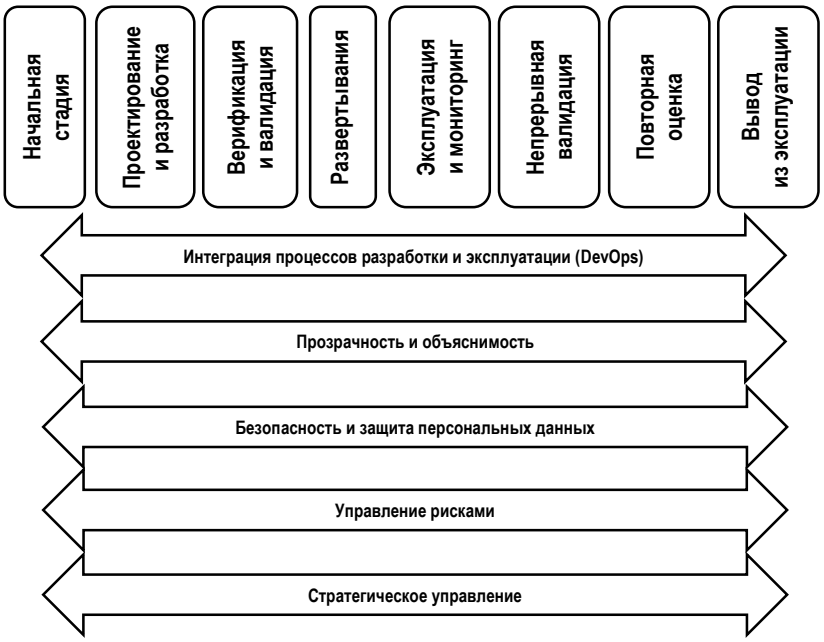
[9] Приказ Министерства экономического развития Российской Федерации от 26 ноября 2024 г. № 752 «Об утверждении Порядка формирования комиссии для установления обстоятельств, при которых при реализации экспериментального правового режима в сфере цифровых инноваций в результате использования решений, разработанных с применением технологий искусственного интеллекта, причинен вред жизни, здоровью или имуществу человека либо имуществу юридического лица, Порядка рассмотрения комиссией случаев причинения вреда жизни, здоровью или имуществу человека либо имуществу юридического лица при реализации экспериментального правового режима в сфере цифровых инноваций в результате использования решений, разработанных с применением технологий искусственного интеллекта, Порядка подготовки заключения комиссии по итогам установления обстоятельств, при которых в результате реализации экспериментального правового режима в сфере цифровых инноваций при использовании решений, разработанных с применением технологий искусственного интеллекта, был причинен вред жизни, здоровью или имуществу человека либо имуществу юридического лица, формы такого заключения» // Официальное опубликование правовых актов. № 0001202412270024.

[10] Этика // Словарь русского языка: в 4 т. / под ред. А.П. Евгеньевой. 4-е изд., стер. М.: Русский язык: Полиграфресурсы, 1999. Т. 4: С–Я. С. 769.

[11] ГОСТ Р 71476–2024 (ИСО/МЭК 22989:2022) Искусственный интеллект. Концепции и терминология искусственного интеллекта: утв. и введен в действие Приказом Росстандарта от 28 окт. 2024 г. № 1550-ст. М.: Рос. ин-т стандартизации, 2024. IV, 57 с.

ISSN 2542-1514 (Print)

экономики и социальной сферы; разработка требований ИБ в отношении технологий ИИ; создание системы оценки соответствия технологий ИИ требованиям законодательства Российской Федерации, в том числе в области ИБ» (пп. «ж»–«л» п. 51 (11)).



Жизненный цикл системы ИИ в соответствии с ГОСТ Р 71476–2024

Причем для тех областей применения технологий ИИ, в которых может быть нанесен ущерб НБ, разработка технологий ИИ должна осуществляться в соответствии с требованиями ИБ доверенных технологий ИИ (пп. «а» п. 51 (10) Указа Президента РФ № 124).

Создание системы оценки соответствия технологий ИИ требованиям в области ИБ связано с системой стандартизации и сертификации технологий ИИ, которая в конечном итоге позволяет разрабатывать доверенные технологии ИИ.

Доверие к технологиям ИИ, по нашему мнению, в первую очередь направлено на создание доверительного отношения пользователей таких технологий к этим технологиям и неразрывно связано с доверием к моделям обработки данных, реализованных в этих технологиях. Важность понимания пользователями процесса обработки данных связана с тем, что технология, имитирующая когнитивные функции человека, может создать иллюзию понимания смысла принимаемых ею решений [17; 18].

Однако это не так. Технологии ИИ на данном этапе развития алгоритмов ИИ не могут объяснить логику принятия своих решений, а также не обладают пониманием с точки зрения человека. Эта иллюзия приводит к тому, что люди неправильно воспринимают технологию ИИ и в результате используют ее не по назначению.

Хотя не только незнание функциональных возможностей ИИ может быть опасно. Понимание потенциала этих технологий и знание их возможностей также может стать причиной негативных последствий, например, если злоумышленники получают доступ к системам ИИ [19].

3. Доверенные технологии искусственного интеллекта

С 2021 г. по настоящее время в сфере ИИ произошли важные изменения, связанные с формированием перечня стандартов. На конец февраля 2025 г. на сайте Росстандарта опубликован 61 стандарт по направлению ИИ[11], из которых 32 приняты в 2022 г.

Исходя из ключевых задач, определенных Указом Президента РФ № 124, таких как обязательное использование доверенных технологий ИИ в областях, где существует угроза НБ, рассмотрим подроб-

[11] Стандарты по направлению «Искусственный интеллект» // РОССТАНДАРТ: федеральное агентство по техническому регулированию и метрологии. URL: https://www.rst.gov.ru/portal/gost/home/standarts/aistandarts.

нее критерии и методы оценки соответствия технологий ИИ, описанных в этих стандартах, требованиям законодательства Российской Федерации.

Под доверенными технологиями ИИ, в соответствии с Указом Президента РФ № 124, понимаются «технологии, отвечающие стандартам безопасности, разработанные с учетом принципов *объективности, недискриминации, этичности*, исключающие при их использовании возможность причинения вреда человеку и нарушения его основополагающих прав и свобод, нанесения ущерба интересам общества и государства» (пп. «ц» п. 5 Стратегии ИИ). Для разработки доверенных систем ИИ создана специальная платформа, позволяющая автоматизировать процессы выявления и устранения угроз, которые могут возникнуть на всех этапах жизненного цикла создания моделей машинного обучения[12].

В этом случае возникает закономерный вопрос: отражены ли в стандартах в области ИИ принципы объективности, недискриминации, этичности, а также другие важные принципы?

Так, в Российской Федерации в 2022 г. разработан Кодекс этики в сфере искусственного интеллекта (далее – Кодекс ИИ)[13]. Однако в самом Кодексе ИИ определения принципа этичности нет.

Судя по всему, включая термин «этика» в название Кодекса ИИ, его разработчики хотели показать, что этичность является основополагающей на всех этапах жизненного цикла технологий и систем ИИ. Однако, как мы уже подчеркивали, использование этого термина по отношению к ИИ вызывает дискуссии.

Несмотря на то, что Указ Президента РФ № 124 связывает разработку доверенных технологий с принципом этичности, мы не можем согласиться с такой постановкой вопроса. Для достижения этой цели необходимо четко определить, что подразумевается под принципом этичности.

Этичность в области ИИ можно рассматривать как систему взглядов и принципов, направленных на защиту моральных и социальных последствий, возникающих в процессе использования технологий ИИ. В этом случае она охватывает широкий круг вопросов, включая ответственность за ущерб, причиненный с помощью ИИ, а также исключение предвзятости и непрозрачности работы алгоритмов обработки данных. Кроме того, этичность затрагивает тонкие грани между обеспечением конфиденциальности и ИБ. Исходя из этого, мы считаем, что принцип социально ответственного поведения в сфере ИИ лучше всего отражает указанные отношения.

Авторами Кодекса ИИ раскрыт принцип социально ответственного поведения в области применения ИИ в связи с одной только областью – это применение технологий ИИ, в то время как этичность в области ИИ затрагивает, в том числе, отношения в области разработки технологий ИИ.

В соответствии с Кодексом ИИ принцип социально ответственного поведения в области применения ИИ определяет, что участники, занимающиеся ИИ, обязаны проявлять ответственность в отношении влияния систем ИИ на общество и граждан. Это включает защиту частной жизни, этичное, безопасное и ответственное использование данных. Участники рынка, занимающиеся ИИ, должны осознавать потенциальный ущерб, который может возникнуть в результате применения технологий и систем ИИ; ответственно подходить к выбору и использованию аппаратных средств и программного обеспечения, необходимых на различных этапах жизненного цикла систем ИИ (п. 2.2 ч. 2 раздела 1).

Таким образом, считаем, что требование социально ответственного поведения в области применения ИИ направлено на организацию доверенной среды пользователей к системам ИИ, поскольку предполагается, что участники рынка, занимающиеся ИИ, должны осознавать при выборе той или иной технологии все возможные риски и делать всё от них зависящее, чтобы не допустить реализации этих рисков [20].

Возвращаясь к двум другим принципам – объективности и недискриминации, также хотим отметить, что в Кодексе ИИ принцип объективности, как и принцип этичности, не имеет самостоятельного определения. В научной литературе данный принцип связывается с такими этапами жизненного цикла ИИ, как его разработка и использование, а также важности защиты данных, обрабатываемых этими технологиями [21; 22], этические последствия предоставления ИИ возможности принимать решения [23], как ценности и принципы, необходимые для обеспечения развития ИИ [24; 25]. Тем не менее в упомянутой литературе не представлено четкого определения данного принципа.

Зарубежные исследователи подчеркивают необходимость систематизации существующих исследований и понятийной категоризации [26].

[12] Платформа Доверенного Искусственного Интеллекта (ДИИ). URL: https://www.ispras.ru/technologies/tai/.

[13] Кодекс этики в сфере искусственного интеллекта // СПС «КонсультантПлюс».

ISSN 2542-1514 (Print)

В ГОСТ Р 59276–2020[14] объективность приравнивается к непредвзятости и связывается со способностью вызывать доверие при функционировании системы, исключающей необоснованное смещение формируемых оценок.

Близко связан с принципом объективности – принцип прозрачности технологии ИИ. Он определен, например, в ГОСТ Р ИСО 26000 и в Предварительном национальном стандарте 963–2024 (ИСО/МЭК 5339:2024)[15]. Прозрачность позволяет заинтересованным сторонам получать информацию о том, для чего предназначена система ИИ, каким образом она разрабатывалась и вводилась в действие. Сюда включена информация о целях, ограничениях, определениях, допущениях, алгоритмах, источниках и сборе данных, безопасности, защите частной информации и конфиденциальности и уровне автоматизации (п. 5.15.8 ГОСТ Р 71476–2024). Прослеживаемость отнесена к элементу прозрачности в качестве потенциального источника этических проблем.

Как видим, принцип прозрачности охватывает такой элемент жизненного цикла системы ИИ, как ее разработка. Иными словами, пользователь технологии ИИ может получить информацию от производителя о принципах обработки данных системами ИИ, методах принятия ею решений. Принцип прозрачности был расширен и дополнен концепцией объяснимости систем ИИ, направленной на обеспечение ясности и прозрачности решений, принимаемых системой ИИ и приложениями, когда с ними взаимодействуют заинтересованные стороны. В таком взаимодействии принципов (прозрачности и объективности) понимание пользователями ИИ ее работы становится более полным.

И, наконец, третий среди определенных в Указе Президента РФ № 124 признаков, связанных с разработкой доверенных технологий ИИ, – принцип недискриминации. В Кодексе ИИ он назван принципом равенства и отсутствия дискриминации. Он призван исключить возможность применения алгоритмов ИИ для обработки и анализа данных, содержащих дискриминационные критерии. В соответствии с Кодексом ИИ участники рынка ИИ (акторы) обязаны следить за тем, чтобы используемые ими алгоритмы и наборы данных, а также методы обработки информации для машинного обучения не приводили к умышленной дискриминации отдельных лиц или групп. При этом следует отметить, что явно заявленные акторами правила работы или использования систем ИИ для различных групп пользователей, разделенных по этим признакам, не могут считаться дискриминацией.

Однако следует отметить, что реализация принципа недискриминации в случае обеспечения НБ приводит к определенным противоречиям. С одной стороны, государство обязано обеспечивать безопасность всех своих граждан, независимо от их расы, этнической принадлежности, религии, пола, сексуальной ориентации или других признаков. С другой стороны, для эффективного противодействия угрозам НБ может возникнуть необходимость в применении критериев обработки данных, которые могут быть рассмотрены как дискриминационные или нарушающие принцип равенства. Но в случае выбора между обеспечением НБ и ограничением принципа равенства государство всегда сделает выбор в пользу НБ.

4. Заключение

Подводя итог проведенному исследованию, можно сделать следующие выводы. В сферах, где существует угроза НБ, использование доверенных технологий ИИ становится обязательным. В соответствии с Указом Президента РФ № 124 доверенные технологии ИИ определяются как отвечающие стандартам безопасности и разработанные на основе принципов объективности, недискриминации и этичности. Их применение должно исключать возможность причинения вреда человеку, нарушения его прав и свобод, а также ущерба интересам общества и государства.

Анализ положений законодательства и норм технического регулирования позволил заключить, что только два принципа, сформулированные в Указе Президента РФ № 124, раскрыты – это принцип недискриминации и объективности.

Несмотря на использование термина «этичность норм», «этичность» в законодательстве в области ИИ не определена в качестве самостоятельного принципа.

[14] ГОСТ Р 59276–2020 Системы искусственного интеллекта. Способы обеспечения доверия. Общие положения: утв. и введен в действие Приказом Росстандарта от 23 дек. 2020 г. № 1371-ст. М.: Стандартинформ, 2021. IV, 11 с.

[15] ПНСТ 963–2024 (ИСО/МЭК 5339:2024) Искусственный интеллект. Руководство для приложений на основе искусственного интеллекта: утв. и введен в действие Приказом Росстандарта от 25 окт. 2024 г. № 70-пнст. М.: Рос. ин-т стандартизации, 2024. IV, 27 с.

Анализ научной литературы позволил сделать вывод, что под этичностью в сфере ИИ мы можем понимать систему взглядов и принципов, направленных на защиту моральных и социальных последствий, связанных с применением различных моделей обработки данных, содержащихся в технологиях ИИ.

Принцип объективности, или непредвзятости, связан со способностью системы вызывать доверие и исключать необоснованное смещение формируемых оценок. Принцип объективности тесно переплетается с принципами прозрачности систем ИИ и защиты и безопасности обработки данных. Для достижения объективности обработки данных системами ИИ необходимо использовать качественные и репрезентативные наборы данных, полученные из надежных источников.

Реализация принципа недискриминации содержит определенные противоречия, которые проявляются в условиях обеспечения НБ, поскольку, несмотря на то, что он призван исключить возможность применения алгоритмов ИИ для обработки и анализа данных, содержащих дискриминационные критерии, при необходимости обеспечения НБ государство может применять иные критерии.

Реализация вышеуказанных принципов, создание доверенной среды пространства применения технологий ИИ, разработка доверенных технологий ИИ – всё это обеспечивает ИБ, эффективность системы защиты интересов, граждан, общества, информации и информационной инфраструктуры, тем самым – НБ.

СПИСОК ЛИТЕРАТУРЫ

1. Бачило И. Л. Правовое обеспечение информационной безопасности на новом этапе развития информационного общества / И. Л. Бачило. – URL: http://igpran.ru/public/articles/BachiloIL.2014.1..pdf.

2. Страхов А. А. Об утечке данных и DLP-системах / А. А. Страхов, Н. М. Дубинина // Криминологический журнал. – 2022. – № 4. – С. 226–232. – DOI: 10.24412/2687-0185-2022-4-226-232.

3. Громов Ю. Ю. Анализ существующих методов защиты информации в современных информационных системах / Ю. Ю. Громов, В. О. Драчев, В. В. Войтюк, В. В. Родин, Т. Г. Самхарадзе // Инженерная физика. – 2008. – № 2. – С. 67–69.

4. Жарова А. К. Предупреждение компьютерных атак типа man in the middle, совершаемых с использованием генеративного искусственного интеллекта / А. К. Жарова, В. М. Елин, Б. Р. Аветисян // Вопросы кибербезопасности. – 2024. – № 6 (64). – С. 28–41. – DOI: 10.21681/2311-3456-2024-6-28-41.

5. Denning P. J. The context problem in artificial intelligence / P. J. Denning, J. Arquilla // Communications of the ACM. – 2022. – Vol. 65, Iss. 12. – P. 18–21. – DOI: 10.1145/3567605.

6. Cheruvalath R. Artificial Intelligent Systems and Ethical Agency / R. Cheruvalath // Journal of Human Values. – 2023. – Vol. 29, No. 1. – P. 33–47. – DOI: 10.1177/09716858221119546.

7. Жарова А. К. Интеллектуальные системы распознавания образов и смысла в системе предупреждения преступлений / А. К. Жарова // Труды по интеллектуальной собственности. – 2024. – Т. 49, № 2. – С. 16–23. – DOI: 10.17323/tis.2024.21708.

8. Жарова А. К. Интеллектуальные системы распознавания образов и смысла в системе предупреждения правонарушений, совершаемых с использованием Сети / А. К. Жарова // Russian Journal of Economics and Law. – 2024. – Т. 18, № 2. – С. 469–480. – DOI: 10.21202/2782-2923.2024.2.469-480.

9. Цыгичко В. Н. Информационные вызовы национальной и международной безопасности / В. Н. Цыгичко, И. Ю. Алексеева ; под общ. ред. А. В. Федорова и В. Н. Цыгичко. – М. : Изд. ПИР-Центра полит. исслед., 2001. – 328 с.

10. Гулабян А. О. Основные направления защиты информации и ресурсов информационных систем / А. О. Гулабян, В. М. Смирнов // Тенденции развития науки и образования. – 2022. – № 88-1. – С. 45–46. – DOI: 10.18411/trnio-08-2022-13.

11. Рузанова Е. А. Средства защиты информации в ЛВС / Е. А. Рузанова // Теория и практика современной науки: взгляд молодежи : материалы III Всерос. науч.-практ. конф. на англ. яз. (Санкт-Петербург, 30 нояб. 2023 г.) : в 2 ч. – СПб. : С.-Петерб. гос. ун-т пром. технологий и дизайна, 2024. – Ч. II. – С. 166–169. – (На англ. яз.).

12. Ройзензон Г. В. Современные подходы формализации понятия этики в искусственном интеллекте / Г. В. Ройзензон // Труды международной конференции по компьютерной и когнитивной лингвистике TEL-2018 (Казань, 31 окт. – 3 нояб. 2018 г.) : в 2 т. – Казань : Акад. наук РТ, 2018. – Т. 1. – С. 306–331.

13. Галлезе-Нобиле К. Регулирование умных роботов и искусственного интеллекта в Европейском союзе / К. Галлезе-Нобиле // Journal of Digital Technologies and Law. – 2023. – Т. 1, № 1. – С. 33–61. – DOI: 10.21202/jdtl.2023.2.

14. Харитонова Ю. С. Правовые средства обеспечения принципа прозрачности искусственного интеллекта / Ю. С. Харитонова // Journal of Digital Technologies and Law. – 2023. – Т. 1, № 2. – С. 337–358. – DOI: 10.21202/jdtl.2023.14.

15. Василевская Л. Ю. Кодекс этики для «искусственного интеллекта»: юридический миф и реальность / Л. Ю. Василевская // Гражданское право. – 2023. – № 2. – С. 19–22.

16. Ларюэль Ф. Два этических начала в технологическом мире / Ф. Ларюэль ; пер. Е. Руднева // Философская антропология. – 2015. – Т. 1, № 1. – С. 49–61.

17. Жарова А. К. Достижение алгоритмической прозрачности и управление рисками информационной безопасности при принятии решений без вмешательства человека: правовые подходы / А. К. Жарова // Journal of Digital Technologies and Law. – 2023. – Т. 1, № 4. – С. 973–993. – DOI: 10.21202/jdtl.2023.42.

18. Бегишев И. Р. Цифровые и природоподобные технологии: особенности регулирования правом / И. Р. Бегишев, А. К. Жарова, М. В. Залоило, И. А. Филипова, А. А. Шутова // Journal of Digital Technologies and Law. – 2024. – Т. 2, № 3. – С. 493–499. – DOI: 10.21202/jdtl.2024.25.

19. Бегишев И. Р. Искусственный интеллект и уголовный закон / И. Р. Бегишев, З. И. Хисамова. – М. : Проспект, 2024. – 192 с.

20. Жарова А. К. Информационно-психологическое насилие в уголовном праве / А. К. Жарова // Информационное общество. – 2024. – № 2. – С. 103–111.

21. Халезов С. А. Этика искусственного интеллекта / С. А. Халезов // Наука. Техника. Человек: исторические, мировоззренческие и методологические проблемы : межвуз. сб. науч. работ. М. : МГТУ ГА, 2023. – Вып. 13. – С. 240–246.

22. Сущин М. А. Три этических вызова приложений искусственного интеллекта в области кибербезопасности (Реферат) / М. А. Сущин, М. Таддео // Этика науки : сб. обзоров и рефератов / ред.-сост. М. П. Булавинова. – М. : Ин-т науч. информации по обществен. наукам РАН, 2022. – С. 175–178.

23. Перевозчикова Д. А. Этика искусственного интеллекта / Д. А. Перевозчикова, М. Ю. Зарипова // Нанотехнологии: наука и производство. – 2024. – № 3. – С. 26–30.

24. Лизикова М. С. Этические и правовые вопросы развития искусственного интеллекта / М. С. Лизикова // Труды Института государства и права Российской академии наук. – 2022. – Т. 17, № 1. – С. 177–194. – DOI: 10.35427/2073-4522-2022-17-1-lizikova.

25. Dodonova V. Ethical Aspects of Artificial Intelligence Functioning in the XXIst Century / V. Dodonova, R. Dodonov, K. Gorbenko // Studia Universitatis Babeş-Bolyai. Philosophia. – 2023. – Vol. 68, No. 1. – P. 161–173.

26. What ethics can say on artificial intelligence: Insights from a systematic literature review / F. V. Giarmoleo, I. Ferrero, M. Rocchi, M. M. Pellegrini // Business and Society Review. – 2024. – Vol. 129, Iss. 2. – P. 258–292. – DOI: 10.1111/basr.12336.

REFERENCES

1. Bachilo I.L. *Legal provision of information security at a new stage in the development of the information society*. Available at: http://igpran.ru/public/articles/BachiloIL.2014.1..pdf. (In Russ.).

2. Strakhov A.A., Dubinina N.M. About data leakage and DLP systems. *Kriminologicheskii zhurnal = Criminological journal*, 2022, no. 4, pp. 226–232. DOI: 10.24412/2687-0185-2022-4-226-232. (In Russ.).

3. Gromov Yu.Yu., Drachev V.O., Vojtjuk V.V., Rodin V.V., Samkharadzce T.G. Analysis of existing methods of information protection in modern information systems. *Inzhenernaya fizika = Engineering physics*, 2008, no. 2, pp. 67–69. (In Russ.).

4. Zharova A.K., Elin V.M., Avetisyan B.R. Prevention of computer attacks such as man in the middle, committed using generative artificial intelligence. *Voprosy kiberbezopasnosti = Cybersecuriry issues*, 2024, no. 6 (64), pp. 28–41. DOI: 10.21681/2311-3456-2024-6-28-41. (In Russ.).

5. Denning P.J., Arquilla J. The context problem in artificial intelligence. *Communications of the ACM*, 2022, vol. 65, iss. 12, pp. 18–21. DOI: 10.1145/3567605.

6. Cheruvalath R. Artificial Intelligent Systems and Ethical Agency. *Journal of Human Values*, 2023, vol. 29, no. 1, pp. 33–47. DOI: 10.1177/09716858221119546.

7. Zharova A.K. The intelligent image and meaning recognition systems in the crime prevention system. *Trudy po intellektual'noi sobstvennosti = Works on Intellectual Property*, 2024, vol. 49, no. 2, pp. 16–23. DOI: 10.17323/tis.2024.21708. (In Russ.).

8. Zharova A.K. Intellectual systems of pattern and meaning recognition in the system of prevention of crimes committed using the Internet. *Russian Journal of Economics and Law*, 2024, vol. 18, no. 2, pp. 469–480. DOI: 10.21202/2782-2923.2024.2.469-480. (In Russ.).

9. Tsygichko V.N., Alekseeva I.Yu. *Information challenges to national and international security*, ed. by A.V. Fedorov and V.N. Tsygichko. Moscow, PIR Center for Political Studies Publ., 2001. 328 p. (In Russ.).

10. Gulabyan A.O., Smirnov V.M. The main directions of information protection and information systems resources. *Tendentsii razvitiya nauki i obrazovaniya*, 2022, no. 88-1, pp. 45–46. DOI: 10.18411/trnio-08-2022-13. (In Russ.).

11. Ruzanova E.A. Means of information protection in LAN, in: *Theory and practice of modern science: the view of youth*, Proceeding of the III All-Russian Scientific and Practical Conference in English (St. Petersburg, November 30, 2023), in 2 parts, St. Petersburg, Saint Petersburg State University of Industrial Technologies and Design Publ., 2024, pt. 2, pp. 166–169.

12. Roizenson G.V. Modern approaches of formalization of the concept of ethics in artificial intelligence, in: *Proceedings of the International Conference on Computational and Cognitive Linguistics TEL-2018* (Kazan, October 31 – November 03, 2018), in 2 volumes, Kazan, Academy of Sciences of the Republic of Tatarstan Publ., 2018, vol. 1, pp. 306–331. (In Russ.).

13. Gallese Nobile C. Regulating Smart Robots and Artificial Intelligence in the European Union. *Journal of Digital Technologies and Law*, 2023, vol. 1, no. 1, pp. 33–61. DOI: 10.21202/jdtl.2023.2.

14. Kharitonova Yu.S. Legal Means of Providing the Principle of Transparency of the Artificial Intelligence. *Journal of Digital Technologies and Law*, 2023, vol., 1, no. 2, pp. 337–358. DOI: 10.21202/jdtl.2023.14.

15. Vasilevskaya L.Yu. The code of ethics for artificial intelligence: a legal myth and reality. *Grazhdanskoe pravo = Civil law*, 2023, no. 2, pp. 19–22. (In Russ.).

16. Laruel F. Two ethical principles in the technological world, transl. by E. Rudneva. *Filosofskaya antropologiya*, 2015, vol. 1, no. 1, pp. 49–61. (In Russ.).

17. Zharova A.K. Achieving Algorithmic Transparency and Managing Risks of Data Security when Making Decisions without Human Interference: Legal Approaches. *Journal of Digital Technologies and Law*, 2023, vol. 1, no. 4, pp. 973–993. DOI: 10.21202/jdtl.2023.42.

18. Begishev I.R., Zharova A.K., Zaloilo M.V., Filipova I.A., Shutova A.A. Digital and Nature-like Technologies: Features of Legal Regulation. *Journal of Digital Technologies and Law*, 2024, vol. 2, no. 3, pp. 493–499. DOI: 10.21202/jdtl.2024.25.

19. Begishev I.R., Khisamova Z.I. *Artificial intelligence and criminal law*. Moscow, Prospekt Publ., 2024. 192 p. (In Russ.).

20. Zharova A.K. Informational and psychological violence in criminal law. *Informatsionnoe obshchestvo*, 2024, no. 2, pp. 103–111. (In Russ.).

21. Halezov S.A. Ethics of artificial intelligence, in: *Nauka. Tekhnika. Chelovek: istoricheskie, mirovozzrencheskie i metodologicheskie problemy*, Interuniversity collection of scientific papers, iss. 13, Moscow, Moscow State Technical University of Civil Aviation Publ., pp. 240–246. (In Russ.).

22. Sushchin M.A. Taddeo M. Three ethical challenges of artificial intelligence applications in the field of cybersecurity (Abstract), in: Bulavinov M.P. (ed. & comp.). *Etika nauki*, Collection of reviews and abstracts, Moscow, Institute of Scientific Information on Social Sciences of the Russian Academy of Sciences Publ., 2022, pp. 175–178. (In Russ.).

23. Perevozchikova D.A., Zaripova M.Y. Ethics of artificial intelligence. *Nanotekhnologii: nauka i proizvodstvo*, 2024, no. 3, pp. 26–30. (In Russ.).

24. Lizikova M.S. Ethical and legal issues of artificial intelligence development. *Trudy Instituta gosudarstva i prava Rossiiskoi akademii nauk = Proceedings of the Institute of State and Law of the RAS*, 2022, vol. 17, no. 1, pp. 177–194. DOI: 10.35427/2073-4522-2022-17-1-lizikova. (In Russ.).

25. Dodonova V., Dodonov R., Gorbenko K. Ethical Aspects of Artificial Intelligence Functioning in the XXIst Century. *Studia Universitatis Babeş-Bolyai. Philosophia*, 2023, vol. 68, no. 1, pp. 161–173.

26. Giarmoleo F.V., Ferrero I., Rocchi M., Pellegrini M.M. What ethics can say on artificial intelligence: Insights from a systematic literature review. *Business and Society Review*, 2024, vol. 129, iss. 2, pp. 258–292. DOI: 10.1111/basr.12336.

ИНФОРМАЦИЯ ОБ АВТОРЕ

Жарова Анна Константиновна – доктор юридических наук, доцент, ведущий научный сотрудник
Институт государства и права Российской академии наук
119019, Россия, г. Москва, ул. Знаменка, 10
E-mail: anna_jarova@mail.ru
ORCID: 0000-0002-2981-3369
ResearcherID: H-4012-2015
SPIN-код РИНЦ: 2240-1467

БИБЛИОГРАФИЧЕСКОЕ ОПИСАНИЕ СТАТЬИ

Жарова А.К. Правовые аспекты искусственного интеллекта в обеспечении национальной безопасности: вызовы и угрозы / А.К. Жарова // Правоприменение. – 2025. – Т. 9, № 4. – С. 36–46. – DOI: 10.52468/2542-1514.2025.9(4).36-46.

INFORMATION ABOUT AUTHOR

Anna K. Zharova – Doctor of Law, Associate Professor, Leading Researcher
Institute of State and Law of the Russian Academy of Sciences
10, Znamenka ul., Moscow, 119019, Russia
E-mail: anna_jarova@mail.ru
ORCID: 0000-0002-2981-3369
ResearcherID: H-4012-2015
RSCI SPIN-code: 2240-1467

BIBLIOGRAPHIC DESCRIPTION

Zharova A.K. Legal aspects of artificial intelligence in ensuring national security: challenges and threats. *Pravoprimenenie = Law Enforcement Review*, 2025, vol. 9, no. 4, pp. 36–46. DOI: 10.52468/2542-1514.2025.9(4).36-46. (In Russ.).